Medical genetics / edited by Victor A. McKusick and Robert Claiborne.

Contributors

McKusick, Victor A. 1921-2008.
Claiborne, Robert.

Publication/Creation

New York : HP Pub. Co., [1973]

Persistent URL

https://wellcomecollection.org/works/vexvfsfa

License and attribution

You have permission to make copies of this work under a Creative Commons, Attribution, Non-commercial license.

Non-commercial use includes private study, academic research, teaching, and other activities that are not primarily intended for, or directed towards, commercial advantage or private monetary compensation. See the Legal Code for further information.

Image source should be attributed as specified in the full catalogue record. If no source is given the image should be attributed to Wellcome Collection.

wellcome collection

Wellcome Collection
183 Euston Road
London NW1 2BE UK
T +44 (0)20 7611 8722
E library@wellcomecollection.org
https://wellcomecollection.org

MEDICAL GENETICS

VICTOR A. McKUSICK and ROBERT CLAIBORNE, EDITORS

A HOSPITAL PRACTICE® TEXT

MEDICAL GENETICS

VICTOR A. McKUSICK and ROBERT CLAIBORNE, EDITORS

CONTENTS

ISBN 0-913800-03-1

HP PUBLISHING COMPANY, INC. • PUBLISHERS
485 MADISON AVENUE, NEW YORK, N.Y. 10022

MEDICAL GENETICS

Edited by
VICTOR A. McKUSICK, M.D.
and
ROBERT CLAIBORNE

More than 15 million Americans have one or another type of genetic or congenital disorder, according to recent figures from the National Institutes of Health.

A significant portion of this surprising prevalence involves infants and children. Genetic and related conditions, in fact, account for from 20% to 40% of all pediatric hospitalizations.

New techniques of intrauterine intervention, cell culturing, immunogenetic study, etc., have produced new, practical genetic knowledge. This, in turn, has increased the scope and relevancy of counseling.

Through other research developments, a number of genetic diseases that had stubbornly resisted treatment are substantially correctable now, if not curable, and many more are showing promising responses experimentally.

Broader acceptance of abortion has also enhanced the ability to prevent genetic disease.

A group of 33 internationally prominent geneticists, many of whom have added significantly to the scope of the field, examine today's parameters of clinical genetics in this important book, which contains completely new material as well as updated versions of the articles first published in **Hospital Practice.**

(continued on back flap)

COVER ILLUSTRATION:
The artist's rendering by Philip Hays *shows a portion of the hemoglobin molecule, with one of its four amino acid chains (a beta). Differentiated in colors are eight altered residues, which together have been implicated in 11 different inherited hemoglobinopathies.*

WELLCOME MUSEUM OF MEDICAL SCIENCE,
183, EUSTON ROAD, LONDON, N.W.1.

GEN 297

MEDICAL GENETICS

It is an old experience
that, through her errors,
Nature often grants us
unexpected insights into
her secrets which are
otherwise a closed domain.

"Über Cystinurie"
A. Loewy, C. Neuberg (1904)

Adapted from

Hospital Practice

with illustrations by

Bunji Tagawa

Chapters 1, 2, 4, 6, 7, 11 (except pages specifically listed)
ALSO: drawings pages 200 and 207

Albert Miller

Chapters 3, 5, 8, 9, 10, 12, 13, 14, 15, 19, 20,
21, 23, 25, 26, 27 (except pages specifically listed)
ALSO: charts pages 66, 74, 77, 189, 192, and 235

Irving Geis

Chapter 18 (except pages specifically listed)
ALSO: chart page 172

Chapter 17	Nancy Lou Gahan
Chapter 22	Dan Bernstein
Chapter 24	Adele Spiegler
Pages 59, 151	Joan Dworkin
Page 125	Enid Kotschnig
Page 138	Zena Bernstein
Page 261	Ann Winclair
Page 266	Gaetano Di Palma

Data and Photo Sources

Page 56
Top photo, J. W. Harris, M.D.,
Proc Soc Exper Biol & Med 75:197, 1950

Page 59
Graph data from S. Charache et al, J Clin Invest 45:813, 1966
and J. Bonaventura and A. Riggs, J Biol Chem 243:980, 1968

Pages 126, 127
Chart data from Am J Med 47:3:395, 1969

Page 135
Radiographs, Martin Welch, M.D., et al, VA Hospital, Oklahoma City,
reprinted from Ann Intern Med 71:539, 1969

Page 156
Chart data from N Engl J Med 281:1385, 1969

Page 241
Photographs reprinted from
HEW Children's Bureau Publication #419

Pages 248, 249, 250
Drawings and chart data from
H. L. Nadler, M.D., J Pediatr 74:139, 1969

Designed by

Robert S. Herald

MEDICAL GENETICS

Edited by

Victor A. McKusick, M. D.

Professor of Medicine, Epidemiology and Biology,

and Chief, Division of Medical Genetics,

Johns Hopkins University, Baltimore

and

Robert Claiborne

Senior Editor, Hospital Practice

HP Publishing Co., Inc. • Publishers • New York

701032

First printing June 1973
Second printing January 1974

WELLCOME INSTITUTE LIBRARY
Coll welTROmec
Call No. QZ

MEDICAL GENETICS
This edition Copyright © 1973 by HP Publishing Co., Inc., New York, N.Y.
Chapters 4, 16, 17, 25, 26 © 1973
Chapters 3, 10, 14, 22, 27 © 1972
Chapters 9, 13, 15, 20, 21, 23 © 1971
Chapters 1, 2, 7, 8, 11, 12, 18, 19 © 1970
Chapters 5, 6, 24 © 1969

Printed in the United States of America

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means without written permission from the publisher:

HP Publishing Co., Inc.
485 Madison Avenue
New York, N.Y. 10022

ISBN 0-913800-03-1

Library of Congress Catalogue Card Number 73-78888

Table of Contents

Contributing Authors

IRWIN M. ARIAS — Professor of Medicine, Albert Einstein College of Medicine; Attending Physician, Bronx Municipal Hospital Center, The Bronx, N.Y.

FRITZ H. BACH — Professor, Department of Medical Genetics and Surgery, University of Wisconsin Medical School, Madison.

MARILYN L. BACH — Assistant Professor, Department of Pediatrics and Pharmacology, University of Wisconsin Medical School, Madison.

CEDRIC O. CARTER — Director, Medical Research Council's Clinical Genetics Unit; Honorary Consultant Geneticist, The Hospital for Sick Children, London, England.

TORBJÖRN CASPERSSON — Director, Institute for Medical Cell Research and Genetics, Medical Nobel Institute, Karolinska Institute, Stockholm, Sweden.

SAMUEL CHARACHE — Associate Professor of Medicine, Johns Hopkins University School of Medicine, Baltimore.

CYRIL A. CLARKE — Emeritus Professor of Medicine, University of Liverpool; Research Fellow, Nuffield Unit of Medical Genetics, Liverpool; and President, Royal College of Physicians, London, England.

C. LOCKARD CONLEY — Professor of Medicine, Johns Hopkins University School of Medicine, Baltimore.

JOSEPH DANCIS — Professor of Pediatrics, New York University School of Medicine, New York City; Career Investigator, National Institutes of Health, Bethesda.

MALCOLM A. FERGUSON-SMITH — Reader in Medical Genetics, University of Glasgow; Honorary Consultant in Medical Pediatrics, Royal Hospital for Sick Children, Glasgow, Scotland.

F. CLARKE FRASER — Professor of Medical Genetics, Department of Biology, McGill University; Director, Department of Medical Genetics, The Montreal Children's Hospital, Montreal, Canada.

JAMES GERMAN — Senior Investigator and Director of the Laboratory of Human Genetics, New York Blood Center, New York City.

ROBERT GUTHRIE — Research Professor of Pediatrics and Microbiology, State University of New York at Buffalo and Children's Hospital, Buffalo.

KURT HIRSCHHORN — Arthur J. and Nellie Z. Cohen Professor of Pediatrics and Genetics and Chief, Division of Medical Genetics, Department of Pediatrics, Mt. Sinai School of Medicine of the City University of New York.

NEIL A. HOLTZMAN — Associate Professor of Pediatrics and Director, Pediatric Genetics Unit, Johns Hopkins University School of Medicine, Baltimore.

R. RODNEY HOWELL — Professor of Pediatrics and Genetics and Director of the Program in Pediatrics, University of Texas Medical School at Houston; Pediatrician-in-Chief, Hermann Hospital, Houston.

DAVID YI-YUNG HSIA — At the time of his death, Dr. Hsia was Professor and Chairman, Department of Pediatrics, Stritch School of Medicine, Loyola University, Maywood, Ill.

MICHAEL M. KABACK — Associate Professor of Pediatrics, University of California at Los Angeles School of Medicine; Associate Chief, Division of Medical Genetics, Harbor General Hospital, Torrance, Calif.

ATTALLAH KAPPAS — Professor and Senior Physician to the Hospital of the Rockefeller University, New York City.

W. EUGENE KNOX — Professor of Biological Chemistry, Harvard Medical School at the New England Deaconess Hospital Cancer Research Institute, Boston; Assistant Physician and Senior Clinical Instructor in Medicine, Tufts University School of Medicine and The Boston Dispensary.

BERT N. LA DU, JR. — Professor and Chairman, Department of Pharmacology, New York University Medical School, New York City.

RICHARD D. LEVERE — Professor of Medicine, State University of New York College of Medicine, Brooklyn.

HUGH O. MCDEVITT — Professor of Medicine and Chief of the Division of Immunology, Stanford University School of Medicine.

VICTOR A. MCKUSICK — Professor of Medicine, Epidemiology, and Biology, and Chief, Division of Medical Genetics, Johns Hopkins University, Baltimore.

WILLIAM J. MELLMAN — Professor and Chairman, Department of Human Genetics, and Professor of Pediatrics, University of Pennsylvania School of Medicine, and Director of the Genetics Clinic, The Children's Hospital, Philadelphia.

ELIZABETH F. NEUFELD — Chief, Section of Human Biochemical Genetics, National Institute of Arthritis, Metabolism, and Digestive Diseases, National Institutes of Health, Bethesda.

JOHN S. O'BRIEN — Professor and Chairman, Department of Neurosciences, University of California at San Diego School of Medicine, La Jolla.

DAVID L. RIMOIN — Associate Professor of Pediatrics and Medicine, University of California at Los Angeles; Chief, Division of Medical Genetics, Harbor General Hospital, Torrance, Calif.

LEON E. ROSENBERG — Professor of Human Genetics, Pediatrics, and Medicine, and Chairman, Department of Human Genetics, Yale University School of Medicine, New Haven.

CHARLES R. SCRIVER — Professor of Pediatrics, Faculty of Medicine, and Associate Professor of Genetics, Biology Department, Faculty of Science, McGill University, Montreal, Canada.

J. EDWIN SEEGMILLER — Professor, Department of Medicine, Director, Division of Rheumatology, and Director, Human Biochemical Genetics Program, University of California at San Diego School of Medicine, La Jolla.

HARVEY L. SHARP — Associate Professor of Pediatrics, University of Minnesota Health Science Center, Minneapolis.

LORE ZECH — Associate Professor in Medical Cell Research and Genetics, Medical Nobel Institute, Karolinska Institute, Stockholm, Sweden.

Acknowledgments and Publisher's Note

The contents of this volume are, in major part, derived from an extended series of articles published in HOSPITAL PRACTICE between October 1969 and June 1973. Where necessary these articles have been brought up to date by the authors.

Because of their significant contribution to the content, it is appropriate that acknowledgment be made to various members of the HOSPITAL PRACTICE staff.

Certainly this text could not have been possible without the efforts of Gertrude Halpern, the Managing Editor of HOSPITAL PRACTICE, who effectively integrated the verbal and visual elements of all of the articles and served as an editorial consultant not only to Dr. McKusick and Mr. Claiborne, but also to many of the individual authors. Similarly, mention should be made of David W. Fisher, Executive Editor, of Irvin Block, Mal Schechter, and Amy Selwyn, Senior Editors, of Samuel Moffat, Contributing Editor, all of whom assisted in the preparation of particular articles, and of Charles Ryberg, Assistant Managing Editor, and Evelyn M. Gottlieb and Doris Ross, Copy Editors, who helped in the processing of the text.

Robert S. Herald, Art Director of HOSPITAL PRACTICE, is responsible for the design of the book. Among those who worked with him were Alfred P. Ingegno Jr., Assistant Art Director, Irwin Kuperberg and Lee Powers, Art Associates, and Stephanie Marcus and Gail Tavares.

Overall supervision of the publication of the book was provided by Herb Cornell, Special Projects Director of HOSPITAL PRACTICE, assisted by Katherine Bloch and Helen Draney. Angel Kuchinski, Helen Lazar, Steven B. Garrett, and Roslyn Yules also made important contributions to preparing the book for publication.

Finally, the index for the entire book was most skillfully and rapidly prepared by Ruth M. Kimmerer.

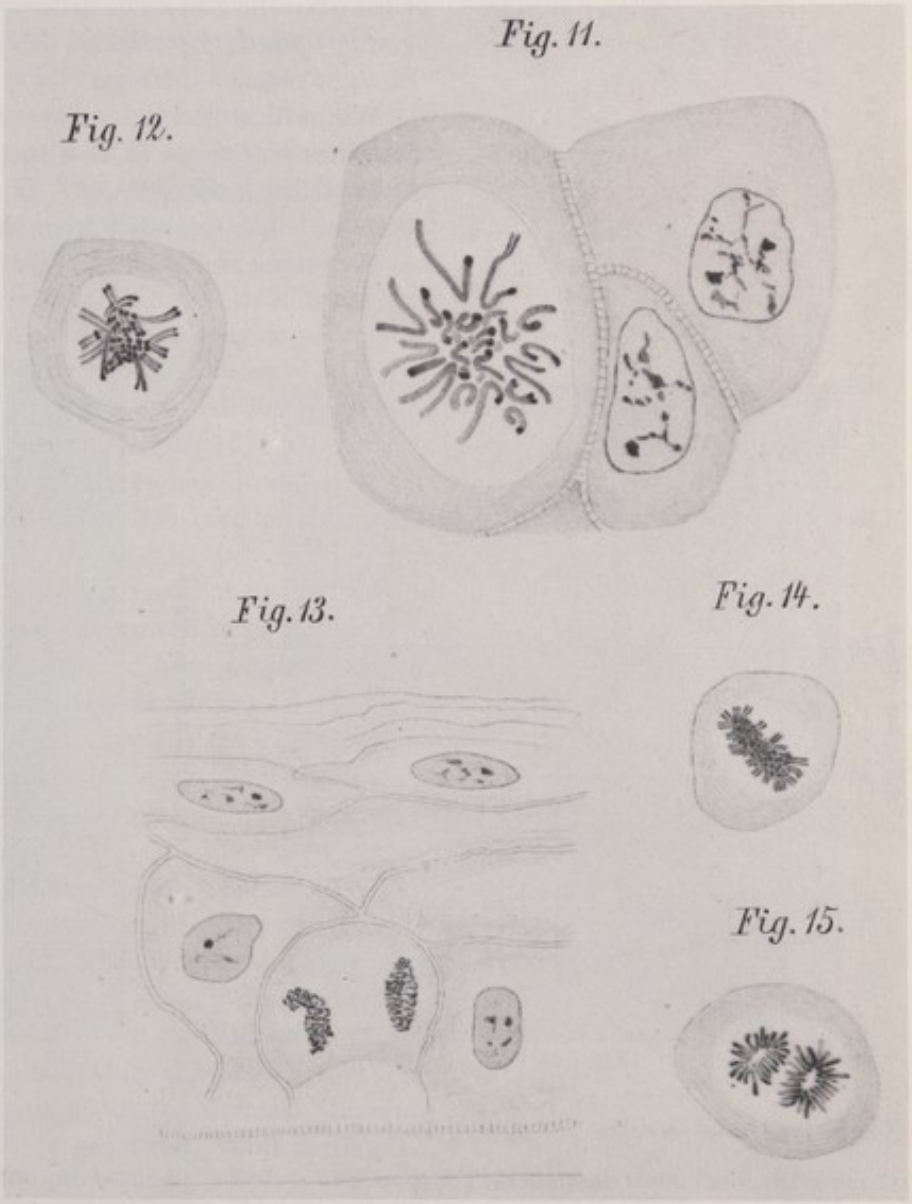
BLAKE CABOT
Publisher

New York
March 1973

Introduction

It is a commonplace that as infectious and nutritional diseases are better understood, controlled, and managed, congenital and genetic disorders assume greater relative significance, as shown by the statistics plotted in graphs on page XIV. To this the increased importance of medical genetics in the last two decades can in large part be attributed. Some disorders, even frequent ones such as cystic fibrosis, and rarer ones, such as the immune deficiency diseases, were not even recognized until rather recently. Nosographically they are post-chemotherapy or post-antibiotic diseases. Cystic fibrosis, first well described by Dorothy Andersen in 1937, did not become fully delineated until after World War II. Bruton's X-linked agammaglobulinemia, the first of the immunoglobulin deficiencies, was not recognized until the early 1950's. Hitherto these cases had been buried in the great mass of infectious diseases and diarrheal and respiratory diseases, acute and chronic.

But another highly significant factor in the increased importance of medical genetics has been the improved techniques for study of the

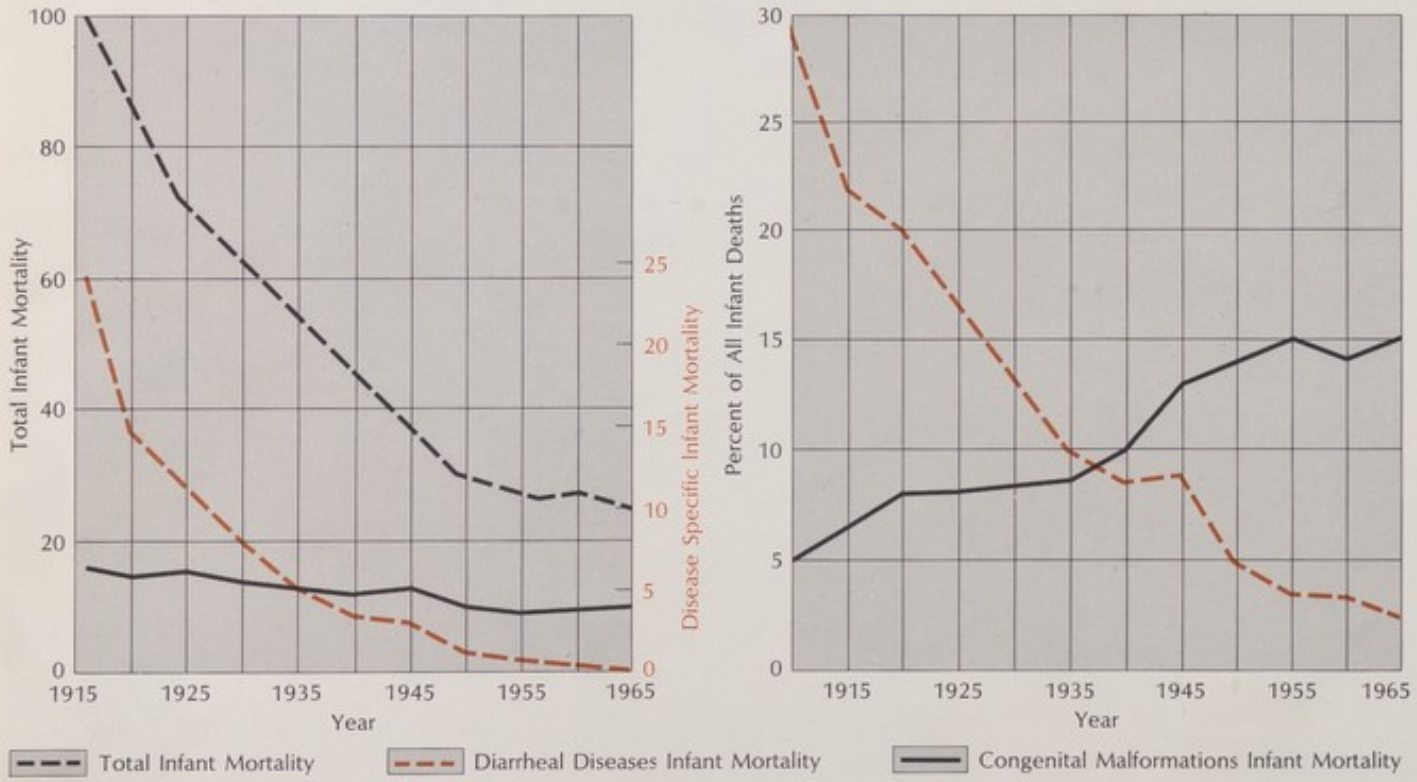


The earliest known renderings of human chromosomes were these drawings based on observations in dividing nuclei from epithelium. They were published by the German microscopist W. Flemming in 1882.

genetics of man, and the increased information, as well as improved diagnosis and management, which has come out of the application of these techniques to the investigation of disease. Relatively facile methods for examining the chromosomes gave the medical geneticist "his" organ or system like the cardiologist or endocrinologist. Human chromosomes were visualized as early as 1880 by Walter Flemming (see photo on page XIII) but the availability of methods for analyzing them in a few drops of blood made a big difference. And, of course, the refined banding techniques give greater precision to the scrutiny of the chromosomes (see Caspersson and Zech, Chapter Three).

Burgeoning biochemical genetics has often provided surrogates for controlled matings that are, of course, impossible in man. Not only can deductions be made, for example, about the evolution of man from the analysis of protein structure, but conclusions about matters such as mode of inheritance are sometimes possible when only a single patient is available (e.g., if both parents have an intermediate level of activity of an enzyme almost completely absent in the patient). Cell hybridization is what Marcello Siniscalco calls "a substitute for sex," another method for getting genetic information which the experimental geneticist obtains through planned matings.

We medical geneticists have also been fortunate that the lowly fibroblast has turned out to have such a wide repertory of enzyme activities (see Mellman, Chapter 16). It was not anticipated, by many at least, that fibroblasts cultured from skin biopsies would be almost as useful as liver tissue, for example, in studying diseases such as homocystinuria



Graph at left shows infant deaths in the U.S. from 1916 to 1965. Rapid decline since 1916 in total infant mortality is paralleled by decline in infant mortality from specific infectious diseases, here represented by diarrheal diseases, including dysenteries. During the same period, infant mortality from congenital malformations declined very little. In graph at right, percentage of infant deaths in U.S. from 1910 to 1965 from diarrheal diseases is compared with percentage for congenital malformations. Data are from Dr. Hymie Gordon, Chairman, Department of Medical Genetics, Mayo Clinic (JAMA 217:1215, 1971).

when direct assay shows no significant amount of the relevant enzyme, cystathionine synthetase, in skin. The development of the prenatal diagnosis of inborn errors of metabolism depends also on this fortunate circumstance; if the fibroblasts of amniotic fluid did not have the same wide range of enzymatic capacities, their deficiency in genetic diseases could not be determined. The exceptions such as phenylketonuria (phenylalanine hydroxylase is normally not present in skin or amniotic fibroblasts) serve to underscore the good fortune of the usual situation (see Dancis, Chapter 24).

Sometimes a cell that does not show a particular function can be induced to do so. An example is the lymphocyte that when stimulated to blastic transformation by phytohemagglutinin shows a number of enzyme activities. Not only can the homozygous state of Garrodian inborn errors be diagnosed, but the heterozygous or carrier state can be detected as well, as shown by the recent work of Joseph Goldstein and colleagues in Seattle and Dallas.

Parenthetically, medical geneticists making use of the study of cells in place of the whole patient have "cashed in" on a reservoir of morphologic, biochemical, and other information in cell biology derived in no small part from study of the famous cell line cultured from the patient pictured on this page. Henrietta Lacks was only 31 when she presented to the Johns Hopkins Hospital in February 1951 and was discovered to have cervical carcinoma. Dr. George O. Gey established the HeLa cell from a biopsy of the carcinoma. Recent review of the histology by Howard W. Jones and his colleagues suggests that it should be labeled adenocarcinoma rather than epidermoid carcinoma. This is an unusual type for the cervical site. However, its atypical histology may correlate with the unusually malignant behavior of the carcinoma. Although, at diagnosis, it was considered to be in an early stage and was "adequately" treated with radium therapy, the patient was dead by August 1951. A less grim correlate of the tumor's singularity may have been its cultural characteristics (until recently no other cervical cancer had been established in culture).

Thus, we have been fortunate in our methods – although this is not to say we do not need others – so that man has proved to be not such an unsatisfactory object for genetic study as it had been supposed earlier. Man has had, to be sure, a strong anthropocentric motivation to exercise his ingenuity in studying his own genetics, especially the genetics of his diseases. In the process he has provided some excellent models for elucidating general genetic, biochemical, and other principles. Some rare diseases, such as the Lesch-Nyhan syndrome (guanine hypoxanthine phosphoribosyltransferase deficiency, see Seegmiller, Chapter 10), have had scientific importance far out of proportion to their frequency of occurrence. Study of the Turner and Klinefelter syndromes (see Ferguson-Smith, Chapter Two) taught us much about the normal mechanism of sex determination in man. As pointed out by William Harvey in 1657 and frequently reiterated by Archibald Garrod in this century, rare disorders always have much to teach us about the normal. "Treasure your exceptions!" was the way one early geneticist, William Bateson, put it.

In thinking about the role of genetics in disease the reader will find it useful, I believe, to view disease in three categories. Some disorders are determined primarily by a single locus. These are, for the most part,

Study of HeLa cell line, cultured from biopsy of adenocarcinoma from patient Henrietta Lacks (above) in 1951, has yielded important morphologic and biochemical information. Until recently no other cervical carcinoma had been established in culture.

rare but there are many of them (see McKusick, Chapter 20) so that in the aggregate they represent a significant body of disease. The mucopolysaccharidoses, PKU, vitamin-responsive inborn errors of metabolism, membrane transport defects, and other conditions discussed here fall in this category.

Other disorders have a numerical or structural abnormality of the chromosomes of the relatively gross sort that can be detected by presently available methods. Most such instances are not inherited – at least not in the usual sense of that word – but they involve the genetic material, the chromosomes, and, therefore, represent one category of genetic disease (see Hirschhorn, Chapter One, and Ferguson-Smith, Chapter Two). Still other disorders are multifactorial (see Carter, Chapter 19). These tend to be common conditions such as essential hypertension and "garden varieties" of congenital malformations such as cleft lip and palate. Environmental and genetic factors, both usually multiple, collaborate in causation. They tend to run in families but do not show the simple mendelian pedigree patterns typical of the first category.

This three-way categorization is to some extent arbitrary: The single gene disorders are influenced by other genes and by the environment and all are, therefore, to some extent multifactorial. Chromosomal aberrations could be viewed as polygenic. Even in the multifactorial disorders the influence of a single locus can often be identified. The classification does have practical usefulness, however, because the approach to the study of the three types differs. Genetic counseling differs also, as indicated by a comparison of the presentations by Fraser (Chapter 21), Hirschhorn (Chapter One), and Carter (Chapter 19), each of which is concerned with a different one of the three.

Since the late 1950's medical genetics has experienced a marvelous weaving together of previously largely separate strands: cytogenetics, biochemical genetics, immunogenetics, and statistical-formal-population genetics. The central loom for one fabric of medical genetics is clinical genetics – the care of patients and families with genetic disorders. Clinical genetics, like other branches of medicine, is concerned with the answers to three questions, to follow the formulation of British biostatistician Bradford Hill: 1) What's wrong? 2) What's going to happen? 3) What can be done about it?

To answer the first question we seek the diagnosis. The diagnostic armamentarium of clinical genetics has increased greatly in recent years, through newly developed chromosomal and biochemical methods, but also through better clinical descriptions ("delineation") of genetic and congenital disorders. Accurate diagnosis is of the utmost importance in answering correctly questions 2 and 3.

The second question is, of course, another formulation for prognosis. "Recurrence risk" is the genetic prognosis involved in genetic counseling. In a condition of late onset such as Huntington's chorea, the question, what is going to happen, may relate not to unborn children but rather to the consultand* himself.

The third question, obviously, is likely to be a function of therapy, or rather management in a broad sense. As in diagnosis, the clinical geneti-

*_Consultand_ is the term suggested by my colleague Dr. E. A. Murphy for the recipient of genetic counseling. _Proband_ and _propositus_ are too formal terms and not sufficiently specific in their meaning. _Client_ has too commercial connotations and _patient_ is often not appropriate.

cist's armamentarium is steadily increasing. This will be evident from the presentation by Howell (Chapter 27), as well as from that on vitamin-responsive inborn errors of metabolism (see Rosenberg, Chapter Seven). Exciting implications for the therapy of lysosomal diseases are discussed in the paper on the mucopolysaccharidoses (see Neufeld, Chapter 14). Preventive aspects of clinical genetics – which are an outgrowth of the second question – are illustrated by prenatal diagnosis and selective prophylactic abortion as a measure to help high-risk families obtain a normal family (see Dancis, Chapter 24).

Over the last decade or so I have collected fourteen genetic misconceptions frequently encountered among physicians. This monograph* should help correct any of these misconceptions which may remain in the mind of the reader. I list the fourteen below with a brief criticism of each:

Misconception 1. Congenital is synonymous with genetic.

Criticism: Congenital merely means present at birth. It has no necessary etiological connotations. Some genetic disorders are not congenital in the usual sense of the word, and many congenital malformations do not have a predominantly genetic cause. For example, Huntington's chorea is genetic but not congenital; rubella embryopathy is congenital but not genetic.

Misconception 2. If a disorder is inherited, a chromosome analysis will show abnormality.

Criticism: Most mendelian disorders have no chromosome abnormality demonstrable with existing techniques. The genetic change is a point mutation or some other alteration in DNA at a level far beyond resolution by microscopic techniques.

Misconception 3. A buccal smear provides full information on the chromosomes.

Criticism: The buccal smear tells only the maximal number of X chromosomes per cell. (The fluorescent staining technique permits similar evaluation in regard to the Y chromosome.)

Misconception 4. If a genetic disorder is dominant, all children of an affected person will be affected; similarly, if all children of normal parents are affected by a genetic disorder, this is evidence of dominant inheritance.

Criticism: This displays ignorance of Mendel's laws and his meaning for the word "dominance."

Misconception 5. When only males (or only females) in a family are affected, for example a father and all four sons but none of his four daughters, this indicates sex-linkage of the disorder.

Criticism: Again ignorance of basic principles of genetic transmission is evident. Male-to-male transmission does not occur with X-linked traits because the father gives his Y chromosome to his sons—that is why they are sons—rather than his X chromosome, which bears the mutant gene.

Misconception 6. A disorder that occurs in multiple brothers and sisters with normal parents is not hereditary.

Criticism: Recessive inheritance is characterized by affected sibs with normal parents, and a recessive disorder is as genuinely hereditary as a dominant one.

Misconception 7. Consanguinity brings out sex-linked disorders, for example, "hemophilia was frequent in the inbred royal families of Europe" (Encyclopedia Britannica).

Criticism: Consanguinity increases the occurrence of homozygous affected females but has no effect on the frequency of affected males.

Misconception 8. The occurrence of a hereditary syndrome composed of two or more manifestations is the result of close linkage on the same chromosome of separate genes, each resulting in one of the individual manifestations.

Criticism: All "mendelizing" syndromes studied in full detail to date have been found to have their basis in the pleiotropic effect of a single mutant gene. Linkage produces no permanent association of traits because even closely linked genes become separated after the passage of a certain number of generations, through the process of crossing over.

Misconception 9. Dominantly inherited disorders tend to increase in severity

*The list of frequent misconceptions and the criticism of each were published in the *Annals of Internal Medicine* 75:4:642, 1971. They are reproduced here with the permission of the editor.

with transmission from generation to generation, a process called anticipation. For example, age of onset, an expression of severity, tends to be lower in affected children than in their affected parents.

Criticism: Bias of ascertainment is responsible for apparent anticipation. Only in milder cases of many dominant disorders does reproduction occur so that on the average offspring are expected to be more severely affected than their parents.

Misconception 10. Inbreeding causes a build-up of "bad genes" in populations.

Criticism: Per se, inbreeding has no effect on gene frequency. It does change *genotype* frequency; it increases the frequency of homozygotes, that is, affected persons in the case of a recessive disorder. If the homozygote is at a disadvantage (implied by "bad genes"), inbreeding actually results in a decrease in the deleterious genes. Such is the basis of the inbreeding "bottleneck" through which some populations are thought to have passed. Through close marriage over millenia, rare recessive genes, a few of which all of us carry in heterozygous state, are weeded out.

Misconception 11. Dominant disorders are common; recessive disorders are rare.

Criticism: There is no necessary relationship between the frequency of a disorder and its mode of inheritance.

Misconception 12. Dominant disorders are more severe than recessive disorders.

Criticism: As a general rule, just the opposite is the case.

Misconception 13. If a couple has had three children born with a given recessive disorder, the chance that the fourth child will also be affected is vanishingly small.

Criticism: As Le Châtelier put it, "Chance has no memory." The risk of an affected child from two carrier parents remains 1 in 4, regardless of previous experience in the family. The analogy to tossing a coin is useful in explaining this basic rule of probability to laymen.

Misconception 14. Genetic disease is not treatable.

Criticism: Although it is true that the underlying genetic defect is, at present, in no instance correctable, many genetic disorders are effectively treated by replacing the missing gene product, by restricting dietary constituents that, because of an enzymatic deficiency, cannot be adequately metabolized, by loading with a vitamin-cofactor of the genetically defective enzyme, and by other measures that influence some step between the gene and the phene.

VICTOR A. MCKUSICK, M.D.

Baltimore
March 1973

Section One

Chromosomes and Their Disorders

1

Chromosomal Abnormalities I: Autosomal Defects

KURT HIRSCHHORN
The Mount Sinai School of Medicine

The year 1956 was a milestone in human cytology, for it was then that the investigators J. H. Tjio and Albert Levan achieved the first accurate count of human chromosomes. This finding, in addition to the techniques (chiefly devised by other researchers) that made it possible, opened a new era in chromosomal investigations. In the years that have since elapsed, our knowledge of human chromosomes, both normal and abnormal, has expanded almost explosively; our techniques for studying them have also expanded, though perhaps not as fast as we could have wished. We now possess a large and increasingly detailed mass of information on chromosomal anomalies, on the processes that engender them, and on their connection with disease – both the congenital conditions resulting from gross abnormalities of the hereditary material and the pathologies developing later in life (notably, various neoplastic conditions) that appear to result from more subtle chromosomal derangements.

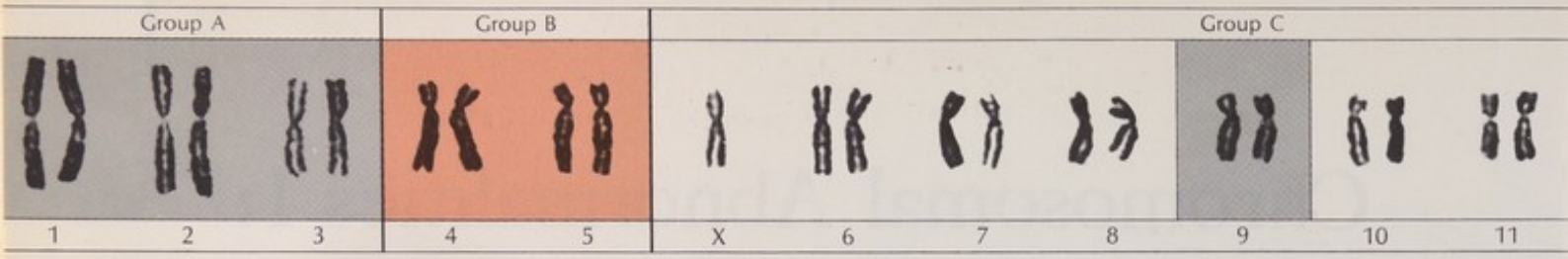
As good a way as any of getting into this subject is to describe the technique of karyotyping, the basic process by which chromosomes are identified and studied.

The first step is to culture human cells of some sort – usually, because they are easily obtained, lymphocytes from the peripheral blood. Under the stimulus of the substance phytohemagglutinin, such cells will divide; the addition to the culture of a compound such as colchicine arrests the division at metaphase, when the chromosomes are discrete and easily visible. Placed in a hypotonic salt solution, the cells absorb water and swell, thereby spreading the chromosomes apart from one another so that they can be clearly distinguished. The preparation is then fixed, stained, and photographed.

The chromosomes appear within the cell as roughly x-shaped objects that differ in size, each consisting of two somewhat worm-like bodies united at a constricted region, the centromere. It is perhaps worth noting that though the x-shape automatically spells "chromosome" to almost any physician, it is in a sense an artifact. Such chromosomes are in fact double chromosomes, united only at the centromere; had division been allowed to proceed, they would shortly have split at that point, forming two complete sets of "normal" single chromosomes for the two daughter cells. For our purposes, however, this refinement can be ignored.

The final step in karyotyping is to cut apart an enlarged photograph of the chromosomes, pair them off, and arrange them in a standardized order. This is based on their lengths and on the position of the centromere; if it joins the two bodies near one end, the chromosome is called acrocentric, if near or at the middle, metacentric – with various intermediate gradations. In this manner one ends up normally with 23 pairs of chromosomes classified into seven groups designated A to G – 44 paired autosomes plus two sex chromosomes, which are, of course, either two X chromosomes in the female or an X and a morphologically quite different Y in the male. [Sex chromosome disorders, which represent something of a special problem in physiology and medicine, are discussed in Chapter Two.]

These gross morphological criteria – size and centromeric position – will not, however, unequivocally identify all the pairs. Thus pairs 4 and 5 (group B) cannot be distinguished one from another; neither can pairs 13, 14, and 15 (group D), and so on. At one time it was thought that the identification problem could be solved, or partially solved, by means of certain secondary constrictions that had been observed on some chromosomes. It now appears, however, that these are rather unreliable guides. Only on chromosome 9 is a secondary constriction con-



As rendering of normal male karyotype makes apparent, grouping of chromosome pairs is based on length; seven groups account for the 22 pairs of autosomes and the two sex chromosomes, each placed in a position appropriate to its size. Within groups some pairs can be distinguished from each other morphologically (e.g., by location of their centromeres); such pairs are shown against a

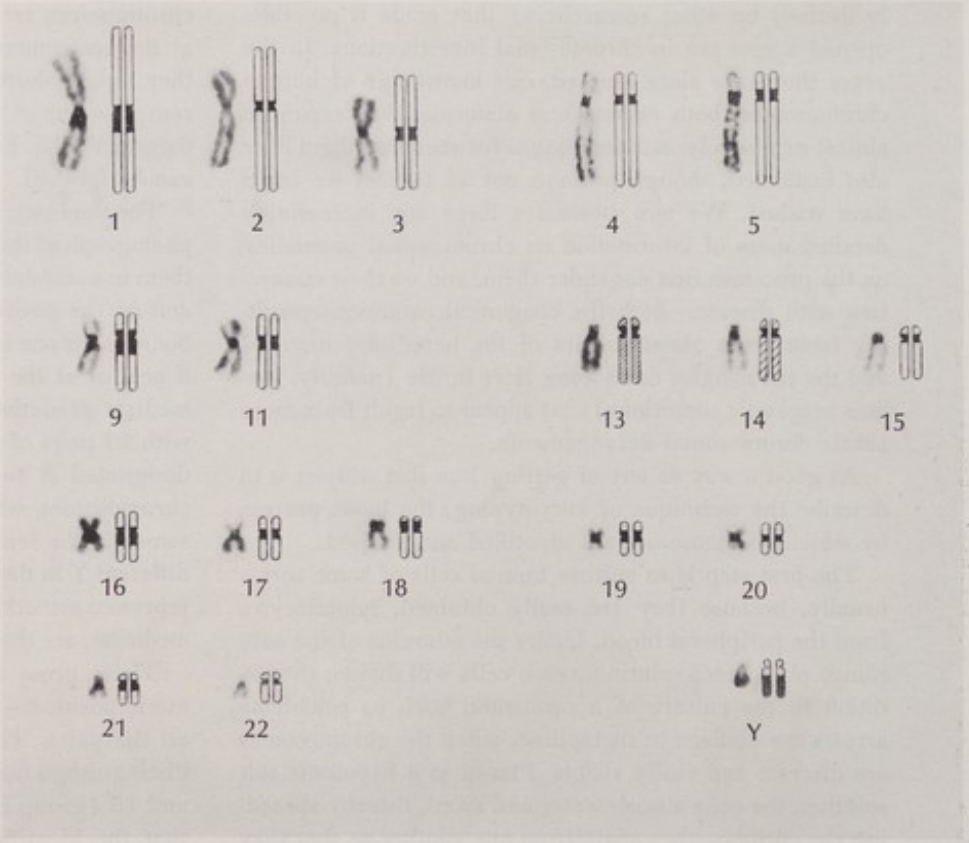
sistently present, and even here there are exceptions.

On two other chromosomes, numbers 1 and 16, constrictions are inconsistently present, but these very inconsistencies have led to some findings interesting enough to justify a brief digression. The constrictions, it appears, are inherited. Moreover, familial chromosome studies by Roger Donahue and his colleagues have established that the constriction on chromosome 1 is inherited in the same pattern within families as is the Duffy blood group – thereby establishing a strong presumption that the gene for this trait is located on that chromosome. Similar studies in England indicate that the gene for haptoglobin lies on chromosome 16.

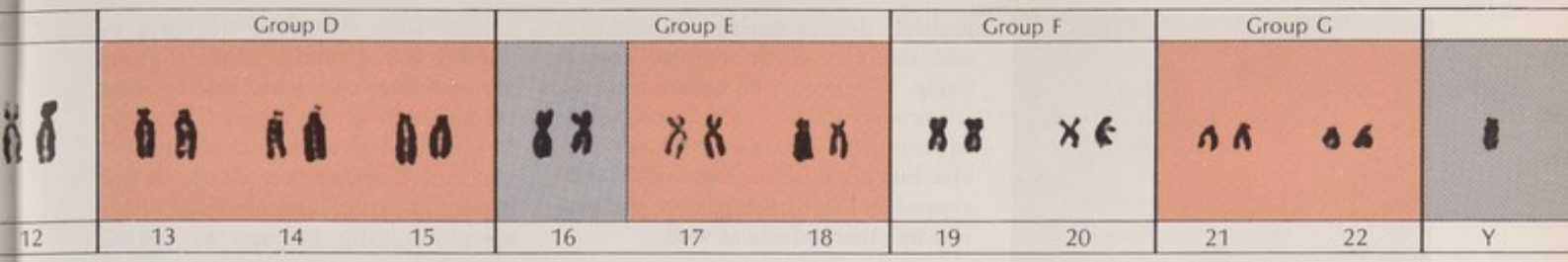
The significance of these findings lies in the fact that they represent the very beginning of chromosome mapping in human beings. In drosophila, the insect which might be said to have made modern animal genetics possible, chromosome maps have attained a complexity comparable to that of a protein molecule's structure; in man, whose generations are measured in years rather than in days, we are very far away from that level of knowledge. It seems likely that more refined study of structural variations in chromosomes will help us to bypass some of the problems inherent in man's leisurely maturation.

It has already been shown, for example, that chromosome "satellites" are also inherited. These are small bodies that may appear on the short arms of any acrocentric chromosome, excepting only the Y, attached to them by a thin stalk. The stalk is thought to represent the so-called nucleolar organizing region of the cell, or possibly an aggregation of ribosomal DNA. What the satellites themselves represent is uncertain, nor do we know, for that matter, why they should appear only on acrocentric chromosomes. It is clear, however, that not only the satellites but variations in their size are inherited, and it can be expected that comparison of their familial patterns with those of other inherited traits will continue the assignment of particular genes to particular chromosomes. A recent study in Edinburgh indicates that such minor morphological variations may be present in up to 3% of the population. None of them, I might add, has been shown to have any pathological significance thus far.

It is clear, however, that assigning a gene to a chromosome will not be terribly informative unless we can in fact identify that chromosome unequivocally. Here a technique discovered several years ago has established distinctions finer than those based on morphology alone and thus is very use-



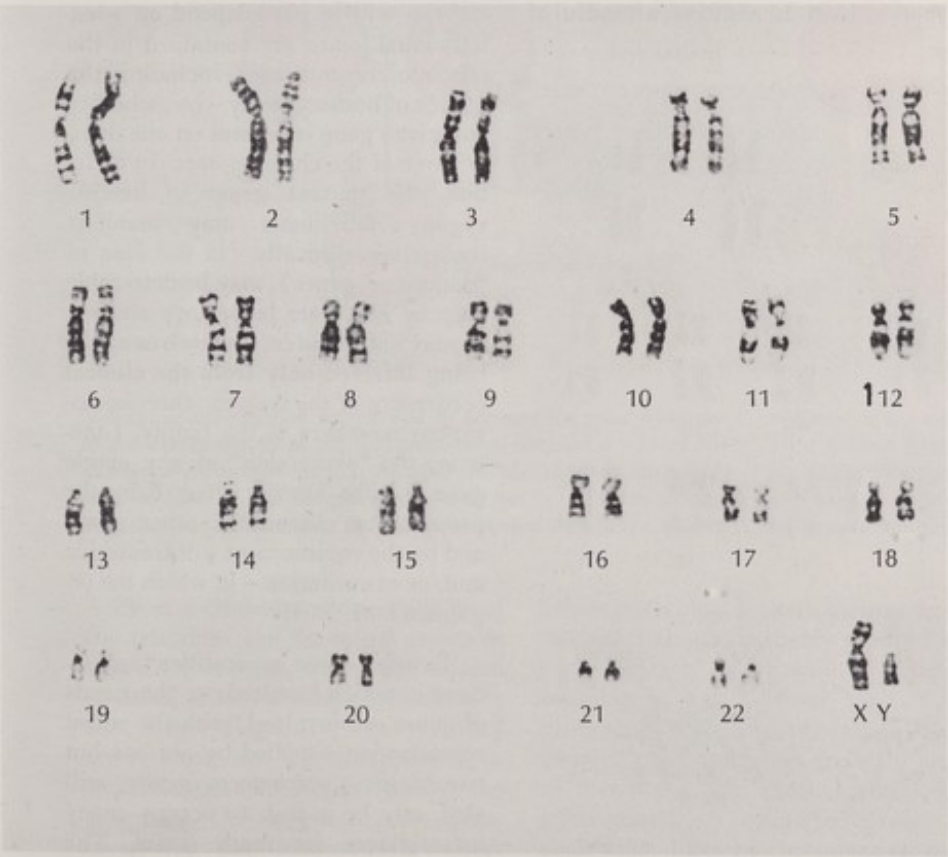
In karyotype above, renaturation and denaturation of the short chain repetitive DNA yields specific staining of chromosome's centromeres, secondary constrictions, and the long arm of the Y chromosome. The X chromosome and the autosomal pairs 6, 7, 8, 10, and 12 cannot be differentiated by this technique.



gray background. Color background indicates pairs that can be distinguished by whether DNA synthesis occurs early or late in cell cycle, as demonstrated by autoradiography. White background indicates those chromosome pairs that cannot be distinguished from others in the group by either technique. One of two X's in the female can be distinguished by autoradiography.

ful. The technique is based on the observation that morphologically similar chromosomes often manifest marked and consistent differences in the timing of their replication. This can be shown by adding an isotopically labeled precursor of DNA (usually tritiated thymidine) late in the division process. A radioautograph of the fixed preparation will then indicate the amount of subsequent DNA synthesis, thereby distinguishing between early replicating chromosomes, which had already finished making their DNA when the labeled material was added, and late replicators into which significant amounts of label were incorporated.

Thus in the group D chromosomes, for example, we find one pair (now arbitrarily designated chromosome 13) that is late-replicating throughout its length, another (14) that is late-replicating only in the region near the centromere, and a third (15) that is early-replicating throughout. Similarly, one can distinguish the B-group pairs, 4 and 5, the E-group pairs 17 and 18 (16, the third member of that group, is identifiable on morphological grounds alone), and probably the pairs of the G group, 21 and 22. Of the 22 autosomal pairs, all but eight can be identified almost always either morphologically or by timing DNA synthesis. Moreover, a combination of techniques developed in the past two years, each based on a different principle, permits accurate identification of each chromosome and its parts. One consists of denaturation and renaturation of the short chain repetitive DNA; this results in specific staining of centromeres, secondary constrictions, and the long arm of the Y chromosome. The second involves specific staining of the bands of chromosomes, including most of the long arm of the Y, with the fluorescent dye, quinacrine (see Caspersson and Zech, Chapter Three). Y chromosomes in nondividing cells can be identified by this method. In the third, most useful technique, chromosomes are treated either by changes in pH or with proteolytic enzymes such as trypsin; specific regions of each chromosome can then be stained with Giemsa, and banding patterns for each arm of each chromosome that are typical for only that chromosome can be observed. Such precise identification enables us to characterize chromosome abnormalities more accurately.



Chromosomes in karyotype above have been treated either by changes in pH or with proteolytic enzymes so that specific regions of each can be stained with Giemsa. This produces banding patterns for each arm of each chromosome that are typical for only that chromosome. All 46 chromosomes can be identified by this method.

Of these, the most obvious are the trisomies, in which the individual has three rather than two of a particular chromosome, for a total of 47 instead of 46. All autosomal trisomies produce severe damage, and indeed nearly all are incompatible with life – in sharp contrast to sex-chromosome trisomies,

Photo above shows Y body in a buccal mucosal cell stained with the fluorescent dye quinacrine.

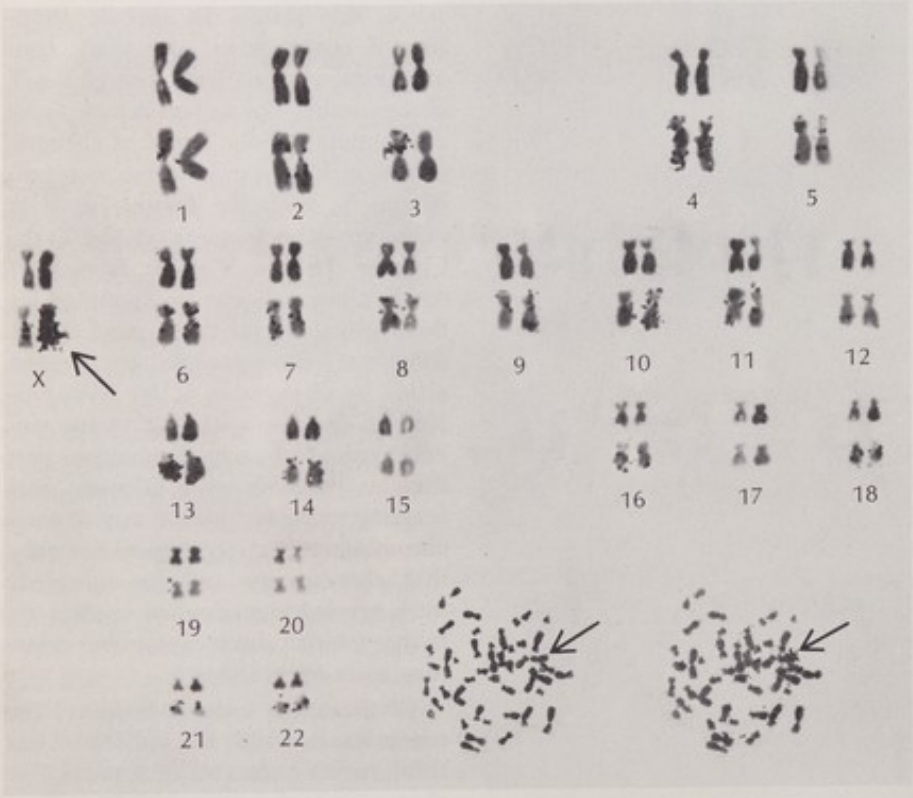
which in some instances seem to produce no detectable abnormality whatever. Trisomies affecting every one of the identifiable autosomes or autosome groups have been found, but the vast majority of these turn up only in early abortion material. Of those observed postnatally, two (trisomies 13 and 18) normally produce death from multiple defects during the first year and the first three months, respectively. The third well-known trisomy, 21, produces Down's syndrome, whose sufferers, though many of them survive into adult life, are severely handicapped by mental retardation and usually by other defects as well.

Recently my associates, Drs. L. Hsu, L. Shapiro, and M. Gertner, and I turned up another trisomy, which had sometimes been confused with Down's syndrome. The new cytological techniques already described indicate that it affects chromosome 22 rather than 21, producing a distinctive and consistent clinical picture. Without going into elaborate detail, one sees a characteristic facies including downturned mouth and microphthalmia, certain bony abnormalities, and, frequently, a history of multiple abortion or long periods of sterility in the mother. As with Down's syndrome, the sufferers frequently die early from congenital heart defects, but if these are not present the children may survive for (at this writing) several years at least. In addition, a handful of reports have described trisomies involving one or another of the C group chromosomes not identifiable by their morphology or by autoradiography, but little more can be said of these beyond the fact that they clearly do not involve an extra X chromosome (which morphologically belongs to the C group).

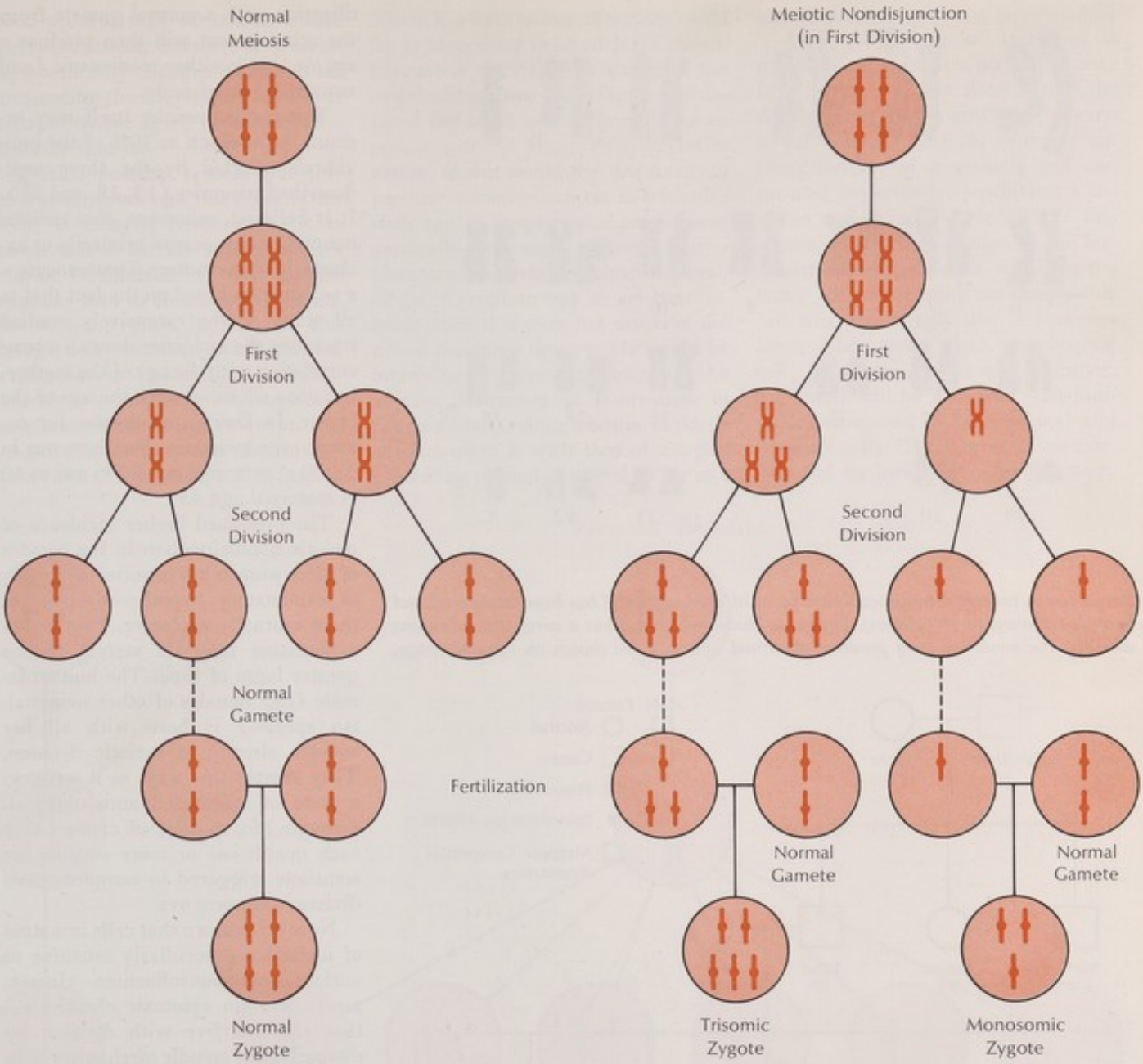
As the above paragraph suggests, there is considerable clinical variation among cases of a particular trisomy. The reasons for this are obscure and can be discussed only in generalities. The existence of three rather than two of a given chromosome obviously sets up severe biochemical imbalances in the growing fetus, but what these are and how they operate can barely be conjectured, and then only by rather distant analogy with much simpler genetic disorders.

We know that each chromosome contains hundreds and probably thousands of different genes and that many of these (some estimates run as high as 30%) differ from one individual to another. Obviously the imbalances set up in a particular trisomic embryo will in part depend on what individual genes are contained in the trisomic chromosomes, including the degree of homozygosity – i.e., whether any given gene is present on one, two, or three of the chromosomes. In addition, the mutant genes in heterozygous individuals may manifest themselves clinically (in the case of "dominant" genes), may be detectable only by elaborate laboratory studies, or may not manifest themselves at all, being inferred only from the clinical occurrence of the trait in other, homozygous members of the family. Likewise, the "expression" of any single gene is often heavily affected by the presence or absence of other genes and by the environment – intrauterine and/or extrauterine – in which the organism finds itself.

To apply these generalities to situations in which hundreds or thousands of genes are involved, with the added complication supplied by not one but two "degrees" of heterozygosity, will obviously be a task to occupy many investigators for many years. The same is true *a fortiori* of the biochemical processes involved in trisomies – a subject on which we are, if possible, even more in the dark.



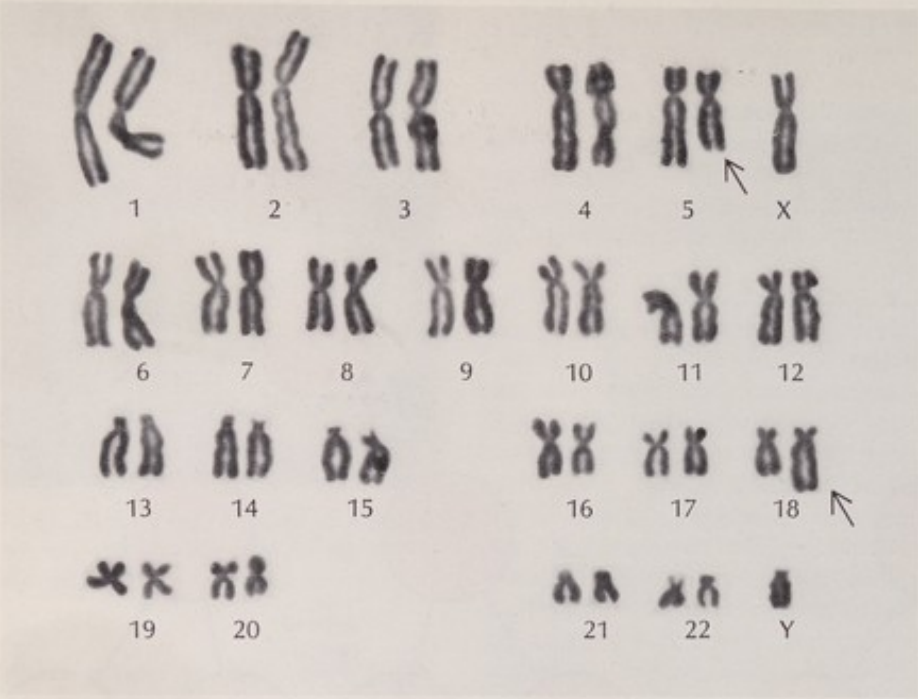
Distinction of chromosome pairs by whether they replicate early or late is made by autoradiography following the addition of a labeled DNA precursor late in cell division. Photo compares nonlabeled and labeled karyotypes of same individual. The late-replicating X of the female can be observed to be heavily labeled (arrows).



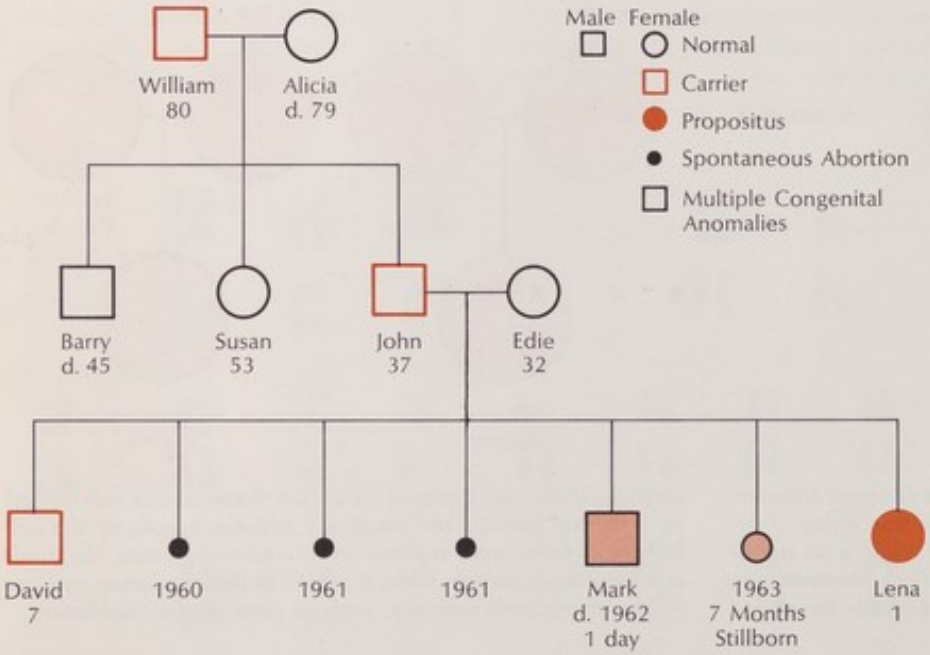
Meiotic nondisjunction is believed to account for most trisomies of the autosomes. If at the first (reduction) division a pair of homologous chromosomes fails to separate, the result is an imbalance in each of the daughter cells: one has both chromosomes, the other has neither. The imbalance is carried over to the second division. If the cell containing the two chromosomes is fertilized by a normal gamete, the result is a trisomic zygote. If the cell lacking a chromosome is fertilized by a normal gamete, the result is a monosomy, which, when it occurs in the autosomes, appears to be incompatible with life, perhaps even with implantation.

From a clinical standpoint, none of the trisomies can be called common; Down's syndrome, the most frequent of them, affects perhaps one infant in 500. Embryologically, however, they are very common indeed, appearing in something like 20% of early abortuses examined. One might expect, therefore, that the monosomies, where there is one chromosome too few rather than one too many, should be equally common, the more so in that the mechanisms thought to produce trisomy (which I shall discuss in a moment) should in theory produce an approximately equal number of monosomies. In fact, however, there is no clinical example of an autosomal monosomy, apart from one doubtful case of monosomy 21, which I and some other geneticists believe to represent a somewhat different chromosomal disorder. Indeed, so far as I know, monosomy has never been observed even in early abortion material. The obvious, and I think the correct, conclusion is that the monosomic zygote never succeeds in implanting itself; it may not survive fertilization. In this area, by the way, the contrast between autosomes and sex chromosomes is much less marked; individuals (or abortuses) with a single Y chromosome have never been observed, and while the single X condition is a well-known clinical entity (Turner's syndrome) the great majority of such individuals perish before birth.

Trisomies are produced by a num-



Karyotype in balanced translocation: a piece of chromosome 5 has been broken off and become attached to 18 (arrows). Patient is clinically normal but a carrier; mechanisms whereby the condition may produce abnormal offspring are shown on opposite page.



The pedigree shown is that of a family with a 5/18 translocation, traced through karyotyping to a carrier grandfather. Only one child in the third generation survives as clinically normal (though a carrier); the propositus survives but is severely retarded.

ber of different mechanisms. The great majority, however, are believed to result from the process known as meiotic nondisjunction. Gametocytes (oocytes or spermatocytes) in the course of their maturation undergo meiosis by two divisions (sometimes the first is called reduction division) in which 46-chromosome cells divide into gametes with (normally) half that number; the "haploid" gametes join at fertilization to form a zygote with the normal 46 chromosomes. Occasionally, however, it happens that meiosis proceeds imperfectly; two gametes formed from a single gametocyte receive respectively one chromosome too few and one too many. Fertilization with a normal gamete from the other parent will then produce a zygote that is either monosomic (and nonviable) or trisomic.

This mechanism by itself may account for as much as 70% of the individuals affected by the three well-described trisomies (13, 18, and 21). It is believed, moreover, that meiotic nondisjunction occurs primarily or exclusively in the maternal gametocyte – a supposition based on the fact that in all three of the extensively studied trisomies, the incidence shows a strong correlation with the age of the mother, but none whatever with the age of the father. In Down's syndrome, for example, the incidence rises from one in 3,000 at maternal age 20 to one in 40 at maternal age 45.

The presumed higher incidence of meiotic nondisjunction in the oocytes of older women has begotten a number of explanatory hypotheses – none of them mutually exclusive. Part of the explanation may lie merely in the greater lapse of time. The human female (like females of other mammalian species) is born with all her oocytes already in meiotic division. They remain this way, as it were, in a state of suspended animation, all through life, except, of course, that each month one or more oocytes are somehow triggered to complete their division and form ova.

Now it is known that cells in a state of meiosis are peculiarly sensitive to various exogenous influences – viruses, x-rays, certain cytotoxic chemicals – that can interfere with division by damaging the spindle mechanism or in other ways. Obviously, the longer a woman lives the longer these environmental agents will be able to have an effect.

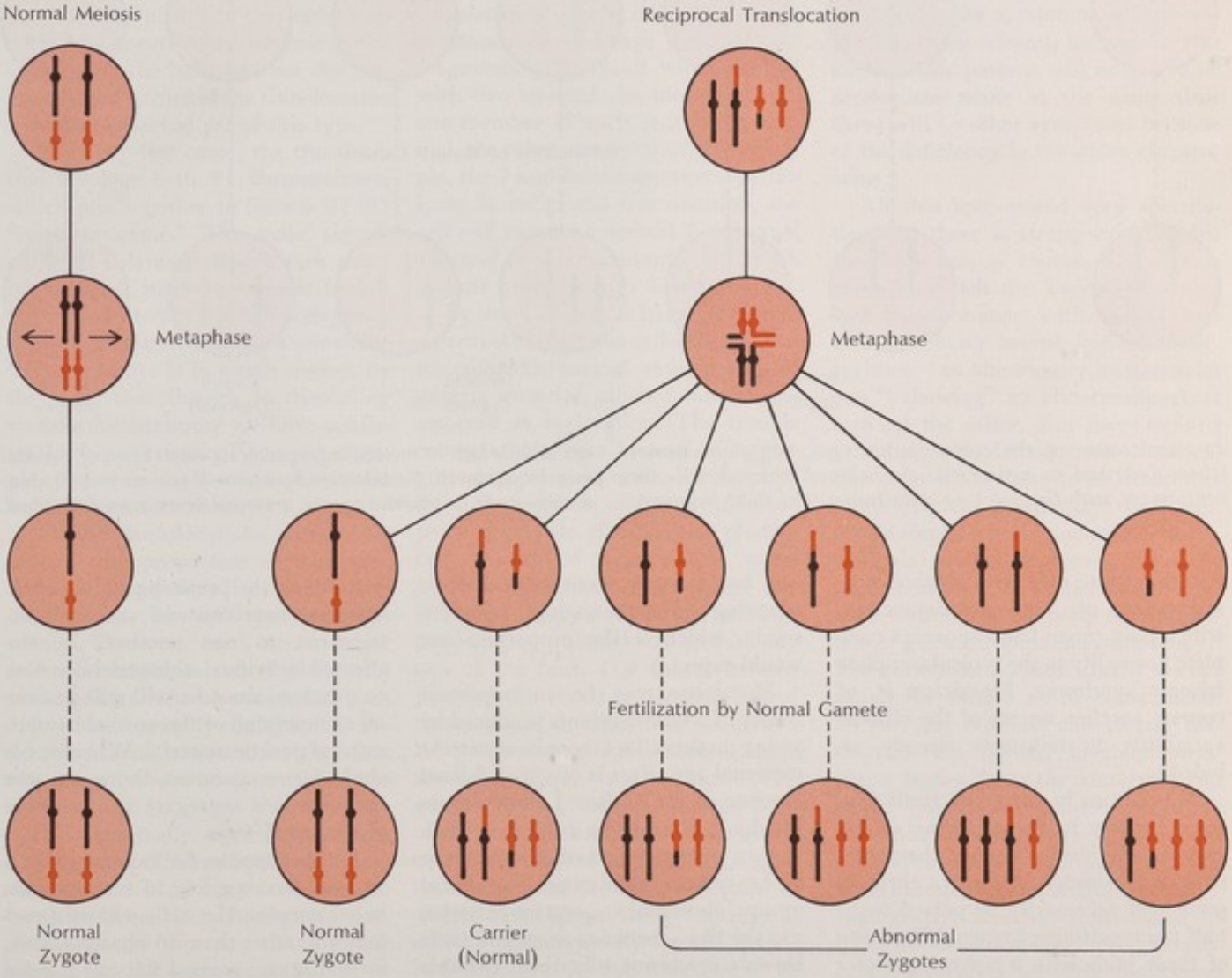
This may well be part of the explanation for trisomies, but it is almost certainly not the whole story. The reason lies in the fact that the simple lapse of time would be expected to produce a simple, linear increase in the "maternal age effect." However, from age 40 on the increase is not linear but exponential. This suggests a second possibility: that the aging processes of the ovary itself are responsible; that these are very potent physiologically we know from the fact that they culminate in complete cessation of ovarian function at menopause. It could well be that these processes

speed up after age 40 and as a side effect occasionally interfere with meiosis, whether by changes in the tissues surrounding the oocytes, by hormonal shifts, or by the aging process itself – whatever that is.

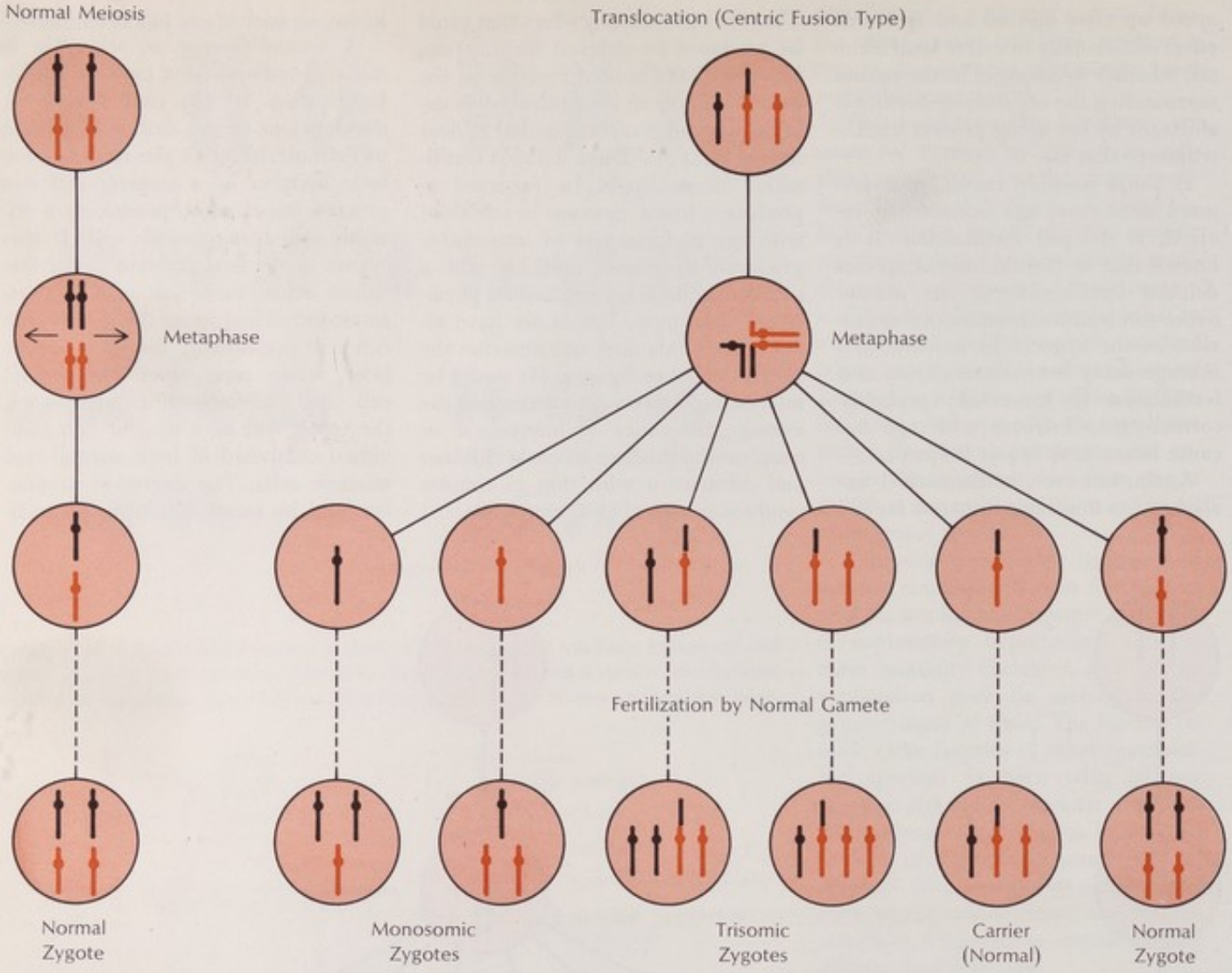
A third possible reason, put forward some years ago but recently revived, is delayed fertilization. It is known that in certain animal species delayed fertilization of the mature ovum can lead to chromosomal anomalies in the zygote. In humans, the average delay between ovulation and fertilization is presumed (probably correctly) to increase with age because intercourse is less frequent.

Again, however, mathematical analysis rules this out as a major factor. There is a limit to any effect that could be produced by delayed fertilization, because of the limited viability of the ovum and sperm; if fertilization is delayed beyond a certain period it does not occur at all. Thus, delayed fertilization alone would be expected to produce a linear increase in trisomies with age as frequency of intercourse gradually decreased, probably with a plateau as the delay reached its physiologic maximum; but as we have already seen this does not describe the actual incidence figures. (It would be interesting, however, to determine the average frequency of intercourse in couples who produce trisomic children and compare it with that of couples producing normal children; so far as I know, no such study has been made.)

A second source of trisomies is mitotic nondisjunction, occurring after fertilization. At any time during the development of the embryo (or even in extrauterine life) the ordinary mitotic division of a somatic cell can proceed imperfectly, producing a trisomic and a monosomic cell. If this occurs at the first division of the fertilized ovum, the result will be a trisomic individual, since the monosomic cell will presumably die. If it occurs later, when more than one normal cell has already been produced, the result will be a mosaic – an individual composed of both normal and trisomic cells. The degree of mosaicism and its location within the body



A reciprocal or balanced translocation occurs when two chromosomes break and their fragments are exchanged. It is not harmful to the carrier individual (who has the full complement of genetic material, merely somewhat rearranged), but depending on the type of centromeric segregation that occurs during reduction division, six major different types of gamete will be produced. After fertilization by a normal gamete, the zygote may produce either a normal individual, another carrier, or one of four types of abnormality, as shown above. All the abnormal zygotes will have some duplication and some deficiency of chromosomal material. The nature of the duplication and the deficiency will vary according to the segregation into haploids during meiosis.



In translocation of the centric fusion type, two chromosomes (those described as acrocentric) are believed to break near the centromere, with the two long arms fusing. The short fragments are usually lost. At meiotic segregation six major types of gametes may be produced. Fertilization by a normal gamete may produce a monosomic or trisomic zygote, a normal or a carrier individual.

will determine the physiological consequences of the nondisjunction, which may range from apparent complete normality to the typical complete trisomy syndrome. Mosaicism is, of course, another source of the clinical variability in trisomies already alluded to.

Mosaicism, in turn, can itself produce trisomy in the offspring of the mosaic individual. If one of the gametocytes is trisomic, the extra chromosome will necessarily be passed on to half the resulting gametes. When one of these joins with a normal gamete, a trisomic zygote will result. (This same process, called obligatory nondisjunction, also operates in completely trisomic Down's syndrome females, who are generally fertile. A few Down's syndrome females have had babies, and about half of their offspring have themselves been trisomic, which is the proportion one would expect.)

Mosaicism may be an important source of trisomic infants produced by young mothers, i.e., those in whom the maternal age effect is minimal. Lionel Penrose of the Galton Laboratory in London has studied a number of such women and has found what he believes to be certain similarities – in blood groups, abnormal fingerprint patterns, and the like – between them and their Down's syndrome offspring. On this basis he has suggested that up to half of them may be undetected mosaics.

The final source of trisomies is translocation. This complicated phenomenon will be explained more fully later on; in this particular context, it amounts to the breakage of two chromosomes near one end and their attachment to one another. In the affected individual, this generally does no damage, since he will still possess all or nearly all of the normal complement of genetic material. When he (or she) forms gametes, however, the chromosomes segregate in a number of abnormal ways.

Let us suppose, for example, that a woman carries a 21/15 translocation in her oocytes. Her cells will then contain 45 rather than 46 chromosomes, including one normal 21, one normal 15, and the translocation combining both. Meiosis can produce six major chromosome combinations in the ova: 21 alone; 15 alone; 15 15/21; 21 15/21; 15/21 alone; and the normal 15 21. When fertilized by a normal

sperm, the first three will produce, respectively, monosomy 15, monosomy 21, and trisomy 15, all presumably nonviable. The latter three will produce, respectively, trisomy 21 (Down's syndrome), a phenotypically normal but "carrier" individual like the parent, and a completely normal individual. Thus in theory the chance that a woman with a 15/21 translocation will produce a Down's syndrome baby is one in three. Empirically, however, the probability is nearer one in five – the difference presumably being due to the lessened viability of the trisomic embryo. Curiously, if the father carries the translocation the empirical chance of a trisomic infant is only about one in twenty; the reason for this is not known. Translocation can also occur de novo in a particular parental gametocyte, in which case it will show up in the baby but not the parents. About a third of the translocation trisomies observed are of this type.

In a very few cases, the translocation involves both 21 chromosomes, which join together to form a 21/21 "isochromosome." The main significance of this exceedingly rare event is that such individuals must invariably produce 21 trisomic offspring, the only alternative being a nonviable 21 monosomy. It is worth noting, by the way, that though in discussing translocation trisomy we have consistently chosen trisomy 21 as our example, the same mechanisms have also been observed in trisomy 13.

Though translocations account for only a tiny proportion of trisomies, they are responsible for many other less obvious chromosomal derangements. These, in turn, are a major source of fetal loss and of a virtually unlimited variety of congenital malformations – usually multiple. To understand them, however, we must first say a word about how translocations occur.

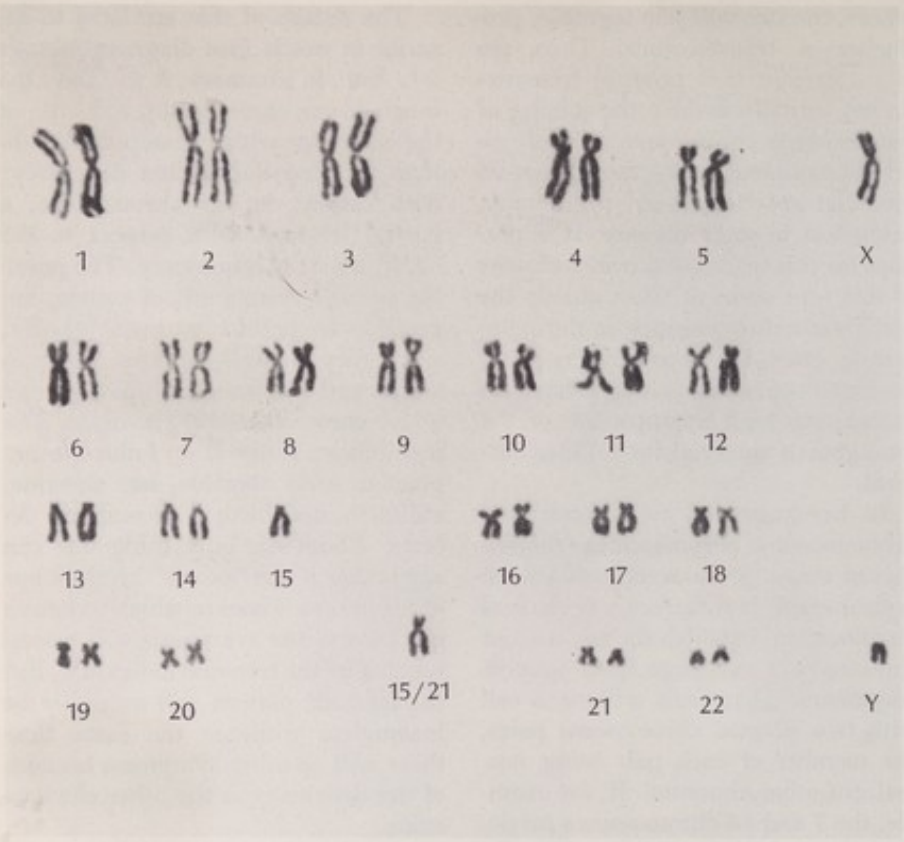
It is generally agreed nowadays that any sort of translocation requires chromosome breakage – and not one break but two. An unbroken chromosome remains discrete, but a break creates "sticky" surfaces at the broken ends. By itself this will merely render the chromosome unstable, ultimately leading to cellular death. (The break may, of course, simply heal itself.) If, however, one of the sticky surfaces encounters another such surface elsewhere, the two will join together, producing a translocation. Thus the translocations that produce trisomies do not actually involve the joining of two complete chromosomes but of two almost complete ones – the broken-off bits that are "left over," presumably being lost in some manner. It is perhaps for this reason that translocations of this type seem to affect chiefly the acrocentric chromosomes; in the metacentric ones, for geometric reasons, the lost fragments would in most cases include too high a proportion of the total genetic material for cellular survival.

In breakages of metacentric and submetacentric chromosomes (though also at times in the acrocentrics) the typical result is therefore a reciprocal translocation, in which the two broken chromosomes exchange their split-off fragments. The result will be a cell with two affected chromosome pairs, one member of each pair being normal, the other abnormal. If, for example, the 7 and 16 chromosomes participate in reciprocal translocation, the cell will contain a normal 7, a normal 16, and two chromosomes that each include parts of both 7 and 16.

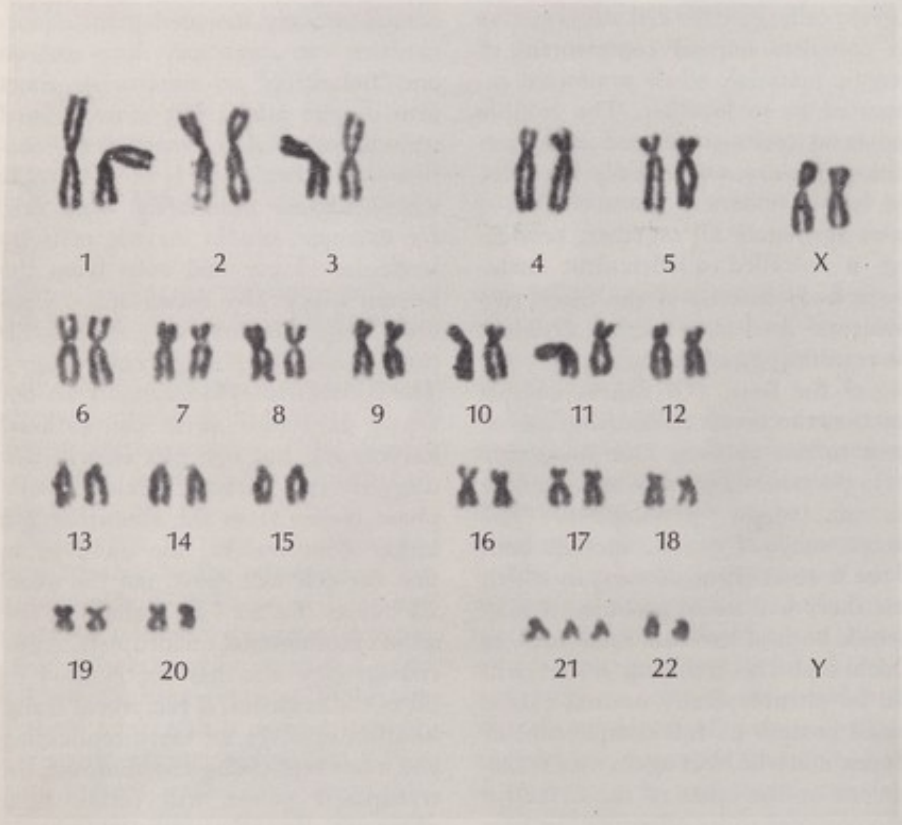
By itself this is not harmful even to a germ cell, since the cell still contains its complete normal complement of genetic material, albeit somewhat reassorted as to location. The trouble comes at meiosis: instead of aggregating in pairs, as normally happens, the four members of the two affected pairs aggregate all together, producing a so-called quadrivalent metaphase body instead of the usual two bivalents. And at reduction division, the resulting gametes end up with any two of the four. (Of course nondisjunction can result in distributions of three to one or even four to none.) As in the trisomy translocations, there are six major possibilities. The gamete may, of course, include both of the normal chromosomes, in which case there will be no problem; it may include both of the abnormal ones, in which case the resulting infant will still be phenotypically normal (since it will possess its full complement of genetic material) but again may transmit one or the other of the defective chromosomes to its offspring. In four of the six cases, however, the gamete will receive one normal and one abnormal chromosome.

The details of this are hard to describe in words (see diagram on page 9), but, in summary, it will have too much of one chromosome, too little of the other, for which reason the condition is called duplication deficiency: with respect to one chromosome, a partial trisomy, with respect to the other, a partial monosomy. The possible somatic results are, of course, impossible to predict or even classify, since they depend on how much is added and how much is missing, and which chromosomes are involved. The possibilities range from failure to implant to early abortion, late abortion, stillbirth, and birth with multiple defects. About the only thing one can say is that if the "excess" involves one of the chromosomes in which trisomies are known, the symptoms will resemble that of the trisomic individual. But the trisomic pattern will naturally be incomplete while at the same time there will be other symptoms because of the deficiency in the other chromosome.

All this may sound very speculative, but there is strong evidence for it. There are, of course, the obvious cases in which the karyotype shows two chromosomes with visible and complementary morphological abnormalities – an abnormally long arm on one "balancing" an abnormally short arm on the other. But more refined techniques can demonstrate translocation even when there is no detectable morphological peculiarity. One can, for example, obtain meiotic cells by testicular biopsy and even from the human ovary (by rather more elaborate techniques involving administration of gonadotropins and culdoscopy). The chromosomes individually are not nearly as visible as in conventional karyotypes, but one can clearly distinguish the normal bivalent metaphase bodies from the abnormal and larger quadrivalent, the more so in that the cell will show, not the usual 23 bodies but 22 – 21 pairs and one set of chromosomal quadruplets. Autoradiography also has contributed its bit to the evidence; if reciprocal translocation involves an early replicating and a late replicating chromosome, the transposed pieces will retain their characteristic timing, so that one will see part of the chromosome replicating earlier or later than the remainder. By these methods it has been shown, for



In patient with Down's syndrome and only 46 individual chromosomes, a centric fusion translocation of 15 and 21 has attached an extra 21 to the second 15. Since patient actually has the genetic material of 47 chromosomes, condition is considered a trisomy.



This patient also has Down's syndrome, but with trisomy 21, the result of meiotic nondisjunction (see page 7). The risk of trisomy 21, the commonest autosomal trisomy, rises with maternal age; this is not true of centric fusion (translocation) trisomies.

example, that up to 15% of idiosyncratically infertile males with apparently normal karyotypes actually possess "invisible" translocations. Their infertility is expressed as chronic embryonic failure, probably before implantation.

A rather special case of translocation is chromosome "deletion," an apparent loss of some chromosomal material. The results include the *cri-du-chat* syndrome, ascribed to the partial loss of the short arm of chromosome 5, and a somewhat similar condition but without the cat-cry feature, due to a similar loss in the morphologically indistinguishable chromosome 4, and two other syndromes which seem to involve losses on, respectively, the long arm and short arm of chromosome 18.

In fact, however, there is good reason to suppose that most or all of these deletions result from balanced translocations in which one chromosome has ended up, so to speak, with the short end of the genetic stick – a process whose results would be morphologically indistinguishable from a deletion. In several *cri-du-chat* cases, for example, one or the other phenotypically normal parent has been shown to be a carrier of a reciprocal translocation involving the short arm of chromosome 5. Where the parental chromosomes are normal, the deletion may well reflect a translocation occurring de novo in the paternal or maternal gametocyte.

The interpretation of deletions as actual translocations is also plausible on simple statistical grounds. Since, as already noted, two chromosome breaks are required if chromosomal stability is to be regained, a deletion must involve not simply breaking off part of one arm but two breaks in the same arm (probably rather close together), with loss of the proximal "interbreak" fragment and a rejoining of the distal fragment to the rest of the chromosome. That the disruptive agent – virus, radiation, or whatever – would score two "hits" so close together would seem considerably less likely than two hits on two different chromosomes. Even two hits on different arms of the same chromosome are unlikely, as we know from the rarity of the so-called ring chromosomes, in which the two broken ends have looped around and joined together.

There are various other theoretically possible results of chromosome breakage, some of which can be detected with our present rather crude methods, some of which cannot. (It goes without saying that the new techniques of chromosome identification mentioned above have already led and will continue to lead to accurate identification of structural rearrangements such as translocations and deletions.) As our techniques continue to improve, however, I should expect many presently "invisible" chromosome anomalies to be implicated in hitherto unexplained cases of congenital malformation – especially multiple malformation, and particularly if the condition is familial.

The same technical advances can be expected to throw light on an obscure but potentially significant aspect of human cytology: the effect of de novo chromosome damage on somatic cells. In contrast with the germ cell anomalies we have been discussing, which produce congenital and usually multiple defects, somatic anomalies would be expected to induce less spectacular, though, on occasion, equally serious disorders – in particular, neoplasms of various kinds.

We know that the chromosomes of somatic cells can be deranged by the same agents that damage germ cells – radiation, viruses, and drugs such as the radiomimetic compounds (e.g., methotrexate, the nitrogen mustards). Both radiation and the radiomimetic drugs are used to treat neoplasia for precisely this reason: their capacity to damage rapidly dividing cells.

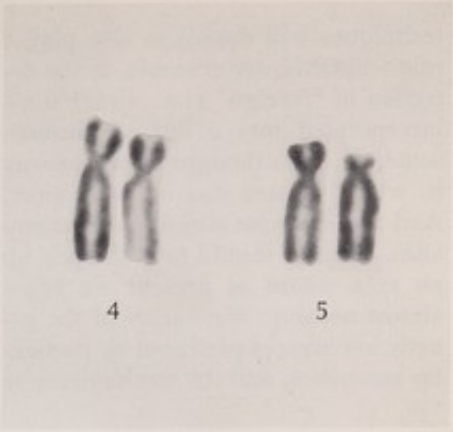
It is worth noting, however, that some forms of somatic chromosome damage appear to be benign, at least so far as the organism is concerned. Thus extensive chromosome breakage can be caused by measles or even by a bad cold, but for some reason the broken chromosomes cannot reunite to form an abnormal (or, for that matter, a normal) pattern. The damaged cells apparently die off, to be replaced by normal cells. The agents cited earlier, however, cause "healable" breakages which lead to various types of more or less stable anomalies, including translocations and apparent deletions.

Some of these may, to be sure, also produce cell death; a dicentric chromosome, with two centromeres, cannot divide normally and the ring chromosome likewise may be ultimately lethal to the cells. In other cases, the abnormal cell will continue to live symbiotically with its normal kindred, reproducing at the same rate and causing no trouble. In still other instances, however, the anomaly may confer some metabolic advantage over the normal cells. They may grow faster and divide more frequently; they may live longer; they may lose the so-called contact inhibition factor, which in wound healing stops cell growth when the wound closes. In chronic myelogenous leukemia (CML), the only neoplastic disease associated with a specific chromosome anomaly, it is strongly suspected that the anomalous cells actually replicate more slowly than normal bone marrow, but live much longer, by this mechanism eventually replacing the normal population. A more subtle derangement that has been hypothesized is the loss of the cell's capacity to synthesize some nutrient normally manufactured by the cell – an amino acid, for example. Assuming that the nutrient is freely available exogenously, the energy normally employed in its synthesis could be channeled into other growth processes, producing a competitive advantage with normal cells.

The supposition that chromosome damage is at the root of many if not all neoplasias is based on a number of well-known facts. First, the same agents that are known to damage chromosomes (such as radiation) are also known to induce neoplastic disease. In addition, chromosome damage has consistently been detected in some forms of neoplasia, particularly the leukemias. As against this, however, we have the fact that in only one type of neoplasia, CML, has a specific chromosome anomaly been detected (the so-called Philadelphia chromosome, an apparent partial deletion of the long arm in chromosome 22 [see Caspersson and Zech, Chapter Three]. As indicated above, this anomaly may be a translocation rather than a deletion.) In others, though chromosome anomalies are consistently present, they are, or at least appear to be, of no consistent type. And in still others, including some forms of leukemia, the chromosome damage (if it exists at all) is undetectable by present techniques.

The hypothesis that visible or invisible chromosome anomalies occurring postnatally are an important cause of neoplastic disease would, of course, appear to conflict with the many observations of a familial predisposition to such pathologies, which would presumably reflect an inherited chromosome disorder. One mechanism by which this conflict might be resolved has been found in Fanconi's anemia and a few other rare diseases (see German, Chapter Four). These diseases are familial, but what is inherited is not a chromosomal anomaly as such but a susceptibility to chromosome breakage and recombination; apparently the chromosomes are abnormally "fragile." Not all these individuals develop neoplastic disease, yet they are enormously more susceptible to leukemia and other tumors. In addition, when fibroblasts from such individuals are cultured and exposed to the virus SV-40 (a known carcinogen), the cells are much more likely to undergo neoplastic change (up to 100 neoplastic colonies per 10,000 cells, compared with the normal rate of 3 per 10,000).

These and other tantalizing developments open major possibilities for future research. There is first the matter of still more precise identification of individual chromosomes – in particular, fragments of one chromosome attached to another. Here electron microscopy, by revealing hitherto invisible differences among chromosomes, can be expected to produce some major advances. Biochemical



Partial karyotype of patient with cat-cry syndrome shows characteristic deletion of part of chromosome 5; the other 5 and the rest of the karyotype are normal. This is the most commonly described of several well-established deletion syndromes.

techniques will doubtless also play a role – notably, for example, in the detection of "foreign" (i.e., viral) DNA incorporated into a human chromosome, which is thought to be one way in which viruses can induce cancer. And more precise definition of anomalies, in turn, should help to open up an area where at present we know almost nothing: the nature of the genetic imbalances produced by particular anomalies, and the mechanisms by which these produce abnormal growth, whether in utero or in the adult cancer patient. Improved understanding of these things would obviously have important implications, not merely as concerns the basic physiology of growth but for therapy as well. Finally, the more pathological anomalies we can detect and define postnatally, the more we can detect prenatally – and thereby prevent by terminating the pregnancy.

Some of these problems will no doubt succumb, at least in part, to refinements of present methods, but others will doubtless require techniques not yet conceived. Still, considering the rapid and continuing advances in both techniques and knowledge up to date, I should not be surprised if the next 15 years of the "chromosome revolution" were as productive, but perhaps not as exciting as the first 15.

2

Chromosomal Abnormalities II: Sex Chromosome Defects

MALCOLM A. FERGUSON-SMITH
University of Glasgow

The preceding chapter discussed some of the anomalies resulting from expressions of defects occurring in the 22 pairs of autosomes. This leaves to this presentation a discussion of those chromosome disorders that result from defects in the two sex chromosomes, XX in the female and XY in the male. What is remarkable is that these two chromosomes by themselves create more "business" for the physician than all the other 44 together.

This comes about not because the sex chromosomes are notably more prone to anomalies than are autosomes, and certainly not because the consequences of sex chromosome disorders are more severe. Rather, it is the other way around: the sex chromosomes loom so relatively large in the field of genetic disease precisely because their disorders are, by and large, *less* damaging, so that the affected individual survives long enough to require the physician's attention. Considering merely the gross chromosomal anomalies, we note, for example, that loss of a single autosome (monosomy) is seldom compatible with life even in utero; in all but an exceptional case of G monosomy the monosomic zygote does not survive implantation, and perhaps does not achieve it. Monosomy Y is no less lethal – but monosomy X individuals survive not merely into the gestational period but, in a small but significant minority of cases, beyond it and into adult life.

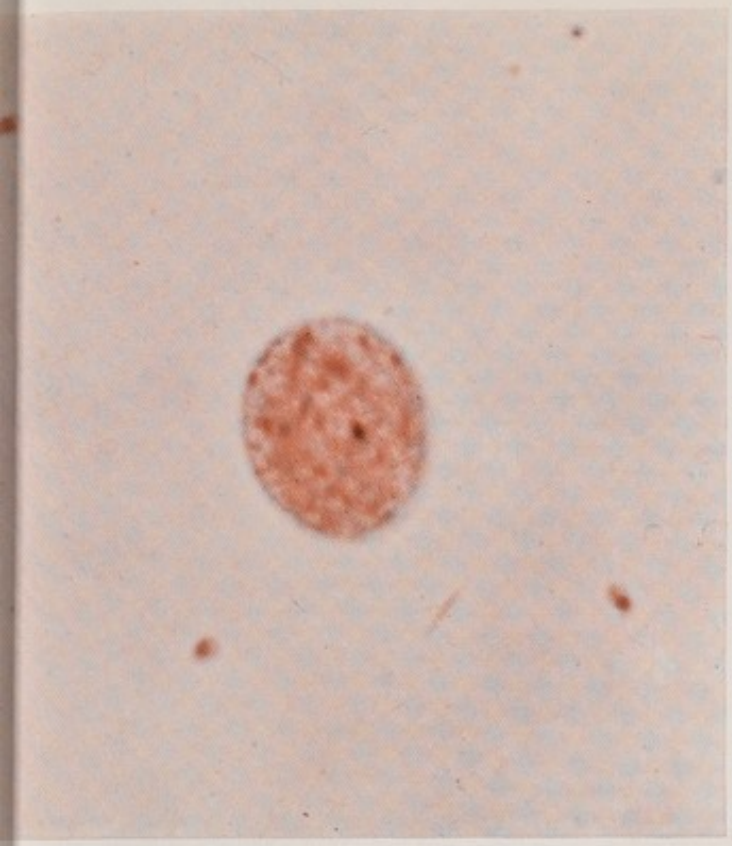
The contrast is even more marked with the trisomies; in the case of the autosomes, the affected individual is invariably severely mentally defective and, with the exception of the G trisomies, most cases perish before birth or shortly thereafter. In sex chromosome trisomies, however, individuals appear to survive to adulthood in many cases without much physical or mental disability; indeed, tetrasomies and even pentasomies, while rare, are not unknown.

The reason for this difference is now completely understandable on the basis that only one X chromosome is genetically active in human cells. Any X chromosome material in excess of this is almost entirely genetically inactive (see Lyon Principle, below).

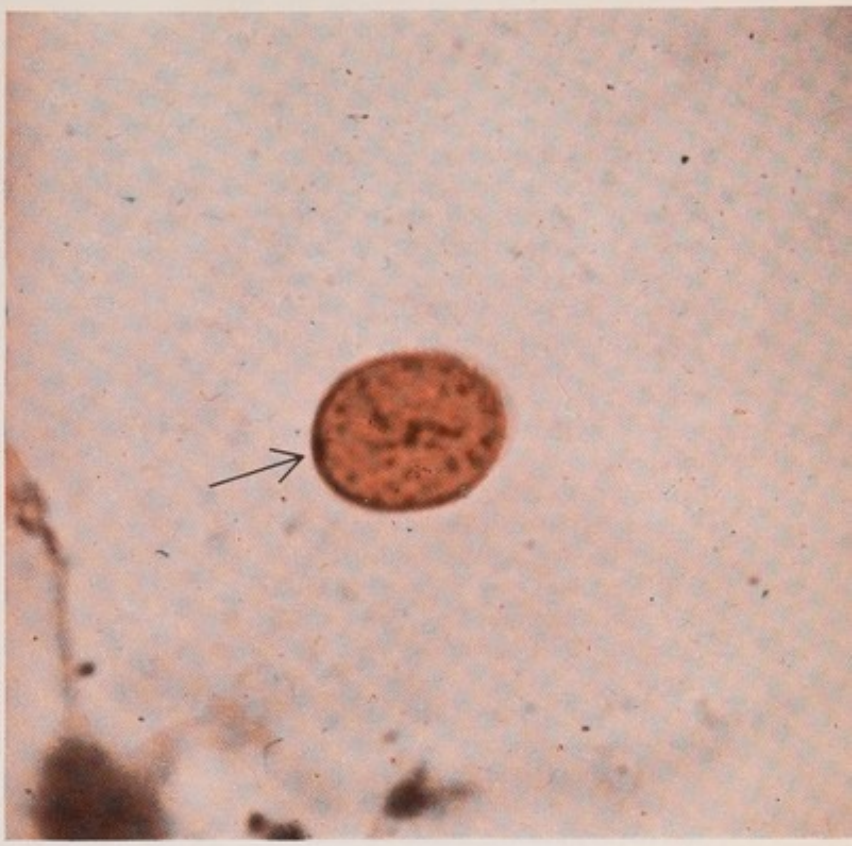
The numerical sex chromosome anomalies – those involving the presence of too few or (much more often) too many normal sex chromosomes – are also the most common. All are thought to represent errors of cell division during gamete formation or early cleavage in the zygote.

During meiosis the normal 46 chromosome complement is halved, one member of each pair passing to each pregametic cell. Since the female gametocyte has two X chromosomes, each of the resulting ova will contain one X; the male gametocyte, with an X and a Y, will produce sperms half of which contain one and half the other. At fertilization, the X sperms engender female zygotes, the Y sperms, males. On occasion, however, through mechanisms which will be described later, the distribution of sex chromosomes goes awry. The result will be an individual with any one of several anomalous sex chromosome complements: monosomy X (XO—Turner's syndrome), XXY (Klinefelter's syndrome), XYY, XXYY, XXX, XXXY, XXXX, XXXXY, and even XXXXX.

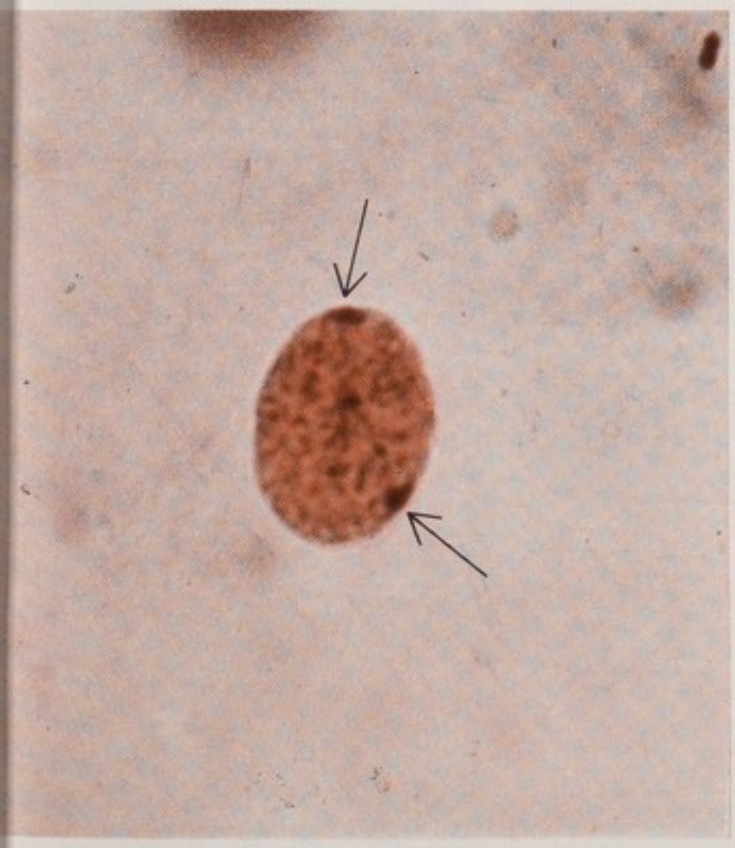
Clearly, the Y variations are much less prolific than X variations; neither YO nor YY cells have ever been observed, and it is presumed that these anomalies are incompatible with life even in utero. One reason is probably the fact that the X plays a much more important genetic role than its male counterpart. It is much larger and, as one would expect, carries many more genetic determinants than the Y. This we know from the many sex-linked con-



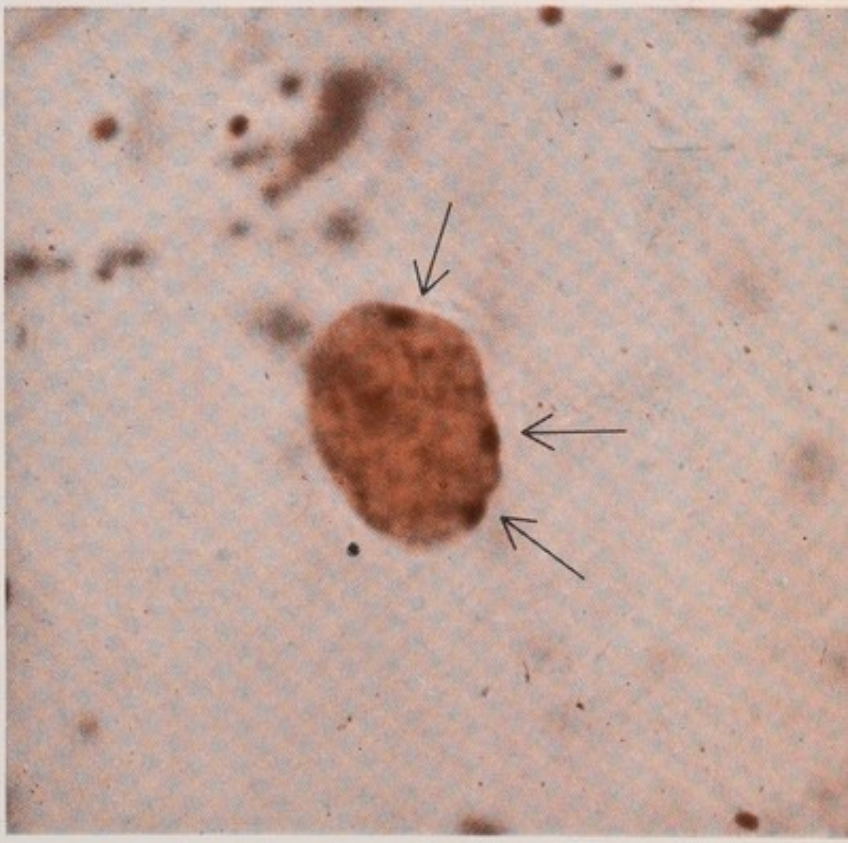
Cells from buccal smears are used for a quick count of X chromosomes. Chromatin negative cell above, from normal (XY) male, has no Barr body, showing only one X is present.



Single Barr body is typical in cells from normal (XX) female. Its small size here suggests partial deletion of one X chromosome, an inference later confirmed by karyotyping.



Two Barr bodies in cell from anatomically normal female show she has XXX anomaly, a condition hard to diagnose clinically since it produces few or no major abnormalities.



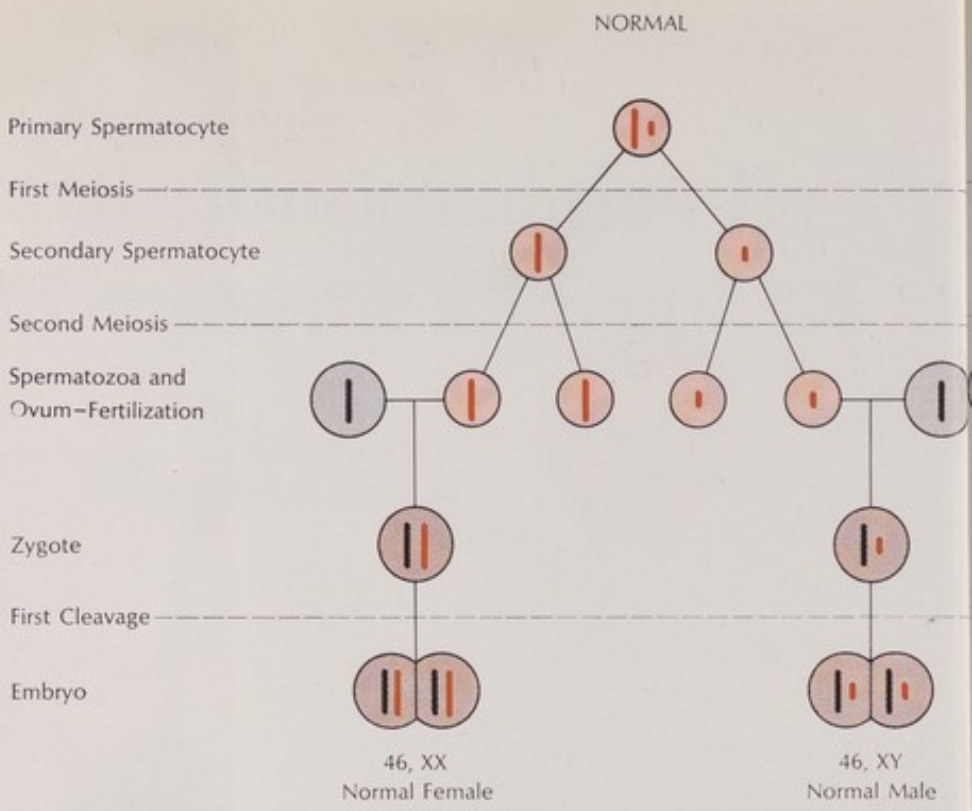
Three Barr bodies show three inactivated X chromosomes in cell from male patient; genotype must be XXXXY. "Invisible" Y chromosome is inferred from individual's gross anatomy.

ditions (color blindness, hemophilia, muscular dystrophy, etc.), the genes for which must, from their 50% transmission from mother to son, but never from father to son, be carried on the X chromosome. Thus the X is evidently considerably more than a mere sex chromosome, since it carries determinants for traits having no connection with the reproductive system. The Y, in contrast, seems to be largely concerned with sex. The search for Y-linked traits comparable to the X-linked conditions just cited has turned up only one possibility: so-called hairy ear, a luxuriant but benign growth of hair on the edge of the pinnae, found principally in Asiatic Indians. This trait is thought by some geneticists, though by no means all, to be transmitted from father to son with the Y chromosome.

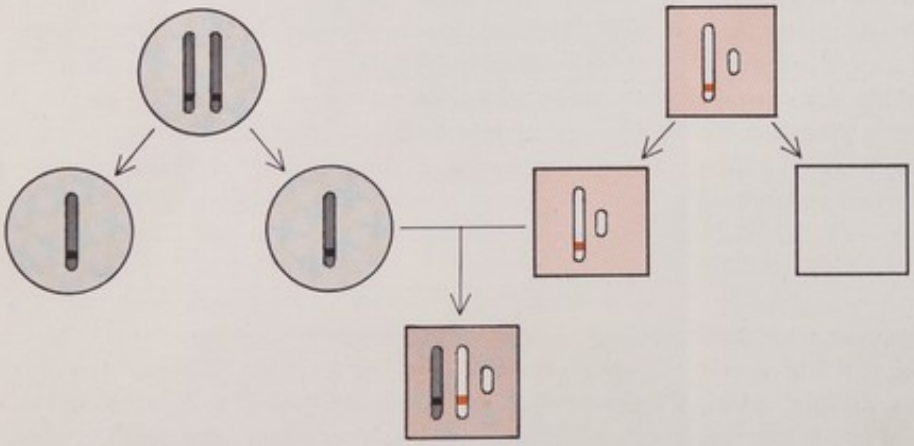
It would appear, then, that while the Y is unnecessary for survival – it is, after all, absent in half the human race – the greater genetic "weight" of the X chromosome makes its absence lethal. But this in turn raises the question of how so potent a chromosome can be present in a double, triple, quadruple, or quintuple dose without destroying the organism.

The answer seems to be connected with one of the most remarkable, most mysterious, and at the same time most useful characteristics of sex chromosomes: X inactivation. In somatic cells, for reasons we still do not understand, only one X chromosome can be biologically active, the remainder being in some fashion "switched off." Nor is this merely a theoretical concept. The substance of the "surplus" X chromosome(s) is condensed into the so-called X chromatin or Barr body (or bodies) which can easily be observed through the microscope as a small mass lying at the surface of the cell nucleus. The X body can be observed following routine histological fixation and staining – most nuclear dyes are satisfactory. A more sophisticated staining technique using quinacrine hydrochloride and examination in ultraviolet light with a fluorescence microscope enables the Y chromosome to be identified in nondividing nuclei. The presence of one or more Y chromosomes is indicated by one or more small brightly fluorescent bodies, termed Y bodies (see Caspersson and Zech, Chapter Three). These derive from the distal segment of the Y chromosome long arm, which is heterochromatic and, like the additional X's, believed genetically inactive.

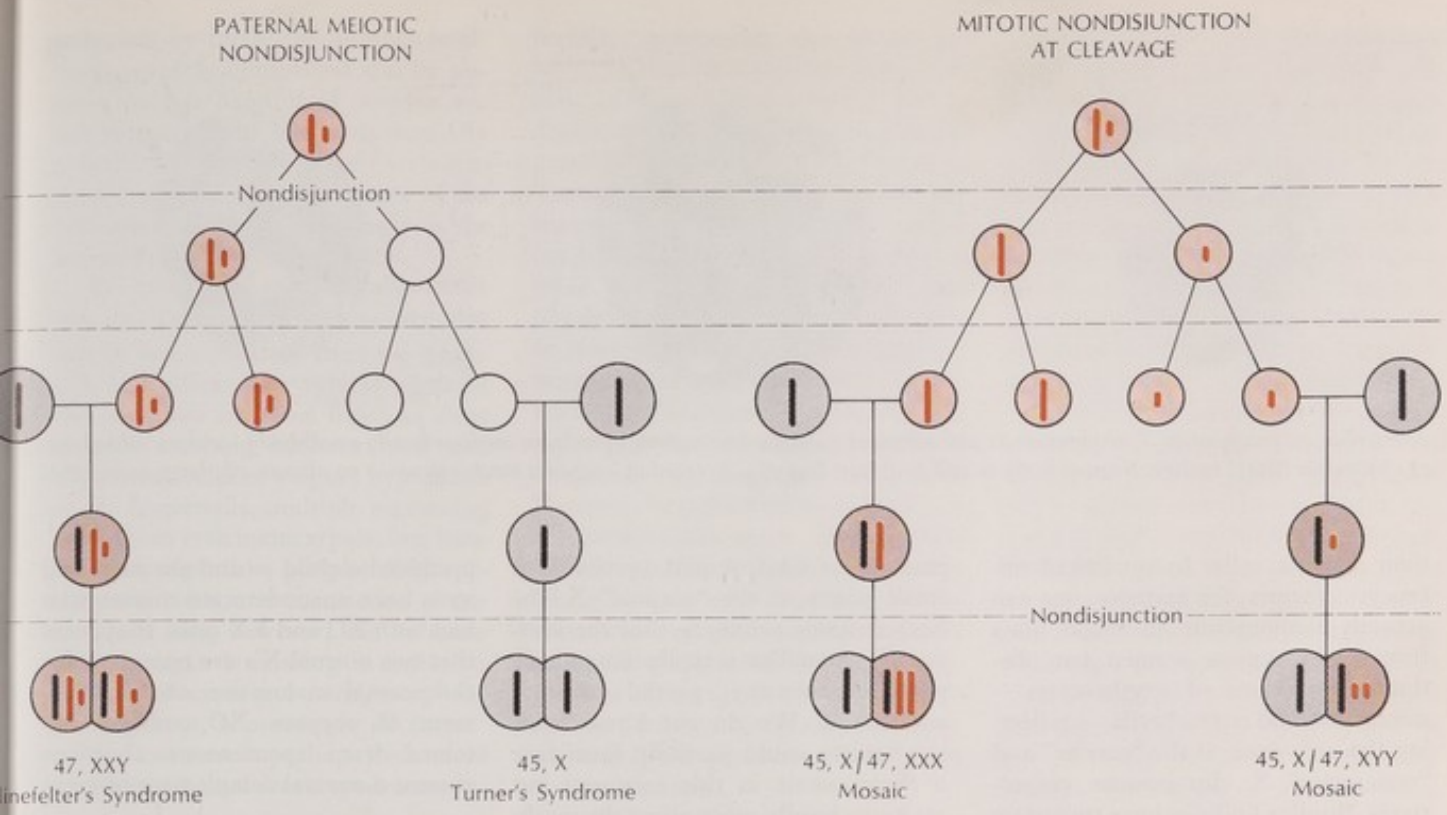
An accurate count of the sex chromosome complement can, in fact, be obtained without the laborious karyotyping required in most chromosome studies; one merely stains the tissue for X and Y bodies. The number of fluorescent Y bodies indicates the number of Y chromosomes in the cell, and the maximum number of X bodies per cell plus one gives the number of X chromosomes. Thus no sex chromatin means XO (Turner's syndrome); one X body means XX (normal female); one X body and one Y body means XXY (Klinefelter's syndrome); two X bodies and two Y bodies means XXXYY; and so on.



Normal meiosis of spermatocytes (above) produces equal numbers of "X" and "Y" sperm that lead to normal offspring. Nondisjunction during meiosis (center) produces



Nondisjunction of paternal gametocyte is proved by inheritance of X-linked blood group. Presence of Xg^a gene (color) in both father and Klinefelter patient shows that the child must have acquired one of his two X chromosomes from the father.



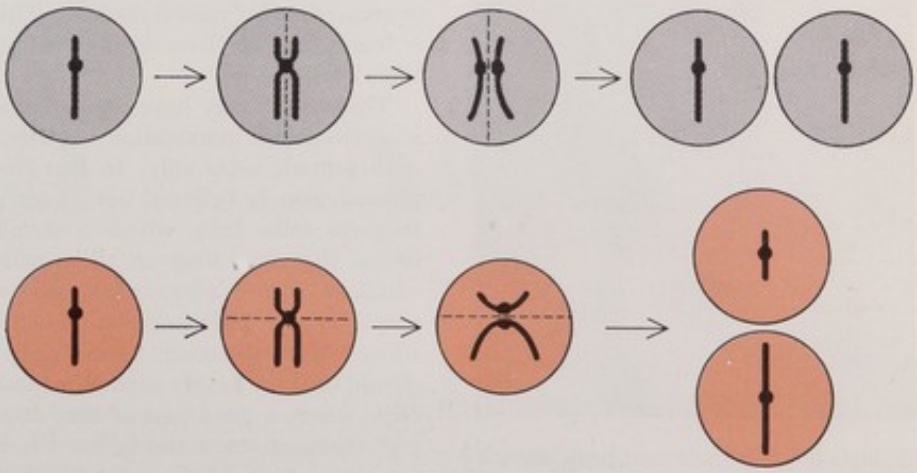
either XXY or XO child; during mitosis after fertilization (right) the result is a mosaic if XO cell line survives, as shown here.

Usually it does not, producing instead a "pure" XXX or XYY individual. Similar processes can occur in oocytes.

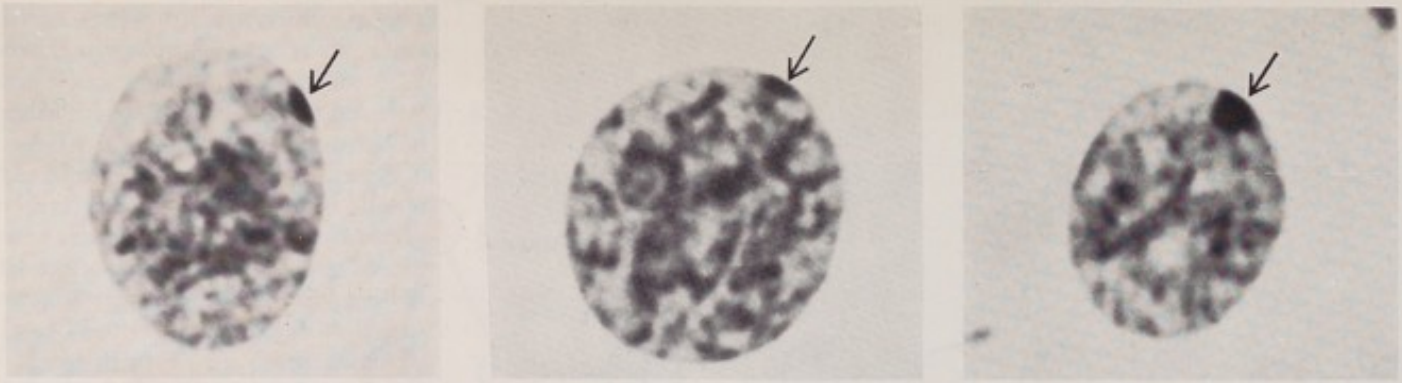
Thus far I have used the standard term "inactivation." A little thought, however, will show that it must be in some sense a misnomer. An XO female can, after all, be distinguished at a glance from a normal XX female, even though both presumably have but one active X; the distinction between the XXY and the normal XY male is also clear, though again both have a single active X.

Inactivation of some sort is surely a reality; too many lines of evidence point that way. There is, for one thing, the fact that with all the other chromosomes (including the Y), monosomy is almost invariably lethal to the zygote. The XO condition, by contrast, while usually lethal, is not immediately or automatically so: it is fairly common in abortuses, and perhaps one XO embryo in 40 survives to term, of which the majority eventually reach adulthood. Again, considering the drastic results of autosomal trisomies (the "least" serious of which is Down's syndrome), one would expect at least severe damage in the XXX individual. In fact, however, few of these women show gross anatomic defects and some show no physical abnormalities whatever. Individuals have survived even tetra- and pentasomies of the X chromosome, whereas with the autosomes such conditions have never been observed. On the basis of this evidence, it must therefore be presumed totally inviable.

The principle of X inactivation, most clearly outlined by Dr. Mary Lyon and thus referred to as the Lyon hypothesis, presumes that X inactivation occurs in cells at random during early embryogenesis. For each cell either the maternal or paternal X is inactivated but, once the decision is made, all the descendants of that cell have the same X inactivated. Females heterozygous for X-linked genes will therefore be expected to show "patchy" expression of those genes in



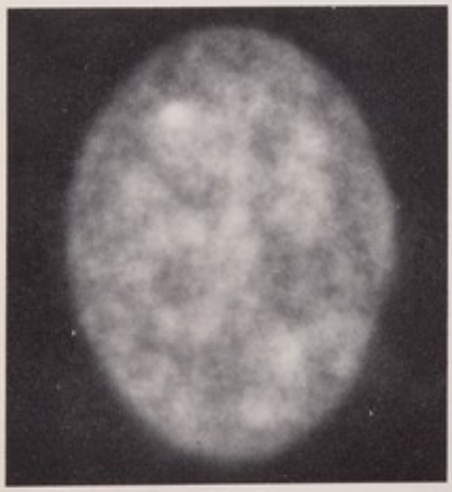
Duplication of X chromosome normally produces two "daughter" X's with one long and one short arm (top sequence). Isochromosomes, with two long or two short lines, are thought to be formed when split occurs at right angles to normal line (bottom).



Anomalies of inactivated X chromosomes are reflected in sizes of chromatin (Barr) bodies. Normal body in cell is at left; loss of material produces an abnormally small body (center); added material in long-arm isochromosome an abnormally large one (right).

their somatic cells. In sex-linked microcytic anemia, for instance, one can actually demonstrate in blood films from heterozygous women two distinct populations of erythrocytes — microcytes and normal cells — reflecting the activation of the "carrier" and "noncarrier" X chromosome respectively. Similar findings have turned up in ocular albinism, choroideremia, and in the so-called manifesting heterozygotes for hemophilia. Similar dual populations have also been demonstrated in cell cultures from carriers of Hunter's syndrome, the Lesch-Nyhan syndrome, and in G-6-PD variants — some cells resembling those from affected males, the rest normal.

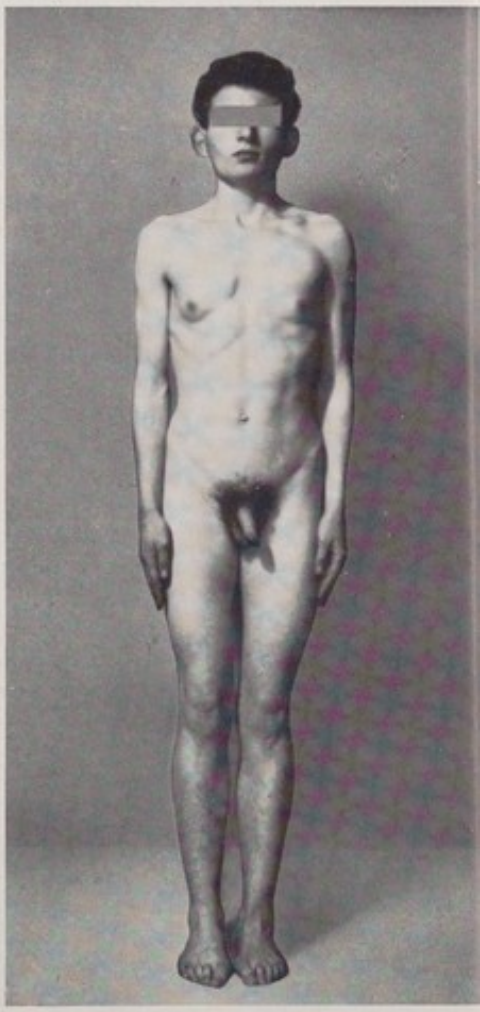
The most plausible explanation for the distinction between XO and XX individuals is that X inactivation is partial, not total. A part (probably a small part) of the "surplus" X (or X's) remains active, so that the various X anomalies actually amount to partial monosomy, partial trisomy, and so on. We do not know what mechanism could partially inactivate a chromosome in this manner – but we know hardly more about the mechanisms that normally and routinely activate and inactivate particular genes in different cell lines of the developing embryo. We do know that such mechanisms must exist to account for the development of structurally and functionally different tissues from cells possessing the same genetic makeup. Obscure though its mechanisms are, partial inactivation does at least offer a plausible explanation for the seemingly contradictory facts just cited.



A single brightly fluorescent body, termed Y body, is seen in the nucleus of a cell from female patient with XY pure gonadal dysgenesis. Nucleus is stained with quinacrine and examined under UV light.

It would also explain another aspect of sex chromosome diseases that I have not yet noted: the frequent contrast between their reproductive and nonreproductive manifestations. This is found in both Turner's and Klinefelter's syndrome.

The reader may have noticed that in discussing X inactivation I ascribed it to somatic cells only. In fact, the phenomenon is believed not to occur in germ cells, from which it should follow that the effect on these cells, which in Turner's syndrome, for example, would lack half their normal active X-chromosome complement, should be more severe than in somatic cells, where a good part of the "missing" chromosome would be inactive in any case. And this is what we find. Such individuals are sterile, as a rule. In the entire literature, I know of only one Turner's syndrome patient who produced a child — and she may well have been an undetected mosaic who had both XO and XX cells. It appears that two normal X's are necessary for the normal maturation and development of oocytes. XO embryos obtained from spontaneous abortions possess a normal complement of germ



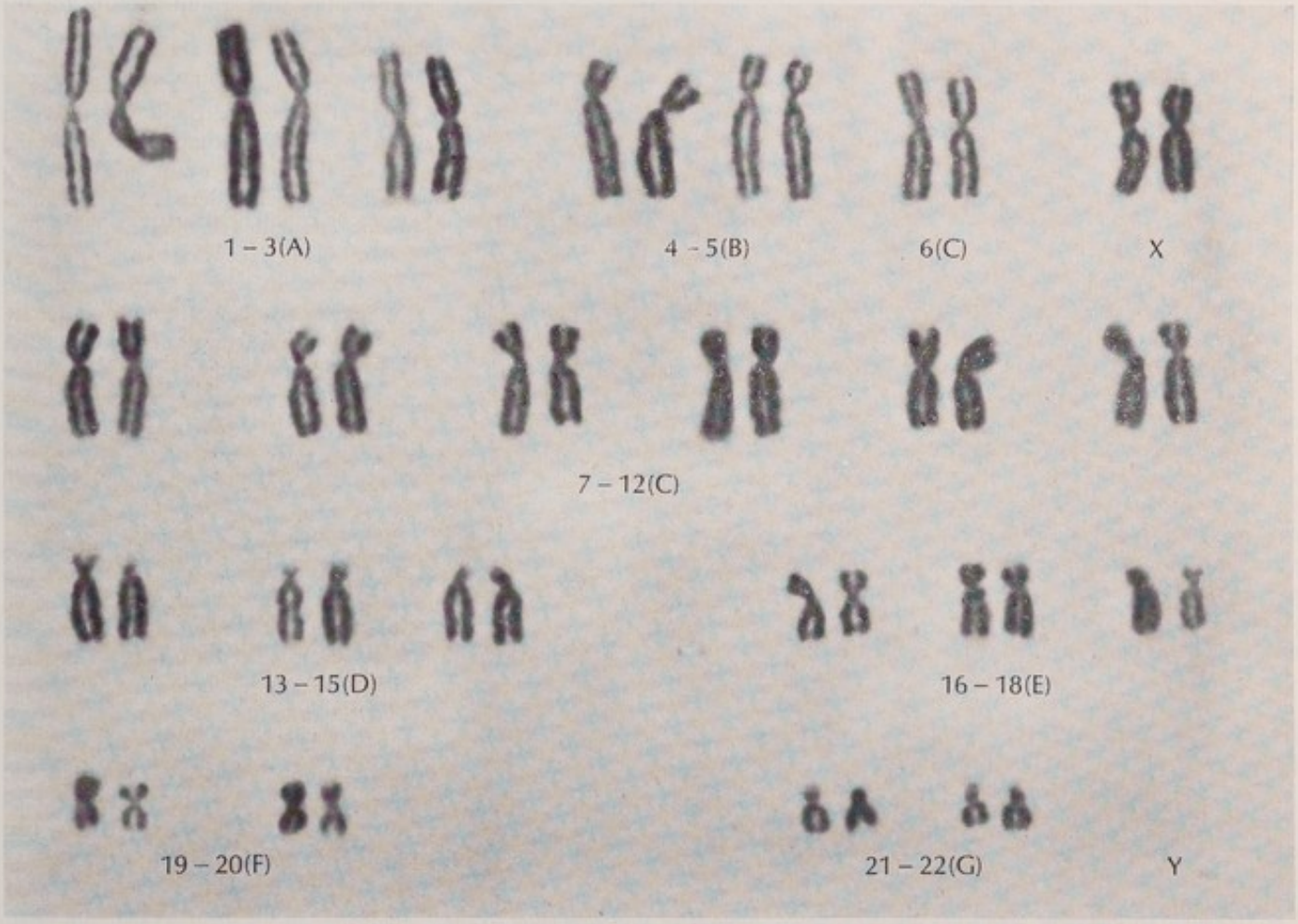
Anatomically male patient has Klinefelter's syndrome, characterized by small

cells, but by birth this number is already severely diminished and by puberty only a handful of oocytes are left as against the hundreds normally present. Similarly in Klinefelter's syndrome, the outstanding feature of the testicular histology at all ages is the severe deficiency of germ cells.

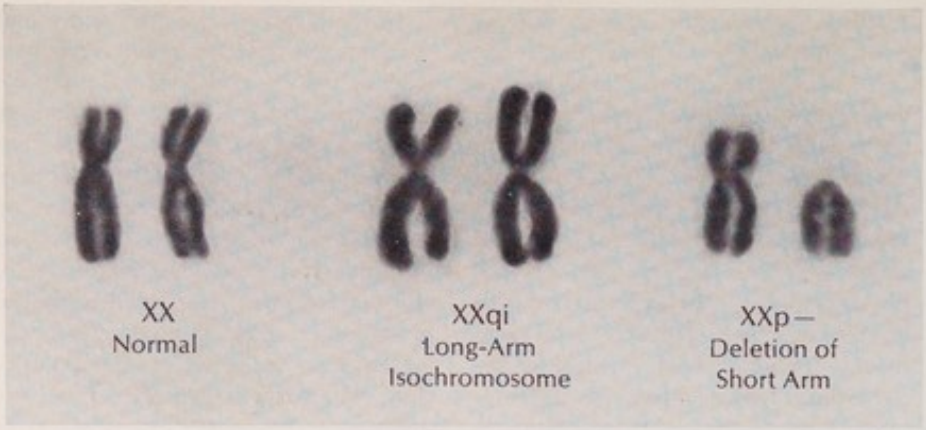
In contrast, somatic abnormalities in Turner's syndrome are distinctly less consistent than the germ-cell anomalies. The typical signs of the condition are often listed as short stature, webbing between neck and shoulders, cubitus valgus, hypoplasia of the fingernails, multiple pigmented nevi, short IVth metacarpals, low hairline, micrognathia, peripheral lymphedema at birth, and coarctation of the aorta. In fact, however, one or more of these signs is missing in a sizable majority of cases and none, except for short stature, is present in all cases.

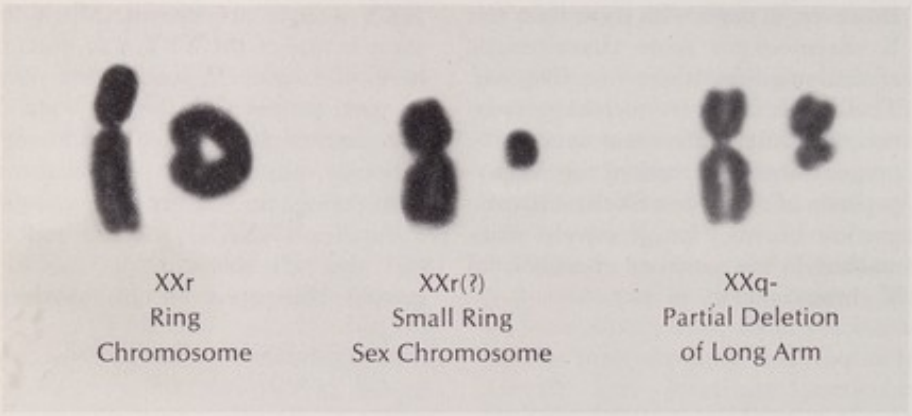
The contrast is perhaps even more marked in Klinefelter's syndrome. Infertility, presumably due to the genetic imbalance resulting from one dose of the Y chromosome and a double dose of X chromosomes in the germ cells, is almost as absolute as in Turner's syndrome. In the somatic tissues, however, thanks to the partial inactivation of the extra chromosome, there is less genetic imbalance and physical and mental abnormality may be minimal. Such signs as occur seem mostly secondary consequences of abnormal sex hormone patterns developing at puberty; indeed the condition is frequently not diagnosed until then. However, in cases with more than two X chromosomes some characteristic clinical manifestations are frequent. These include severe mental retardation, proximal radio-ulnar synostosis, prognathism, and marked tubular hypoplasia of the testes. Skeletal disproportion becomes progressively more marked in the presence of additional Y chromosomes.

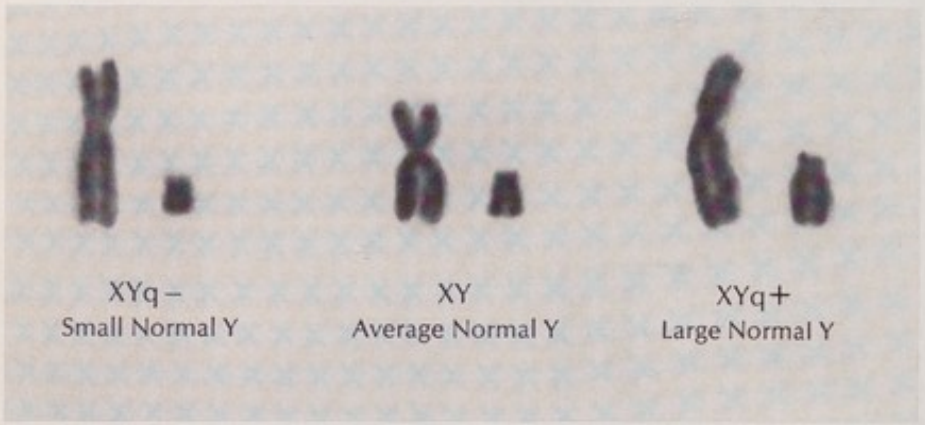
In the XXX woman, the additional chromosome appears to interfere little with fertility – evidently the presence of an additional X is far less serious to the developing oocyte than the absence of one – and its partial inactivation ensures that it will have little or no effect on somatic cells. The reason for this may be that the additional X chromosome may be lost during the proliferation of the oocytes by a mechanism termed selective nondisjunction, which is known to occur normally in certain small mammals. Certainly no XXY or XXX offspring of XXX women are known. Much the same is true of the XYY man, though here, of course, X inactivation plays no part; presumably the additional Y chromosome is too small to derange seriously the genetic mechanisms. Concerning the higher chromosome anomalies (XXXX, XXXY, and so on) the rule seems to be that the greater the excess of chromosomes,

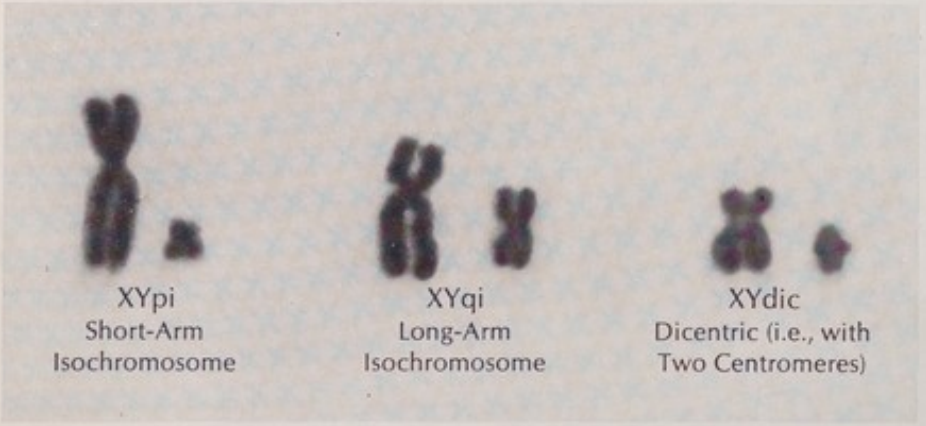


testes and gynecomastia. His karyotype reveals an apparently normal female chromosome complement. Probable explanation is translocation during paternal meiosis, with X and Y chromosomes exchanging material. Thus one X here is presumably X_y.









Morphologic chromosome anomalies are very varied; some have provided clues to the genetic function of the long and the short arms. Anomalies are found in both X and Y; in addition, however, the normal Y has a much greater size range.

the greater the incidence and degree of physical and mental abnormality that will be seen.

Here is perhaps the place to set down some other significant but obscure points connected with the major sex chromosome anomalies. The first has to do with height, one of the few traits that seem to have any consistent association with these disorders. As already noted, XO individuals are abnormally short; it now appears that any excess of either sex chromosome (and particularly the Y) makes for unusual tallness – not, indeed, necessarily exceeding the limits of height distribution in the normal population, but taller than one would expect from the family pedigree. And again, the greater the excess, the greater the abnormality. The reasons for this are wholly unknown. It does not appear to stem from sex hormone abnormalities, since the unusual height usually manifests itself before puberty, nor have studies of growth hormone secretions in these individuals revealed any anomalies that would account for the fact.

At least as puzzling is the matter of mental capacity. Some of the first studies of these conditions purported to show a consistent, or fairly consistent, relationship between sex chromosome anomalies and mental retardation. (An exception was Turner's syndrome, in which such retardation as exists is demonstrable only by rather special techniques; specifically, the sufferers tend to score normal or nearly so on tests of "verbal IQ" but average somewhat below normal on "performance IQ.")

It was quickly pointed out, however, that many early studies had been carried out among inmates of mental institutions and homes for the feebleminded, so that the finding of intellectual impairment was built-in; some even went so far as to declare that the sex chromosome disorders had no influence on mental development. Further research, especially on intellectually "normal" populations, now indicates that the truth is somewhere in between. All the commoner anomalies are compatible with normal mental development. Nonetheless, the incidence of the disorders among mentally deficient populations is distinctly greater than would be produced by

chance alone – four times or more the expected figure.

It is possible, of course, in at least some of these conditions (XXY and XXX, for instance) that the normal-range individuals may be those with a better-than-average genetic endowment whose potential was stunted by the chromosome abnormality; studies of such individuals in relation to their families' mental ability would be instructive but have not (so far as I know) been made. Certainly most sex chromosome retardates are "high-grade," and in no way comparable to the idiots and imbeciles produced by, say, untreated phenylketonuria or autosomal aberrations. Once again, the more extreme chromosomal deviations seem to produce more extreme and more consistent intellectual consequences; some of these deviations have been observed only in severely retarded individuals. With most of these rare disorders, however, the number of cases is still so small that any firm generalizations about them are dangerous.

One should not close this section on mental consequences without mentioning the so-called criminal effects of the XYY condition. A surprisingly high frequency (about 4%) of XYY's was first shown in a group of criminals, and peculiarly violent and incorrigible ones at that. Again, however, further studies have established that the XYY condition is by no means uncommon; in fact, its incidence is now estimated as about 1 in 700 males. If all these individuals – at least 50,000 in Great Britain alone – were violent criminals, the rest of us would not be very safe! Paralleling the case of mental retardation, the XYY condition does seem, for reasons at present unclear, to be more common among violent criminals, but to consider it as some sort of automatic criminal stigma would be wildly at variance with the facts.

Turning now to the mechanisms that produce numerical anomalies in the sex chromosomes, we find, unsurprisingly, that they parallel those observed in the autosomal disorders. The two sex chromosomes of either the male or female gametocyte may in one way or another fail to separate properly at meiosis, producing gametes with neither or with both sex chromosomes. Alternatively, after fertilization a similar anomalous segregation may occur at mitosis, producing one cell line with excess chromosomes and another with a deficiency (i.e., XO or YO), the latter presumably dying off sooner or later. (See illustrations on pages 18 and 19.)

When we come to the factors that engender these anomalous segregations we are on less certain ground. Maternal age might be expected to play a part, as is known to be the case in many autosomal anomalies, but the data support this only in the case of the XXY anomaly. Here, in fact, we have been able to amass some very detailed and refined evidence. In families possessing sex-linked traits (color blindness, the Xg blood group) it is sometimes possible to identify, through study of the inheritance of the trait, the parent in whom nondisjunction of the sex chromosomes occurred. Thus if a couple of whom the mother is Xg (a-negative) and the father Xg (a-positive) produce an XXY son who is Xg (a-positive), he must have received the positive gene via an extra X chromosome from his father (plus, of course, the Y); if he is negative, he cannot have received the paternal X but must have got two X's from the mother. And interestingly, though not surprisingly, in cases where the extra X is traceable to the mother, maternal age seems to be a potent influence, whereas when the father is at "fault," parental age —either maternal or paternal – plays no apparent role in the occurrence of anomalies.

When we consider the sources of the XO anomaly, maternal age is found not to be a factor, which makes for a puzzling contradiction. The same processes that are deemed to produce XXY should with equal probability produce XO. To be sure, the actual relative incidence of the two conditions is heavily biased by the immense intrauterine mortality of the XO embryo, but even allowing for that one would expect a maternal age effect to manifest itself in both or neither. Even more puzzling is the fact that studies by the methods just described show that in some 72% of XO individuals the single X was contributed by the mother, i.e., in nearly three out of four cases the condition was produced by paternal nondisjunction. If nondisjunction occurred with equal frequency in both male and female meiosis, one would expect twice as many $X^{M}O$ as $X^{P}O$ conceptions (YO conceptions are inviable). It is possible that the greater than expected excess of $X^{M}O$ conceptions is influenced by loss of the paternal X at fertilization. Furthermore, even in the minority of XO patients where maternal nondisjunction can be demonstrated, the age of the mothers appears to follow normal distribution patterns. More research is obviously needed here.

It is perhaps worth noting that autoimmune mechanisms are suspected of playing a role in at least some sex chromosome disorders. Autoantibodies to various tissues turn up relatively frequently in XO individuals and their parents. But we are again faced by the contradiction with XXY, where a normal incidence of autoantibodies has been demonstrated. This poses in another form the question of how a mechanism that produces chromosome deficiency in some gametes could fail at the same time to induce chromosome excess in an equal number (or, of course, vice versa).

Where mitotic nondisjunction occurs at a later stage of development, when a normal cell line has already been established, the result may be a mosaic. (Indeed certain types of mosaics –e.g., XX/XO– may be formed even at the first division, one chromosome being "lost.") In fact, the clinical material now on record includes a comprehensive listing of the possibilities – XX/XO, XY/XO, XX/XXY, XY/XXY, etc. The actual phenotypic manifestation of mosaicism depends of course on the proportion of abnormal cells and their location within the body, and in fact a whole series of gradations have been described ranging from the "classic" Turner's or Klinefelter's syndrome to a physically normal male or female with lowered (or even normal) fertility.

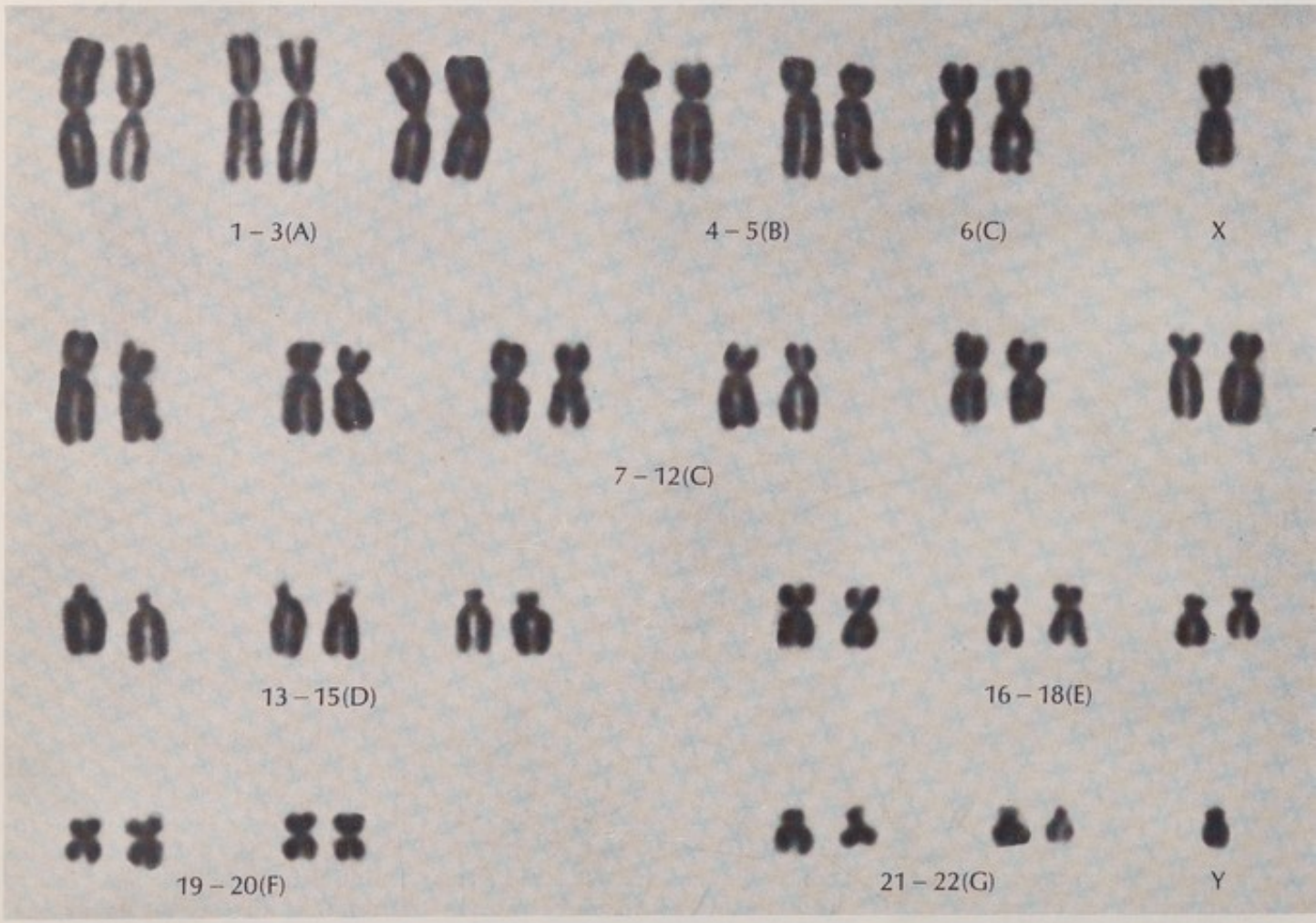
If the mosaic includes one cell line with a Y chromosome and another without it (e.g., XO/XY, XX/XXY) the result may well be an intersex. In XX/XY and XX/XXY mosaicism the individual may be that rare form of intersex, a true hermaphrodite, and have both testes and ovaries. Sometimes the true hermaphrodite will be

"lateral," with an ovary on one side and a testis on the other; in other cases either or both gonads will be an ovotestis. In the case of the XO/XY mosaic intersex, the "ovary" will usually be the "streak gonad" characteristic of Turner's syndrome, showing only a small bit of ovarian stroma without demonstrable germ cells.

One of the most interesting types of intersex is the XX/XY individual. Since it is difficult to hypothesize any very plausible meiotic or mitotic mechanism that would produce two lines of *normal* cells, it is not surprising to learn that many and perhaps all of these persons are chimeras – individuals whose cells are derived from more than one zygote. This most likely occurs as a result of double fertilization of a binucleate egg. Such cases may show two quite distinct sets of blood groups, reflecting two genetically distinct populations of erythrocytes. Interestingly, XX/XY chimeras of this sort have been produced experimentally in mice by fusing blastocysts. The result may not be merely a hermaphrodite; in the case of a black mouse–white mouse fusion, a black-and-white-striped animal may be the product.

These findings would suggest that to get both ovaries and testes in the same individual would require two different cell lines, one with a Y, one without. Against this, however, is the fact that the majority of true hermaphrodites have an apparently normal female XX karyotype. A possible explanation here could be an interchange of material between the father's X and Y chromosomes, possibly as a result of accidental crossing over at first meiosis. This would produce sperm cells containing either X_y or Y_x chromosomes. Fertilization of a normal ovum by the former would produce an individual with two apparently normal X's whose sex chromosome complement was in fact XX_y. Random inactivation of the normal and abnormal X would establish two cell lines that would respectively produce ovaries and testes.

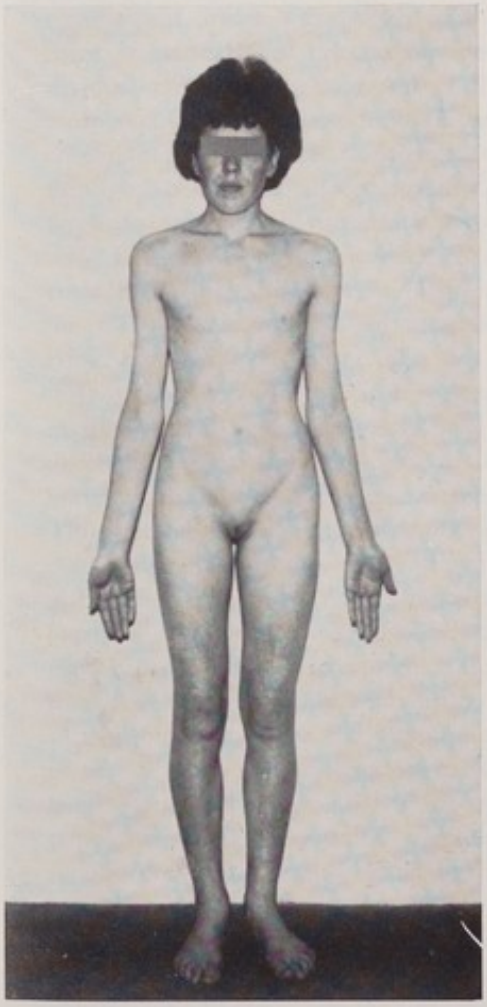
All this may sound highly speculative, but there is some confirmatory evidence, coming from the Xg blood group studies already cited. In one case, an XX-hermaphrodite child of an Xg (a-positive) father was found to be Xg (a-negative), despite the fact that it should have received the gene in its father's X chromosome. A plausible explanation is that interchange in the father moved a bit of the X, containing the blood group gene, onto the Y, replacing the missing bit of X with the portion of the Y containing male determinants. Similar XY interchanges might well explain some other puzzling sex phenomena, including XX men with Klinefelter's syndrome and XY women with pure



Anatomically female patient at right suffers from pure gonadal dysgenesis; karyotype (above) is that of a normal male. In contrast with XX male hermaphrodite shown earlier, this disorder is thought to be developmental rather than chromosomal: if

gonadal dysgenesis.

The matter of interchange takes us into the area of subtler anomalies in which the sex chromosomes are normal in number, consistent with the phenotypic sex, but more or less abnormal in structure. The most important of these are various cases of Turner's syndrome (usually atypical ones) in which both X's are present but one of them is in some way incomplete. This may involve apparent deletion of part of the chromosome or an alternative, more complex mechanism through which the so-called isochromosome is formed. The normal X is submedian (i.e., having a long and a short arm) though not extremely so; the isochromosome X, however, is metacentric, being apparently composed either of two long arms or (very rarely) two short arms. This is thought to come about through misdivision at the centromere during mitosis, with the split occurring at right angles to the usual cleavage (see diagram on page 19) to form gametes with a long-arm and short-arm isochromosome respectively.



gonads fail to form, the result is a female, regardless of chromosomal sex.

Any condition involving loss of even part of the X material produces infertility – the usual streak gonad of Turner's syndrome – thus providing further evidence that two *complete* X's are needed for normal development of the oocytes. Particularly interesting, however, are the somatic differences. The long-arm isochromosome or the short-arm deletion produces the typical Turner's syndrome; somatically, it is as if the loss of the short arm was equivalent to the loss of the entire chromosome. By contrast, the long-arm deletion and the short-arm isochromosome, involving partial or even total loss of the long arm, produce infertility but none of the other Turner's symptoms – short stature, neck webbing, and the like. Thus the somatic symptoms in Turner's syndrome are due to loss of short-arm material.

We can perhaps speculate a bit further. We have noted earlier that X inactivation appears to involve only a part of the chromosome; it is also possible to show that where one X is defective, it is almost always that one which is inactivated; the Barr body is noticeably larger or smaller than usual, reflecting an excess or deficiency of genetic material. (This "preferential" inactivation, by the way, is another example of the organism's remarkable compensatory mechanisms for dealing with physiologic abnormalities.) It would therefore seem likely that X inactivation affects primarily the long arm, the loss of which apparently produces no somatic damage; the short arm would remain more or less active (if present), since its absence produces marked somatic abnormalities.

Similar structural abnormalities have been found in the Y chromosome, and again the isochromosomes are the most interesting. The short-arm isochromosome (i.e., loss of the long arm) produces an anatomically male individual who has the short stature and other somatic features that would, in a female, be pathognomonic of Turner's syndrome. The long-arm isochromosome (i.e., loss of the short arm) produces an anatomically female individual who has no somatic abnormalities. It would appear from this, first, that the male sex determinants are located mainly on the short arm and, second, that part of the long arm of the Y is to some degree genetically homologous with part of the short arm of the X. However, these conclusions must be considered very tentative at present, since they are drawn from observations based on very limited material.

(I might interject at this point that identification of some Y anomalies is seriously hampered by marked variations in the size of the normal Y, which has been found to range from about half the size of a G group chromosome to a size equivalent to a D group chromosome. The more extreme cases, when first observed, were thought to indicate abnormalities, but investigation of male relatives revealed that they are characteristic of male lineages and perhaps even of some regional populations. It is therefore unsafe to reach conclusions about suspected Y anomalies without family chromosome studies. The size of the Y, by the way, does not affect virility.)

Enough has been said, I think, to suggest the breadth and interest of the information we have already acquired from studies of sex chromosome disorders, as well as some of the problems that still await solution. One of the latter that has not been mentioned is whether, and to what extent, these anomalies predispose to disease in later life. There is some evidence, for example, that XO/XY mosaics are at greater risk of developing tumors in their dysgenetic gonads than, say, "pure" XO individuals. Likewise, there are suggestions that XXY (Klinefelter's) patients may be at relatively higher risk of breast cancer (gynecomastia is a common feature) than are normal males. Some of the added risks may well result from hormonal abnormalities, but here too more information is needed. In Klinefelter's syndrome, for instance, testicular biopsies have shown consistent and major abnormalities of the interstitial cells, yet it is still not entirely certain whether the biosynthetic pathway for androgenic hormones is or is not normal.

The solution to these and other problems will doubtless come largely through studies of larger and more

clearly defined populations – that perennial demand of human geneticists – in particular of "normal" populations of newborns and adults. I should also like to make a personal plea for greater depth as well as greater breadth of investigation – and reporting. We have already seen, for instance, the important role of blood groups in advancing our understanding of meiotic nondisjunction, human chimeras, and true hermaphrodites. I would urge that studies on blood and serum factors as well as on sex-linked traits, such as the Xg blood group, in the family pedigree be made wherever possible and incorporated into the published report.

3

Chromosome Identification by Fluorescence

TORBJÖRN CASPERSSON *and* LORE ZECH
Karolinska Institute

Progress in study of chromosomes, normal and abnormal, has been heavily dependent on our ability to identify these structures unambiguously. The modern era in human cytogenetics may be said to have commenced in 1956 with the first accurate enumeration of the human chromosome complement, which was quickly followed by the development of standardized criteria for identifying several individual chromosomes in terms of their size and gross morphology. Another leap forward occurred with the discovery that several other chromosomes could be identified through labeling and autoradiography, which revealed differences in the speed with which their DNA replicated itself. Both these advances were followed by bursts of research – which, however, were limited in each case by the residual ambiguities in chromosome identification (see Hirschhorn, Chapter One).

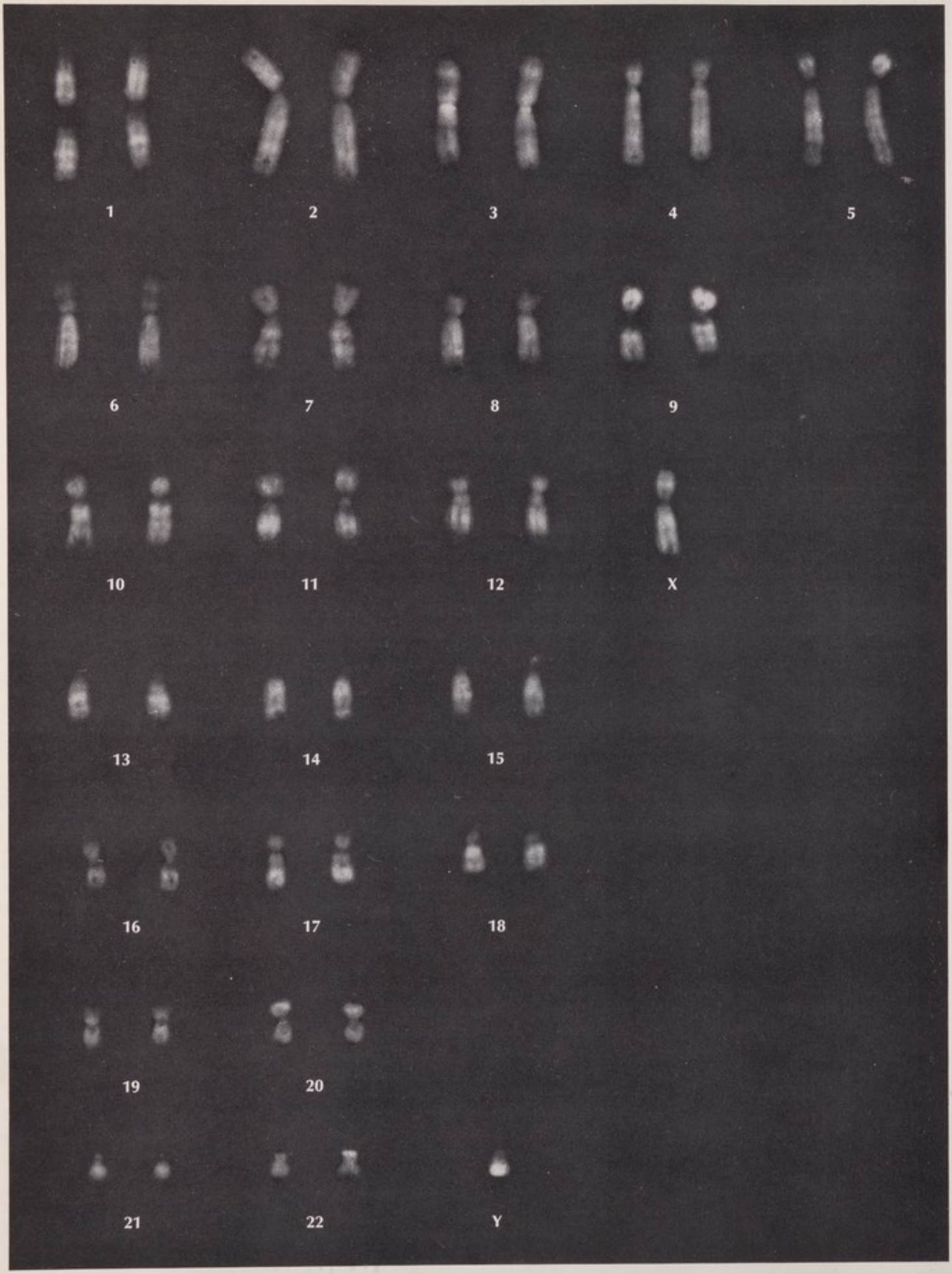
It is therefore with particular pleasure that our group at the Karolinska Institute has been able to report the development of a technique whereby every one of the 46 human chromosomes can be unambiguously identified. This method, involving the staining of chromosomes with certain fluorescent compounds and their examination under ultraviolet light, is very much simpler and quicker (as well as more comprehensive) than the autoradiographic approach. An additional advantage is that it yields its information in a form readily analyzed by computer. This, in turn, promises to open the way to automated processing of chromosome preparations – a capability that has long been recognized as essential if the cytogenetics of large populations is to be studied effectively. The technique, finally, does not merely identify individual chromosomes but also portions of chromosomes – an accomplishment that can be expected greatly to expand our understanding of such anomalies as translocations, in which a part of one chromosome becomes attached to another.

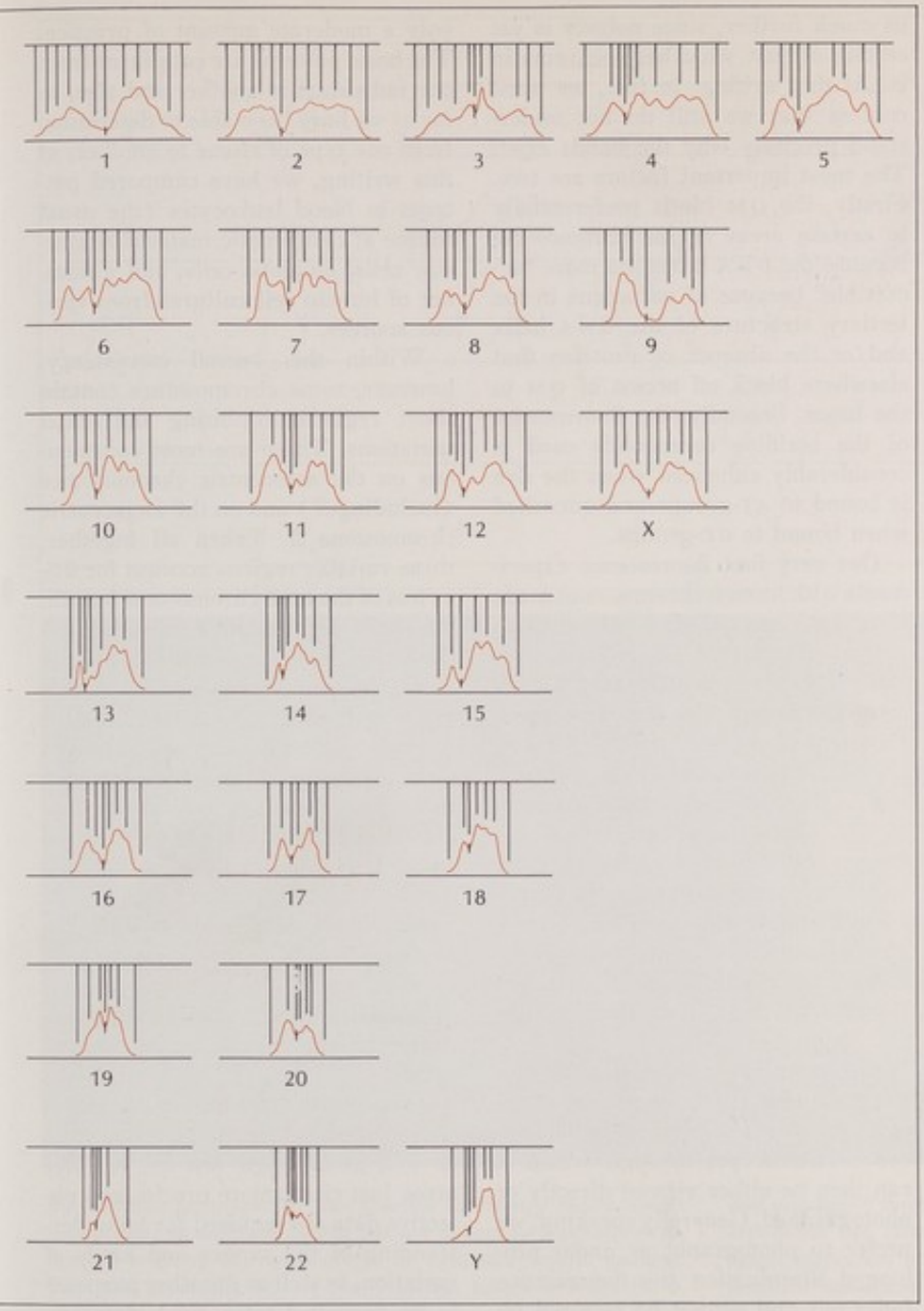
The ultimate implications of this advance will, of course, become apparent only over the course of many years. Already, however, work in our own and many other laboratories has yielded substantial scientific dividends in both clinical and "pure" cytogenetics.

When we began our work, cytologists could identify only 13 of the 24 different types of human chromosomes (22 pairs of autosomes, plus the X and Y sex chromosomes). Seven of these could be singled out in conventionally stained preparations on the basis of their size and the position of the centromere, the junction of the four arms of the metaphasic chromosome; another six could be distinguished by means of the cumbersome autoradiographic method. Regions within a single chromosome could be identified only in very few cases – apart, of course, from the long and short arms.

Our first attempt to resolve this impasse employed the techniques of ultramicrospectrophotometry. This involves illuminating a chromosome and measuring the degree to which the spectrum of the transmitted light is changed by the presence of DNA or by the presence of Feulgen stain, which binds to DNA. Carried out in modern, automatized apparatus, this gives a quick reading of the amount of DNA present. Unfortunately, however, the accuracy of the reading depends on the size of the object relative to the wavelength of the light used; for even 5% accuracy, the object can be no smaller than 1 μm (10^{-6}m) — and the smallest human chromosomes are not much bigger than that. Moreover, the metaphase chromosome contains a high concentration of nonaqueous material, with a consequent high refractive index, producing refraction and diffraction in the transmitted light that aggravate the error. In fact, we found by experiment that this method would satisfactorily identify only the largest of the human chromosomes, giving us little more information than previous techniques.

We next attempted to refine this approach by scanning the length of the chromosome, thus obtaining a curve showing changing DNA concentration from one end to the


1
2
3
4
5
6
7
8
9
10
11
12
X
13
14
15
16
17
18
19
20
21
22
Y



All 24 human chromosome types can be identified unequivocally by treating preparations with quinacrine mustard (QM) stain and examining them under a fluorescence microscope, as can be seen at left in micrographs of the normal karyotype (original magnification x2500), which have been grouped according to conventional numbering (Exp Cell Res *62:491, 1970). The chromosomes are shown in metaphase; the original prints were photographically enhanced to bring out the characteristic, brightly fluorescent QM bands. Such visual examination identifies individual chromosomes and also singles out regions within them, but maximum precision for more detailed study is obtained by recording photometrically the faint fluorescence present along the entire length of the chromosomes. The result is to produce an individual curve for each chromosome, as shown above on this page. These curves of the 24 human chromosome types were enhanced for contrast so that the precise location of even quite faint QM bands could be determined (vertical lines above curves). By this means the curves facilitate identification of regions within the chromosomes for more detailed study* (Exp Cell Res *70:476, 1972). While curves (more than 30,000 examples have been studied) show some variation from one individual to another, in all instances the curves for any one given chromosome all resemble each other more than they resemble those of any other chromosome. The curves also make possible computerized classification of chromosomes and chromosome regions.*

other. In cells of the rye plant, which has seven chromosomes of much the same size and shape, we found that the curves were constant and reproducible – and sufficiently different to distinguish one chromosome from another. Work with Chinese hamster cells gave similar results and also suggested that the curves might be used to characterize particular chromosome regions, which would, of course, have been most desirable. Preliminary studies of human chromosomes gave encouraging results, though again primarily with the larger chromosomes. However, the method was so laborious, and required so much instrumentation, that we decided to try another approach.

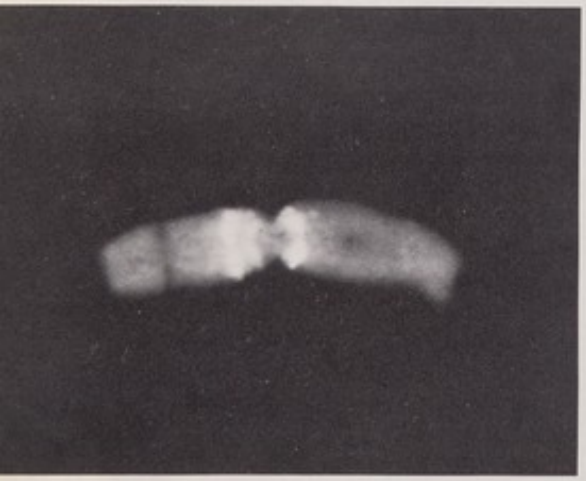
Since we had already determined that the kind of information we wanted would mean working at the very limit of resolution of ordinary light microscopy, it occurred to us that fluorescence might be the answer. It has been known for some years that with self-luminous objects the resolving power with a given wavelength of light is in effect doubled. Moreover, since we would be dealing with emitted rather than transmitted light, the problems of refraction and diffraction cited above would be largely eliminated. Finally, whereas spectrophotometric scanning involves taking readings of a large number of points, which then must be summed up, fluorescence measurement would involve merely determining the total quantity of light emitted by a particular chromosome region, which can be accomplished with much simpler equipment.

Preliminary experiments using small drops of a solvent containing DNA combined with suitable fluorochromes showed that it was possible to detect fluorescence from as little as 10^{-16} gm of the nucleic acid – a sensitivity a thousand times greater than with the spectrophotometric method, bringing us to within one or two orders of magnitude of the amount of DNA in a single gene. We then began looking for a stain that was fluorescent and would bind to DNA – specifically, to its base residues. We chose this way because conventional stains, which bind to phosphate or carbohydrate groups, had shown no significant details along the length of the chromosome, while there was reason

to think that a base-binding stain would do so.

In collaboration with Dr. Edward J. Modest of the Children's Cancer Research Foundation, Boston, we selected for our first trial the compound quinacrine mustard (QM), which we believed would bind firmly to one of the nitrogen atoms in the nucleotide base guanine. When we stained karyotypes of the broad bean plant (*Vicia faba*) with QM and examined them under ultraviolet light, we immediately saw very distinctive banded patterns on the chromosomes; work with some other plant species yielded similar results. The bands do not show up in all chromosomes of all plants, but when they do they are strictly characteristic of the particular chromosome.

It naturally occurred to us to wonder what these conspicuous bands – the brightest of them is up to three times as bright as the dark regions – indicated. Comparison of them with curves for DNA distribution, obtained by the spectrophotometric method previously cited, made clear that they were *not* indicating DNA concentration, since this was relatively even from one end of the chromosomes to the other. Other experiments revealed a certain parallelism between the distribution of the bands and that of heterochromatin, but this did not get us much further, since nobody is yet certain of just what heterochromatin is. At this writing, in fact, we must confess that we still do not understand precisely why the bands exist. The most important factors are two. Firstly, the QM binds preferentially to certain areas of the chromosome because the DNA bases are more "accessible" because of variations in the tertiary structure of the DNA helix and/or the absence of proteins that elsewhere block off access of QM to the bases. Secondly, the fluorescence of the acridine compounds used is considerably enhanced when the dye is bound to AT-groups and quenched when bound to GC-groups.



The first experiments in fluorescence identification of chromosomes were made with plant material. Shown above is the largest chromosome of the broad bean plant (Vicia faba), *which has several very distinct QM bands; measurement showed fluorescence of these bands to be 2 to 3 times more than in other parts of the chromosome* (Exp Cell Res *49:219, 1968*).

Our very first fluorescence experiments with human chromosomes made clear that they, too, showed distinctive individual patterns. The technique we worked out (to anticipate a little) involves staining with QM at a concentration of 50 μg/ml for 10 to 30 minutes, at a pH of 6.5 to 7. The preparation is then washed in three changes of buffer solution, sealed with a coverslip, and stored in a refrigerator until one is ready to examine it. (Similar techniques can be used with several other fluorescent compounds. However, of the 20-odd we have tried, QM remains the best from almost every standpoint.)

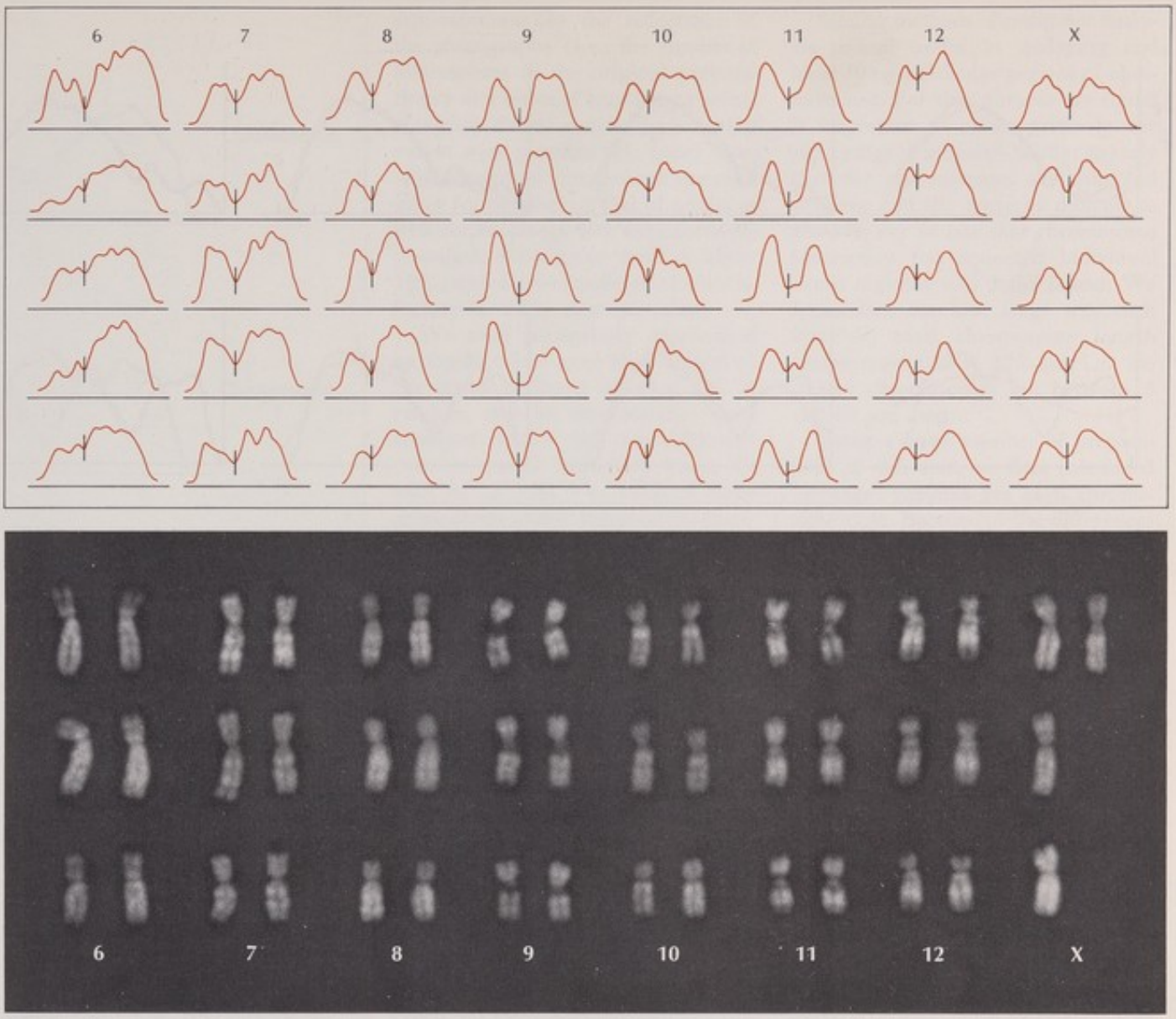
For examination, the slide is illuminated with UV light on the stage of a special UV microscope, a fluorescence microscope, through which it can then be either viewed directly or photographed. Generally speaking, we prefer to photograph, as under prolonged illumination the fluorescence fades out and cannot be restored by restaining. (This disadvantage can be obviated by other staining methods to be described later; these, however, have disadvantages of their own.) While the patterns show up in both negative and positive prints, the latter are generally preferable, since by using suitable copying paper the pattern contrast can be enhanced.

The actual patterns are almost impossible to describe in words. However, careful examination of the photographs illustrating this article should make clear that they are quite adequate for chromosome identification. In our experience, any cytologist or cytologic technician can learn to identify chromosomes by this method with only a moderate amount of practice. The basic patterns are consistent from one individual to another and also, so far as we have been able to determine, from one type of tissue to another; at this writing, we have compared patterns in blood leukocytes (the usual source of cytogenetic material), marrow cells, amniotic cells, and a number of human cell cultures from various sources.

Within this overall consistency, however, some chromosomes contain short regions exhibiting individual variations. These are most conspicuous on the acrocentric chromosomes (including Y) and on the metacentric chromosome 3. Taken all together, these variable regions account for 2% or less of the total chromosome length, which is far too little to introduce confusion into the identification process. They are, however, of potential value in themselves – notably, in the forensic determination of kinship, since all our experience indicates that individual variations are inherited in accordance with simple Mendelian rules. It is worth noting in this connection that variation occurs not merely between individuals but even more commonly between homologous chromosomes of the same individual – presumably reflecting differences between paternal and maternal chromosomes.

While visual examination serves well enough to identify chromosomes, as well as to single out the variable areas just cited, more precise and objective data are required for an understanding of the ranges and kinds of variation, as well as for other purposes to be described later. Such data can be obtained by measuring variations in light output along the length of the chromosome with a technique called fluorometry. For maximum precision, one must use the chromosome preparation itself, but for most purposes it is sufficient to use a photomicrograph, and the necessary apparatus is much simpler.

The micrograph – either positive or (occasionally) negative – is placed on a motor-operated microscope stage that is illuminated either from below (negative print) or above (positive). The light from the micrograph is focused on a slit diaphragm, so arranged as to be equivalent to a 0.2 x 2.5 μm rectangle in the preparation.



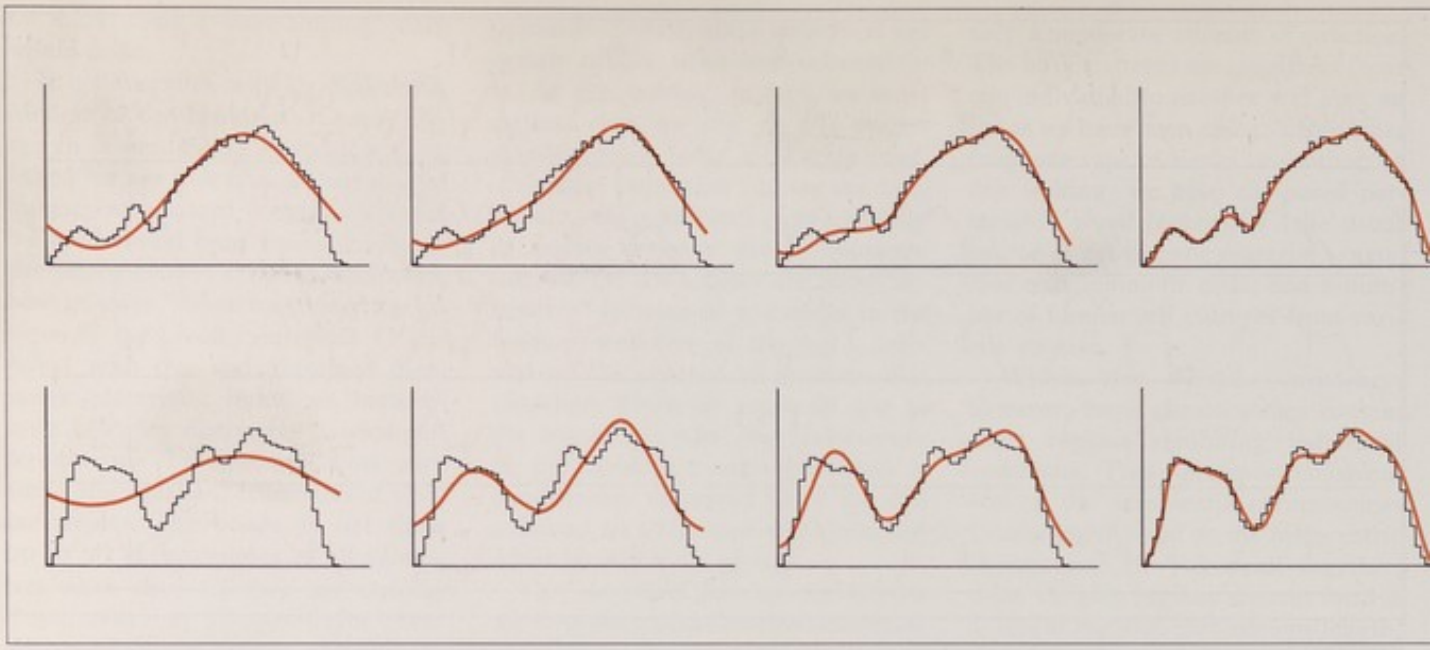
By previous techniques it was possible to classify chromosomes into groups of different length but not always to separate members of a group by morphologic or autoradiographic analysis. This was especially true of the C group, which includes the X and numbers 6 through 12. As can be seen above, such individual separation can now be accomplished by fluorescence, both by visual inspection of the QM bands and, even more effectively, from the associated fluorometric curves (Hereditas 67:97, 1971).

By moving the stage, the slit can in effect be moved along the entire length of the chromosome; by means of a beam-splitter, part of the light is passed into an eyepiece within which is marked the orientation of the slit, so that the operator can adjust the photograph for proper "scanning." The remainder of the light is passed into a photomultiplier and amplifier, producing an electric current that controls the Y-input of an ordinary XY recorder, the X-input being controlled by the motion of the stage.

In certain cases – e.g., if the chromosome varies considerably in width – we use a circular diaphragm with an "equivalent" diameter of 0.2 μm, which is moved along the length of a single chromatid. The result in either case is a curve showing the variations in light intensity along the entire length of the chromosome or chromatid. If desired, the machine can also be hooked up to a digitizer and tape-puncher, producing a series of numerical values that can be fed directly into a computer.

Extensive studies of these curves, based on more than 30,000 examples, show that they are highly reproducible and consistent. As one would expect from what was said earlier about the fluorescence patterns themselves, there are variations in the curve of a given chromosome from one individual to another and often between two paired chromosomes of one individual, but despite these variations the curves for a particular chromosome are all clearly more like one another than like any other curve. Moreover, the curves possess enough detail to permit characterization of regions within chromosomes.

Before saying more about the curves and what they can be used for, we should say a word about some other staining techniques that are now coming into use for chromosome identification. There are a consider-



Because of the difficulty in distinguishing the chromosome types in human group C, Fourier analysis (also sometimes called harmonic analysis) was applied to extract significant features of the directly derived fluorescent functions and help establish that the different patterns of the individual chromosomes were statistically significant. Shown above is the application of the approach to chromosomes 6 (top) and 7, both in C group. The photometrically determined patterns are seen as "staircase" curves together with

able number of these, but nearly all of them involve pretreating a karyotype preparation in some fashion (with saline, NaOH, heat, etc.) and then staining it with Giemsa stain.

These techniques produce banding patterns very like those produced by fluorescence staining, though differing in certain consistent respects, and in general seem to be quite as reproducible as the fluorescent bands. They have the advantage that they are permanent – they do not fade with illumination, as do the fluorescent preparations – and can be examined in an ordinary light microscope, as against the specialized and expensive fluorescence instrument. This naturally represents a considerable advantage for routine clinical work. As against these advantages, the resolution is only half that of the fluorescent image, and the Giemsa preparations do not lend themselves well to conversion into the intensity curves just described. But the field of chromosome staining is advancing so rapidly that we may well see before long a method that combines most or all the advantages of the fluorescence and Giemsa techniques. Meanwhile, however, subsequent observations cited in this article will refer to fluorescent preparations unless otherwise stated.

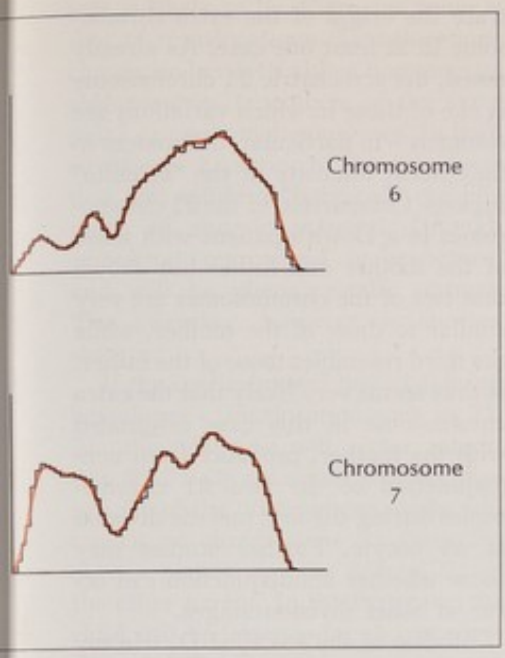
We should also mention, more or less in passing, that the fluorescence patterns found in ordinary somatic cells during mitosis can also be detected in the bivalents of the germ cells during meiosis. Since the chromatin in the latter is more diffuse, the patterns are less clear and details are lost, so that while some bivalents can always be picked out, others, with more subtle patterns, can only be identified in some cells. Nonetheless, insofar as the patterns can be distinguished, they are the same as those found in somatic cells. More certain identification of individual bivalents has come with the addition of another staining technique that permits identification of the centromere in these cells.

Perhaps the most significant implication of these advances is what they tell us about the distribution of chiasmata – the areas in which the two parts of the bivalent join to exchange genetic material. It turns out that these crossings do not occur at random but rather are far more frequent in certain chromosome regions – in particular, the ends of the long arms – than in others. This finding should eventually prove useful in mapping the location of human genes on specific chromosomes, a study still in its early infancy. In addition, the capacity to identify bivalents is proving of more immediate value in the study of chromosome abnormalities.

Having achieved a means of recording and characterizing chromosome patterns objectively, we then sought methods whereby the patterns could be compared and analyzed automatically, by computer, which is the obvious approach to statistical analysis of variations in specific patterns and to the long-sought goal of automated chromosome sorting. We have also found that certain computer techniques can simplify the classification of chromosomes and (especially) chromosome regions.

As already indicated, the curves obtained from chromosome photographs can be "digitized" into long sequences of numbers. However, raw data of this sort are of little use in comparing different patterns, the more so because the numbers do not represent absolute values of fluorescence but rather values relative to a more or less arbitrary baseline. Moreover, this baseline may vary from one chromosome or chromosome preparation to another. What we were interested in was patterns – the position of the inflection points on the curves, the relative magnitude of the peaks and troughs, and so on.

It seemed to us that the most profitable approach would be via the mathe-



the smooth curves that evolved as the computer utilized (in order from left) 1, 2, 3, 6, and 8 harmonically related waves (Hereditas *67:105, 1971).*

matical technique known as Fourier analysis, which can be performed by computer, using the digitized data as raw material. This method depends on the mathematical fact that any curve can be approximated, as closely as desired, by a series of sinusoidal waves that are harmonically related (meaning that their wavelengths are in the ratio 1, 1/2, 1/3, 1/4, 1/5, etc.). Putting it another way, a curve can be defined to any desired degree of precision by a series of number pairs, the first of each pair representing the amplitude of the harmonic and the second, its "phase angle" relative to the initial point of the curve. (A complete "cycle" of a sinusoid is defined as 360°, so that the phase angle defines the point on the cycle that coincides with the beginning of the curve under analysis.) We found, in fact, that all of our curves could be adequately described by eight complex Fourier coefficients, that is, by a sequence of 16 numbers. For more precise definition, we added information on the length of each chromosome (for analytical purposes these had been mathematically equalized).

After some preliminary experiments, we tested this analytical technique on a collection of 912 chromosomes. To simplify acquisition of the raw data, we arranged to measure semiautomatically the reflectance of the photographs (i.e., the amount of fluorescence in the original preparations) at a series of equidistant points along the chromosomes; the resulting values were punched into paper tape, which was then run through the computer for processing. Out of a total of 912 chromosomes 802 were correctly classified, for a gross error of about 12%, which we considered very satisfactory for a first approximation.

We were particularly encouraged by the fact that more than a third of the errors occurred more or less at random, that is, chromosomes were confused that in fact were different even on casual inspection (e.g., 14 with 8, 15 with Y). Most of these gross errors could doubtless be eliminated by including additional data on centromere position. Another source of error was the fact that the chromosomes were classified individually, as they had to be in this heterogeneous collection of photographs. In normal karyotyping, however, a computer program could take account of the fact that no more than two chromosomes of a given pattern will occur – e.g., it could in each case pair the two *most* similar patterns and reject apparent correspondencies with others.

The main reason for the errors, however, is the fact that in the actual preparations small differences in the degree of contraction can occur in different parts of the same chromosome. The bands, though still present, may be somewhat shifted in position. We believe that this factor can be corrected for by more sophisticated programming. Altogether, we are convinced that the accuracy of computer sorting can be very substantially increased, and that normal karyotypes can be classified with a high degree of accuracy. Implicit in this is the expectation that the computer method can automatically and rapidly single out aberrant chromosomes, or at any rate suspiciously deviant ones, in large numbers of karyotypes. To this end experiments are being run with an automatic microscope and film scanner developed at the Jet Propulsion Laboratory in Pasadena. Preliminary results with the C-group chromosomes (the hardest to distinguish by conventional methods) are promising, and we are now moving on to work with whole karyotypes.

Similar methods of computer analysis proved useful in analyzing and identifying particular regions of chromosomes. For this purpose we found it important to select precisely the right stage of mitosis – neither so early that the chromosomes are thin and winding and the patterns difficult to observe, nor so late that chromosome contraction has squeezed individual bands together and detail is lost. We found that the best stage was such that the total chromosome length ranges from 100 to 200 μm, i.e., the largest chromosomes are between 5 and 10 μm long.

Using a large number of preparations of this sort, we first calculated "average" patterns for each chromosome type. Because of the differences in contraction within a single chromosome, already cited above, this gave us only a rough picture of the banding pattern. We then subjected these averages to mathematical contrast enhancement by computer; this technique, by exaggerating the inflections of the curves, brings out weak bands that would otherwise be masked by adjacent, stronger bands. By comparing the contrast-enhanced individual patterns and superimposing the prominent bands on the average pattern curves, we were able to map the location and expectable strength of all the consistently observable bands; these can now serve as reference points for identifying different chromosome regions. At the same time, we were able to characterize more precisely the narrow areas on some chromosomes where individual variability is expectable.

We have also been able to put electronics to work in another way, in examining not the pattern curves but the chromosomes themselves. This should be particularly useful in situations where it is necessary to examine a large number of karyotypes, for example in studying the effects of mutagens or carcinogens. For this purpose we feed light transmitted through a fluorescence negative (or a Giemsa-stained preparation) into a TV camera. The resulting picture is displayed on a screen in arbitrarily chosen enlargement and its contrast can be enhanced precisely as with a home TV set. Moreover, brightness, contrast, and size can be equalized from one chromosome or preparation to another; these qualities,

as already noted, vary appreciably in the original preparations. We can fit the camera to a microscope if we wish to observe the original preparation. The picture can also be scanned electronically along any chosen line. The result, displayed on an oscilloscope, approximates the fluorescence patterns already cited. Finally, with the addition of a second camera, two chromosomes (e.g., members of the same pair) can be displayed side by side, thus further improving the possibilities of detecting small aberrations in the patterns.

Having recounted at some length the methods of fluorescent chromosome identification and some of the technical advances it has engendered, it is time to discuss the practical results. In summary, these techniques have enabled researchers to identify extra chromosomes specifically, instead of assigning them to a group of similar chromosomes; to spot minor anomalies more certainly; and to define structural rearrangements of the genetic material more precisely. The following selection of findings from our own and other laboratories suggests some of the current accomplishments and future possibilities.

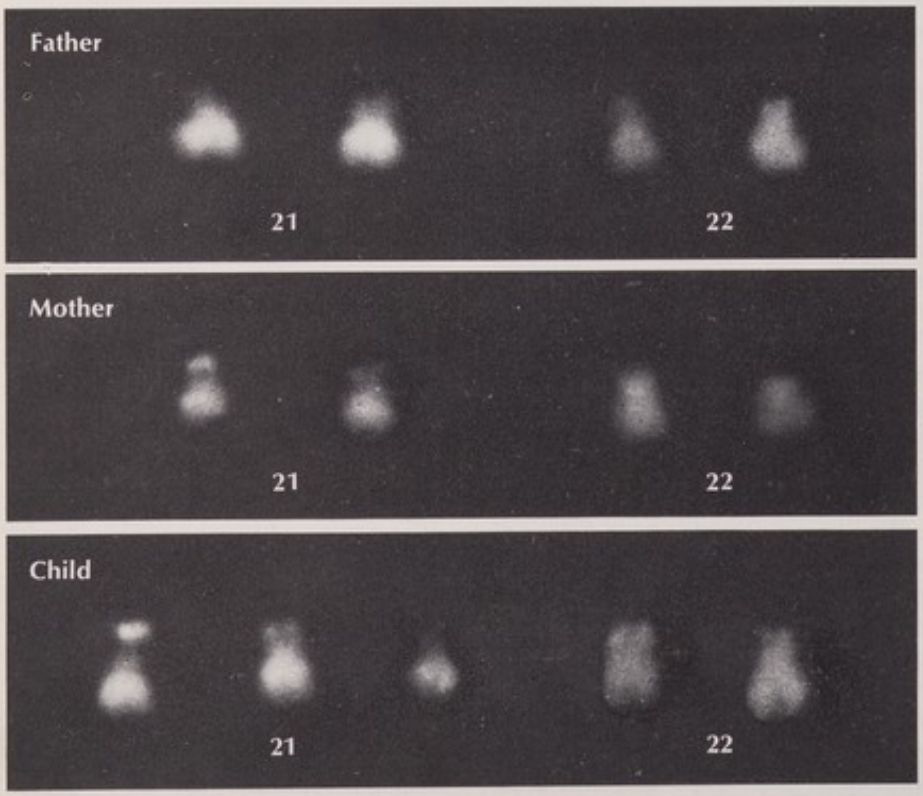
It has been known for many years that patients suffering from Down's syndrome have an extra chromosome in the G group, but it was impossible to be certain whether this was a third chromosome 21 or a third 22, which is morphologically very similar. Conventionally it was called 21. Studies of fluorescent- and Giemsa-stained preparations from patients with Down's syndrome, in our own and other laboratories, have now confirmed that the chromosome is the smaller of the G group chromosome pairs. Strictly speaking this is chromosome 22, but so firmly fixed in people's minds is the "fact" that Down's syndrome is trisomy 21 that the numbering will be changed. It had already been suspected that the smaller G-group pair was involved in Down's syndrome, but the evidence of the banding patterns makes it certain. Variations in fluorescence have also made it possible to trace the origin of the extra chromosome in at least one case. As already noted, the acrocentric 21 chromosome is one of those in which variations are common – in particular, differences in fluorescent intensity of the "satellite" regions. Comparison of the 21 chromosomes in a Down's patient with those of the mother and father has shown that two of the chromosomes are very similar to those of the mother, while the third resembles those of the father. It thus seems very likely that the extra chromosome in this case originated with the mother, probably from nondisjunction of the two 21 chromosomes during the first meiotic division of an oocyte. Further studies may show whether nondisjunction can occur at other division stages.

Studies of the so-called D_1 trisomy have confirmed the belief, also based on autoradiography, that the chromosome involved is in fact number 13 (rather than 14 or 15), whether the defect is a "regular" trisomy or a so-called Robertsonian translocation, in which the two acrocentric chromosomes are fused into a single body.

We have also been able to throw a certain amount of light on several anomalous cases in which apparent G trisomies showed a clinical picture not typical for the well-known Down's syndrome. Examination of two of these cases revealed that the extra chromosome, though morphologically very similar to both 21 and 22, in fact resembled neither of them in its fluorescent pattern. Apparently it is a structurally abnormal chromosome produced by some rearrangement of the hereditary material, though the details remain to be determined.

Fluorescent analysis has clarified the nature of a congenital condition hitherto described only as C-group trisomy. While a variety of data had made clear that the extra chromosome was not the X chromosome, it had not been otherwise identified. Of five such cases that have been subjected to fluorescent analysis, all show an extra, though apparently otherwise normal, chromosome 8. The clinical signs show certain similarities – mainly skeletal abnormalities and moderate retardation – in some though not all cases. However, clinical comparisons are difficult because most of the reported cases are mosaics.

An area in which fluorescence anal-



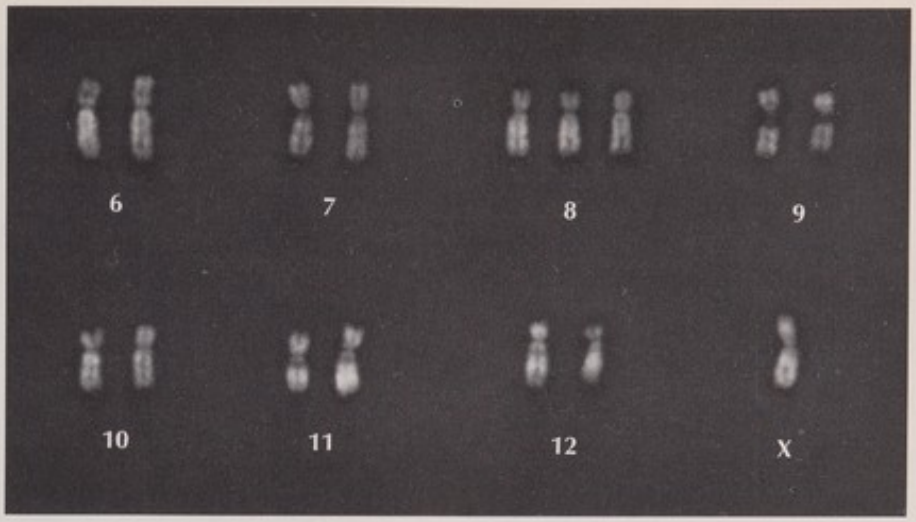
A third chromosome of the G group has long been known to be present in patients with Down's syndrome; whether it was a third 21 or 22 was uncertain. Fluorescent studies have confirmed that the extra chromosome actually belongs to the smaller of the two G-group pairs, making the trisomy conventionally "22." But since trisomy 21 has been so firmly linked to Down's syndrome, the numbering of the G chromosomes will be changed. As can be seen above, this child with Down's syndrome probably obtained his extra chromosome from his mother, whose 21 is very similar (Hereditas *70:153, 1972).*

ysis is likely to be of special value is that of translocations. Translocations can occur between either homologous chromosomes (members of the same pair) or nonhomologous ones. In both situations, a carrier of the translocation will retain a "balanced" complement of genetic material (though somewhat rearranged in structure) and will be phenotypically normal. The offspring, however, is another story.

If the carrier parent has a balanced homologous translocation such as 21-21, the offspring will either inherit this 21-21 chromosome and the normal 21 of the other parent, or he will inherit *no* chromosome 21 from the carrier parent and the normal 21 of the other parent. In the first case, the child will be trisomic for chromosome 21 and suffer from Down's syndrome. In the latter case, the embryo or fetus will be monosomic and be aborted. A monosomic condition is generally considered too unbalanced for viability.

In the case of a nonhomologous translocation such as 21-22, however, the offspring may either be normal (having inherited the normal chromosome 21 of the carrier parent) or a balanced carrier like the parent, or he may possess three chromosomes 21 (Down's syndrome). The importance of this information in genetic counseling is self-evident.

An example is the case of two mentally retarded brothers with a constellation of abnormalities resembling acrocephalosyndactyly. In both, conventional karyotyping showed that one of the G chromosomes had a larger-than-normal long arm; the same abnormality apparently showed up in the father who, however, was physically normal. Fluorescence analysis demonstrated that the father in fact possessed a balanced translocation in which a portion of the long arm of chromosome 1 had attached itself to number 22, while the two sons had inherited the elongated 22 but not the curtailed 1. They thus suffered from a partial trisomy 1. On closer examination of the original karyotypes, it became apparent that one of the father's 1 chromosomes was indeed somewhat shorter than those of the sons. At the next pregnancy in the family, prenatal karyotyping showed that the fetus possessed a balanced translocation like the father – and subsequently the family's first healthy child was born.



Fluorescence identification has established that a congenital condition previously known only as C-group trisomy is trisomy 8; one of five examples thus far studied is shown above. All three chromosomes 8 appear grossly normal. Clinical signs usually include skeletal abnormalities and moderate retardation (J Med Genet *9:1, 1972*).

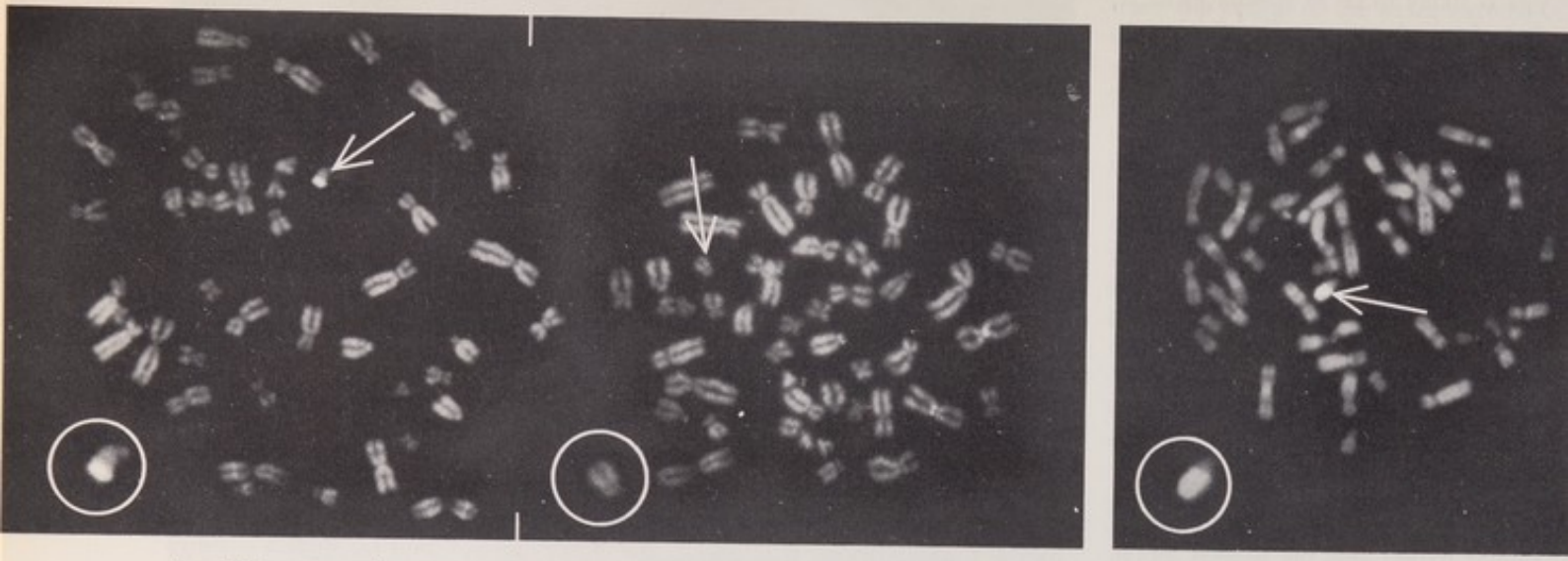
It is worth noting that the fluorescence analysis in this case showed not merely the general nature of the translocation but enabled one to say, within quite narrow limits, where the orginal breakage in the chromosome 1 had occurred. It seems safe to say that further work will enable us to identify even more subtle chromosomal rearrangements, including the so-called reciprocal translocations in which two chromosomes exchange material – often with little or no alteration of their gross morphology. Already one research group has been able to define in detail a complex rearrangement involving no fewer than three different chromosomes. With continued progress along these lines, we can expect more meaningful correlations between chromosome anomalies and the associated clinical picture, an area in which much confusion exists at present. There is also reason to believe that we can begin locating genetic "markers" on particular chromosomes.

The Philadelphia chromosome, found in most cases of chronic myeloid leukemia, occurs only in blood and bone marrow of patients with the disease and therefore cannot be considered a genetic marker. Most authorities had favored the hypothesis that it was a deleted (shortened) chromosome 21; in part this was because many patients with Down's syndrome suffer from leukemia. Furthermore, patients with chronic myeloid leukemia show low levels of leukocyte alkaline phosphatase in contrast with elevated levels of the enzyme found in Down's patients. Fluorescence analysis has made clear that this explanation for the Philadelphia chromosome is not correct. The simplest explanation of our results is that it is a deletion of part of the non-Down's G chromosome, 22 if 21 is to be considered the one involved in Down's syndrome.

It seems likely that other malignant disorders will turn out to have specific marker chromosomes that cannot now be identified with conventional techniques. Recently, in fact, it has been reported that 10 out of 12 patients with Burkitt's lymphoma possessed an additional terminal fluorescent band on the long arm of one chromosome 14.

Thus far we have spoken of anomalies involving only the autosomes. This, however, excludes the majority of chromosome abnormalities, which affect the sex chromosomes.

So far as inappropriately duplicated (or absent) X chromosomes are concerned, fluorescence analysis has added little to existing knowledge, since these could already be easily "counted" by observing the so-called Barr bodies in stained cells, and identified in karyotypes by means of autoradiography. Some interesting studies, however, have been done on XX males – individuals with anatomically

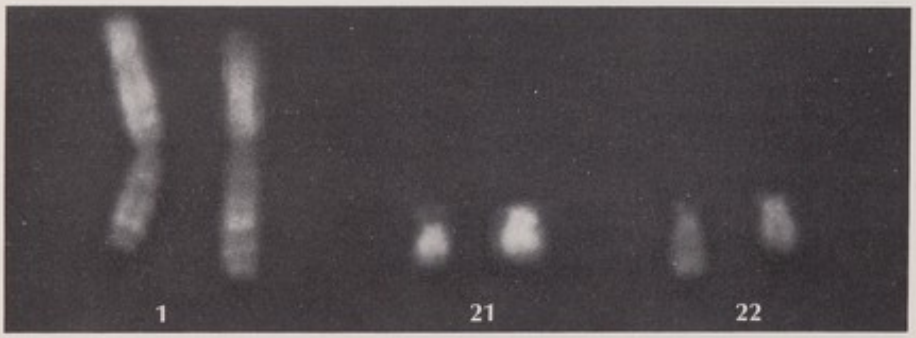


In addition to conditions referable to inappropriate numbers of sex chromosomes, clinical manifestations can be related to mosaicism, such as XO/XY; differences in size and fluorescence patterns of the Y between fathers and their sons are providing new information on these problems. In the paired micrographs above, the karyotype of a father is at left, that of his XO/XY mosaic son

normal male genitalia and testicular dysgenesis somewhat like that in Klinefelter's syndrome, but apparently without the Y chromosome present in Klinefelter's (XXY). These individuals, plus a number of XX hermaphrodites, are exceptions to the generally accepted principle that male genitalia cannot develop in the absence of the Y chromosome. Explanations put forward include undetected mosaicism of several sorts and translocation of part of the Y chromosome to another locus. So far as the latter is concerned, fluorescence analysis has thrown some light on the problem.

Since the long arm of the Y is brightly fluorescent, it should be easily detectable if translocated. In fact, of 13 XX males and one XX hermaphrodite whose karyotypes have been analyzed, none show any evidence of such a translocation. It should be noted, however, that in the Y chromosome it is the short arm that is generally believed to contain the "sex factor(s)" essential for gonadal differentiation – and this arm is also only weakly fluorescent, so it might be overlooked if translocated.

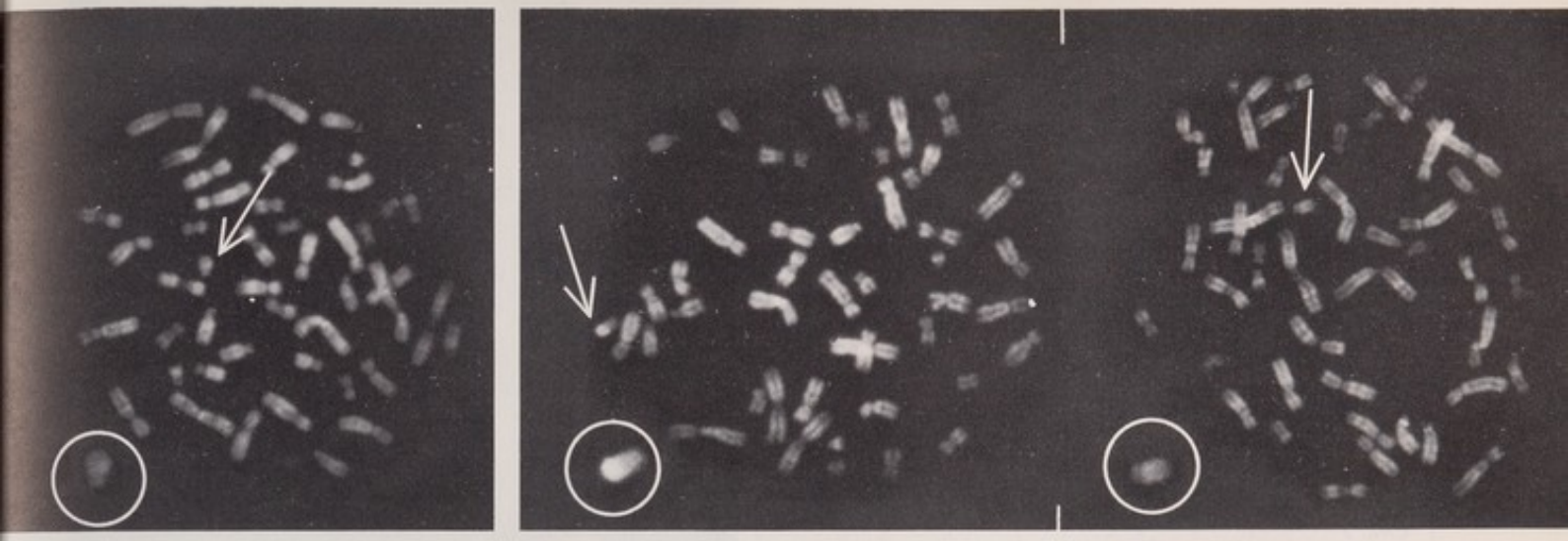
Concerning the Y chromosome, we have learned a good deal more. It is well known, for example, that this body can vary greatly in size in normal individuals – from less than 1.5 to nearly 4 μm. Fluorescence analysis indicates that the variation is probably restricted to the brightly fluorescent portion, the distal section of the long arm, while the remainder of the body is constant in size. This provides further evidence that the sex-determining factor(s) resides in the constant short arm or the nonfluorescent part of the long arm, since variations in the size of the fluorescent part of the long arm have no apparent effect on fertility or virility.



Translocation of genetic material can occur either between members of the same chromosome pair (homologous) or different pairs. While in both situations a balanced carrier will be phenotypically normal, the outcome in the offspring is different. In the homologous situation, the offspring will be either trisomic or monosomic (i.e., unviable). In the nonhomologous situation, the offspring may be normal, balanced, or possess extra chromatin, equivalent to partial trisomy. The fluorescence micrograph above demonstrates a balanced translocation between 1 and 22 in an apparently healthy carrier father; two sons, however, inherited the father's elongated 22 and his normal 1, producing a partial trisomy 1 that was associated with mental retardation and an acrocephalosyndactyly-like syndrome. Prenatal karyotyping in the next pregnancy showed that the fetus had a balanced translocation like the father's; a healthy child was born.

Fluorescence analysis has also helped to remove a misconception concerning the XYY anomaly. Since none of these individuals with this anomaly had been observed to produce XYY sons, some geneticists had invoked a special mechanism that eliminated the second Y from all spermatocytes. Fluorescent staining, however, shows that in fact a sizable fraction of the spermatocytes contain two Y's, so that in theory, at least, the individual should be able to beget XYY sons, and in fact cases have been reported of XYY fathers who had an XYY son.

The current feasibility of identifying small chromosomal aberrations more accurately has facilitated studies of the effects of ionizing radiation and other chromosome-breaking mutagens. Fluorescence analysis gives explicit information about the chromosomes that are involved in such radi-



*at right. In each case the Y is pointed out by an arrow and enlarged in the inset. Note the greater size and more intense fluorescence of all the paternal Y's. Karyotyping showed that the patient at far right had a translocation involving the short arm of chromosome 2 (see also diagram below). Phenotypically he was a normal male (*Hereditas *68:317, 1971).*

ation-induced aberrations as breaks, reciprocal translocations, rings, dicentrics, and fragments.

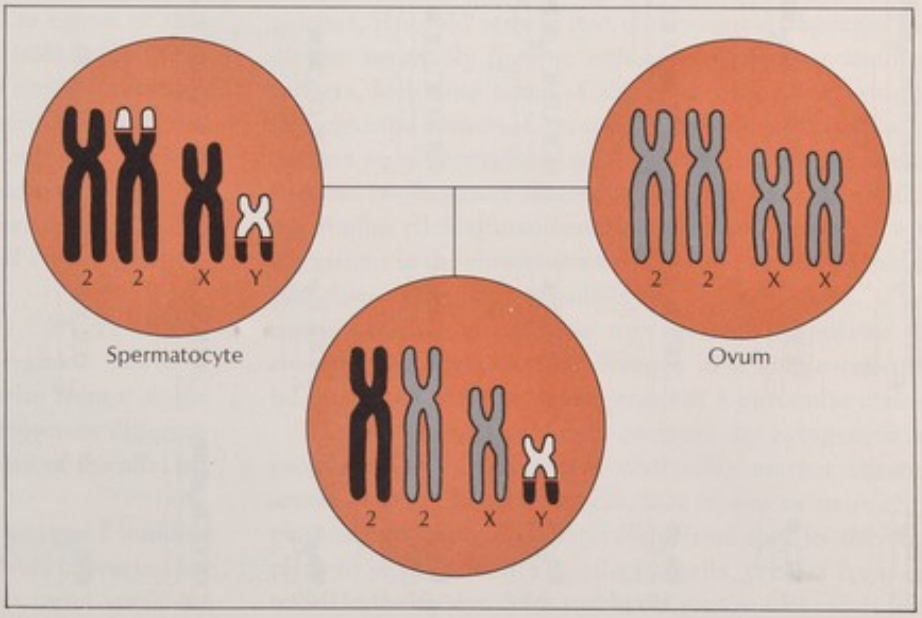
In such work the number of metaphases to be analyzed is always much greater than that required for clinical purposes. It is possible to reduce greatly the time required for pinpointing karyotype aberrations, by using a TV-based procedure for rapid identification of chromosomes and chromosomal regions. This instrument permits a twenty-fold increase in the speed of metaphase analysis. With the capacity to analyze so many metaphases comes an opportunity for statistical studies on the frequencies of different types of chromosomal breaks.

By this means, we have examined approximately 500 metaphases, which were obtained from X-irradiated human lymphocytes derived from 10 individuals. Two irradiation doses were used, 226 rad and 56 rad. The results are presented in the chromosome map shown on page 38.

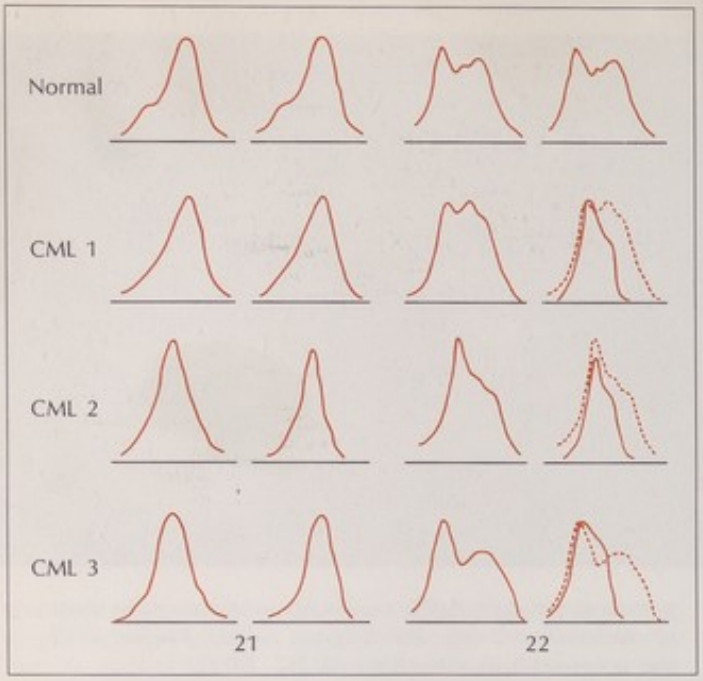
Unexpectedly we observed a difference in radiation-sensitivity between different chromosomes and chromosomal regions. The locations of about 380 individual breaks are plotted on the schematic chromosome map cited above. The regions are marked in accordance with the protocols of the Paris Conference (1971) on Standardization in Human Cytogenetics. The figure also shows the sensitivities of different parts of each individual chromosome. Any conclusions, however, should be drawn with great caution since the size of the investigated material is still relatively small.

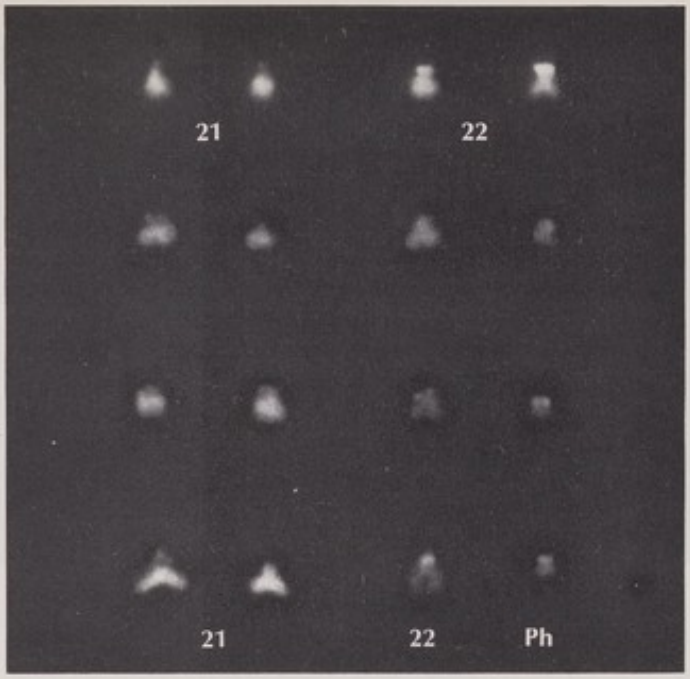
In the foregoing discussion, we have deliberately limited ourselves to the more interesting and significant clinical implications of fluorescence analysis. A number of other findings have been reported elsewhere, some of practical significance, others of primarily theoretical interest, since they tend mainly to confirm conclusions already reached through older cytogenetic techniques. Before closing, however, we would like to mention briefly some applications of fluorescence analysis in the field of research.

The technique has proved itself useful in identifying the chromosomes of other mammals – notably, for example, those of the mouse and pig – most of which are so similar that they cannot be identified by conventional techniques. Another interesting finding concerns the tobacco mouse, which has 14 (7 pairs) metacentric and 12 (5 pairs and X and Y) acrocentric chromosomes. The distinct fluorescence patterns in these chromosomes have con-



In patient whose karyotype is at top right, a reciprocal translocation between the short arm of chromosome 2 and the long arm of the Y was considered to have occurred de novo in a spermatocyte, with the son inheriting the abnormal Y and the father's normal 2.

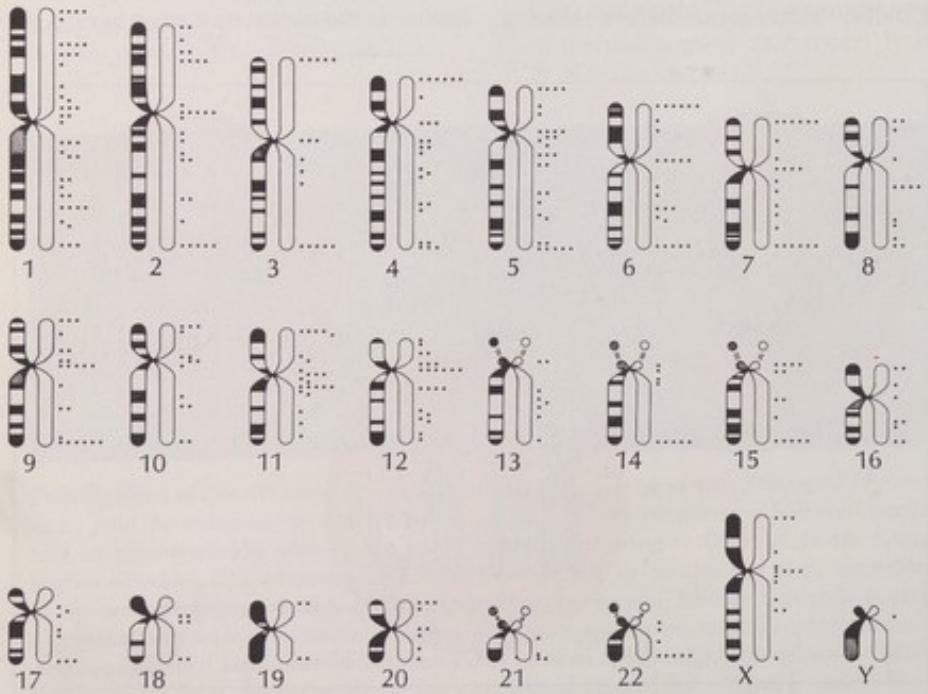




The Philadelphia chromosome found in patients with chronic myelocytic leukemia had been thought to be a shortened 21 (because of the association of Down's syndrome with leukemia). Fluorescence analysis has demonstrated that this explanation is not correct; rather, it is a deletion of the non-Down's chromosome in group G (22 if the one involved in Down's syndrome is classified 21). The photomicrographs show the Philadelphia chromosome at the right of pair 22 in three patients with the disease. Fluorometric curve of normal 22 is superimposed (dotted line) on those of the abnormal 22's (Exp Cell Res 63:239, 1970).

firmed the belief that each metacentric chromosome is homologous to two acrocentric chromosomes of *Mus musculus* that have fused in their centromere. How and why this fusion should have taken place in the course of evolution is a fascinating but still unanswered question.

Fluorescent identification of human and infrahuman chromosomes should prove of special value in the genetic analysis of somatic cell hybrids. It has been known for some years that cells from different species (e.g., mouse and man) can be fused so that the nucleus contains chromatin from both animals. In the mouse-man hybrid cell, most of the human chromosomes are eliminated quite rapidly, though in no particular order. If the cell then continues to show certain characteristically human biochemical activities (enzymes, antigens, drug resistance), i.e., properties not found in normal mouse cells, it is evident that the genes controlling these activities must be located on whatever human chromosome(s) remains in the cell. It has already been demonstrated, for example, that the human thymidine kinase gene is located on one of the E-group chromosomes. It should be obvious that the more precise chromosome identifications made possible by fluorescence staining and related techniques can greatly facilitate the job of relating specific biochemical properties to specific chromosomes. This is the essential prerequisite for any extensive mapping of human genes.



Karyotype diagram above shows chromosome breakage points after x-irradiation. Approximately 380 individual breaks have been plotted (Exp Cell Res 75:541, 1972).

Since fluorescence staining is only about three years old, a great many of its possible applications and implications are still to be explored. We believe, however, that what has already been accomplished with its aid demonstrates its remarkable utility as a tool for unlocking the secrets of heredity.

4

Oncogenic Implications of Chromosomal Instability

JAMES GERMAN
The New York Blood Center

Cytogenetic analysis very often reveals abnormalities in the affected cells in human cancers. Abnormal mitotic figures, abnormal numbers of chromosomes per cell, and abnormalities of chromosomal structure are to be found in most instances of human leukemia as well as in solid tumors. Most observers view these changes as a single phenomenon and but one expression of the uncontrolled cell growth that characterizes cancer, without assigning them an etiologic role. Others feel that a shifting chromosomal complement could be of importance in the evolution of the tumor. Nonetheless, the possibility has not been ruled out that chromosomal change may be a primary event in some human malignant disease. Mutation involving one or more genes could prove to be the essential cellular change that converts a normal cell to cancer or that permits emergence or progression of at least some forms of cancer. One possibility is that once a specific mutation has occurred, affected cells might be more susceptible to conversion to cancer by an unrelated oncogenic factor. Then again, perhaps the role of viruses in relation to chromosomal change is such that the latter permits a viral genome to combine linearly with a cell's own genome, initiating conversion to cancer.

In considering the problem stated by the title of this paper, three observations deserve consideration:

1) Many human cancers appear to arise from a single cell; often this cell has one or more chromosomes differing in morphology from any in the complement of the affected individual's noncancerous cells.

2) Chromosomal instability, i.e., an increased number of disrupted and rearranged chromosomes, characterizes the cells of certain persons who are, because of some environmental situation, at an increased risk of cancer.

3) Several rare genetic disorders, to be discussed here in some detail, which are characterized both by chromosomal instability and an increased risk of cancer, share with well established "causes of cancer" – certain forms of radiation, certain chemicals, and certain viruses – the ability to produce increased numbers of cells having *de novo* chromosomal rearrangements.

Before documenting these three observations, however, a certain technical limitation should be mentioned. Only limited information is obtainable from the standard methods of preparing cells for cytogenetic study. Subtle rearrangements often go undetected, a situation that may be changed when technical advances, such as quinacrine fluorescent or Giemsa banding techniques (see Caspersson and Zech, Chapter Three), come into wider use in the study of cancer. Even taking this technical limitation into account, it would appear that chromosomal abnormalities are not invariably present, either in human or nonhuman cancers, including some of the most extensively studied experimental tumors. Conceivably, it may be that in some cancers no abnormalities will be detectable by any means because two separate classes of cancer exist, both exhibiting similar clinical manifestations and pathology but with alterations in the chromosome complement involved in only one class. Once this possibility is acknowledged, it becomes possible to examine with greater objectivity the concept that chromosomal changes in a single cell may be associated with the development of a particular cancer.

In part this concept rests on work by cytogeneticists establishing the presence of distinctive marker chromosomes in many human cancers, that is, one or more chromosomes differing in morphology from any in the complement of the patient's unaffected cells, present from cell to cell in the cancer. Such marker chromosomes result from structural rearrangement, either intrachromosomal or interchromosomal. These are most often translocations, deletions, or inversions (see Hirschhorn, Chapter One). The

new chromosome configuration is not only stable enough to permit it to pass successfully through mitosis but the cell's genome, although altered, permits cellular proliferation. In this case a cytogenetically distinct subpopulation of cells develops and may be identifiable amidst the host's normal cells.

It is well known that additional abnormalities, particularly abnormal mitotic figures and cells with extra or missing chromosomes and sometimes additional chromosomal rearrangements, are often also evident, but the marker chromosome tends to be the constant finding from cell to cell of the mutant subpopulation, detectable both at the primary site of the cancer and in its metastases. When examined by standard cytogenetic techniques the marker usually appears unique to a given patient's cancer. Thus, although two patients may have similar cancers in terms of pathology and clinical expression, as a rule there is little apparent similarity in the marker chromosomes seen in both cases.

The important exception is chronic myelogenous leukemia, inasmuch as the same marker, the Ph[1] chromosome, appears in the neoplastic cells of nearly all (close to 90%) patients. (In contrast, note that although this distinctive chromosomal abnormality has been associated with chronic myelogenous leukemia, usually no marker chromosomes, or for that matter any other chromosomal aberrations, have been associated with acute lymphocytic leukemia.) As it happens, the abnormality in chronic myelogenous leukemia, affecting the long arm of one of the smallest of the chromosomes (number 22), is especially well located for microscopic recognition (see Caspersson and Zech, Chapter Three). Conceivably, specific markers appearing in other forms of leukemia or solid tumors are being overlooked because of the technical limitations referred to earlier.

Since chromosomal rearrangements are essentially never seen in the complements of unirradiated cells of nonmalignant tissues, the presence of the same marker chromosome from cell to cell throughout the cancer cell population strongly suggests that this population is derived *from a single cell* that also had in its complement the same marker (or markers).

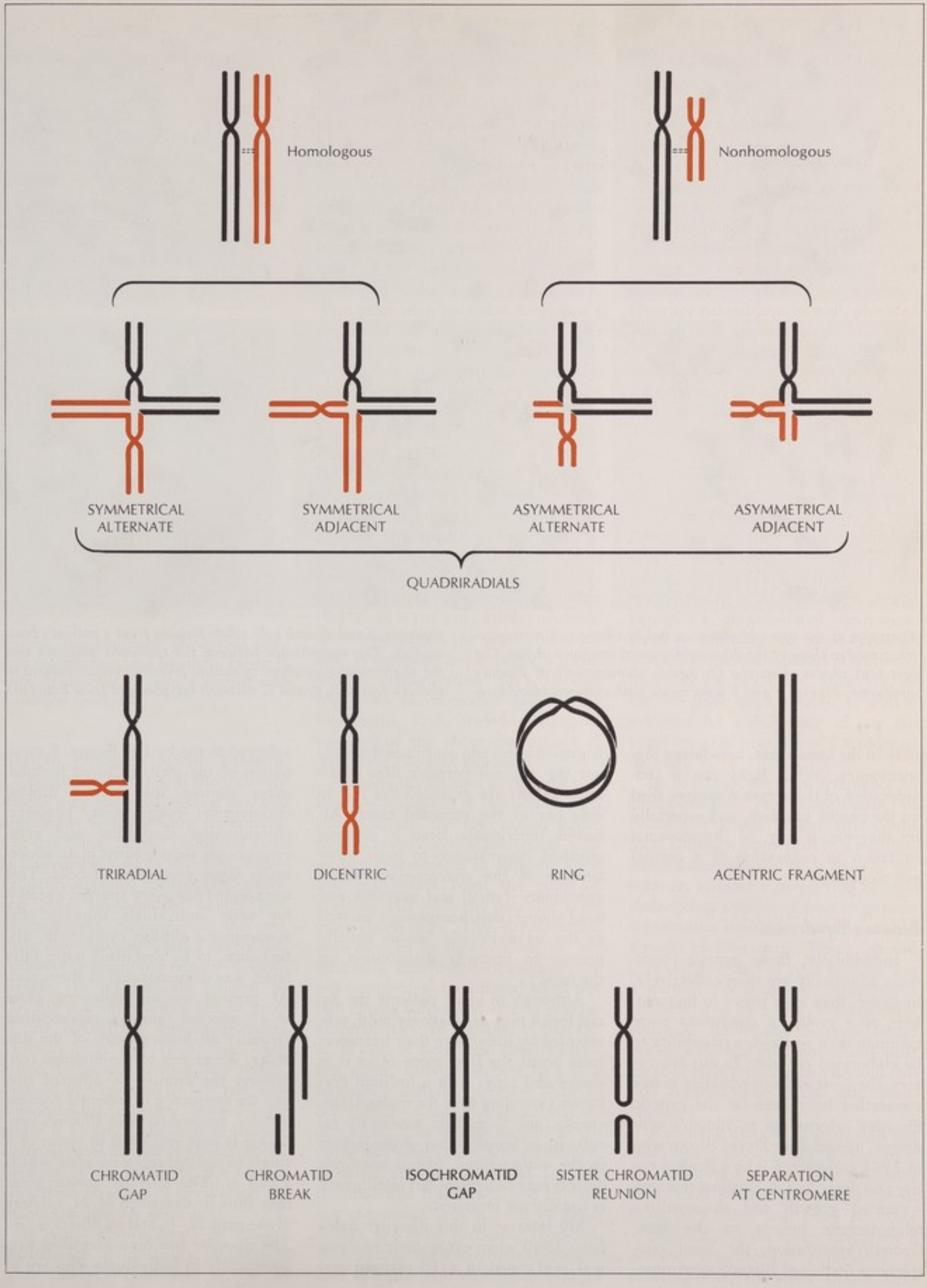
As soon as the association of distinctive marker chromosomes and cancer was recognized, several investigators concluded that many cancers evidently originate in a single cell in which some chromosomal changes had occurred. For example, C. E. Ford, studying the cytogenetics of primary reticular neoplasms in the mouse, observed that many of them represented clones of cells containing one or more marker chromosomes. By analysis of these markers, he could trace development of neoplastic cell populations by proliferation from one ancestral cell. Atkin, performing cytogenetic studies in women with different types of cancer (including adenocarcinoma of the ovary and carcinoma of the cervix and the body of the uterus), found that for any one case there usually was a stem line with the same abnormal chromosomal complement from cell to cell.

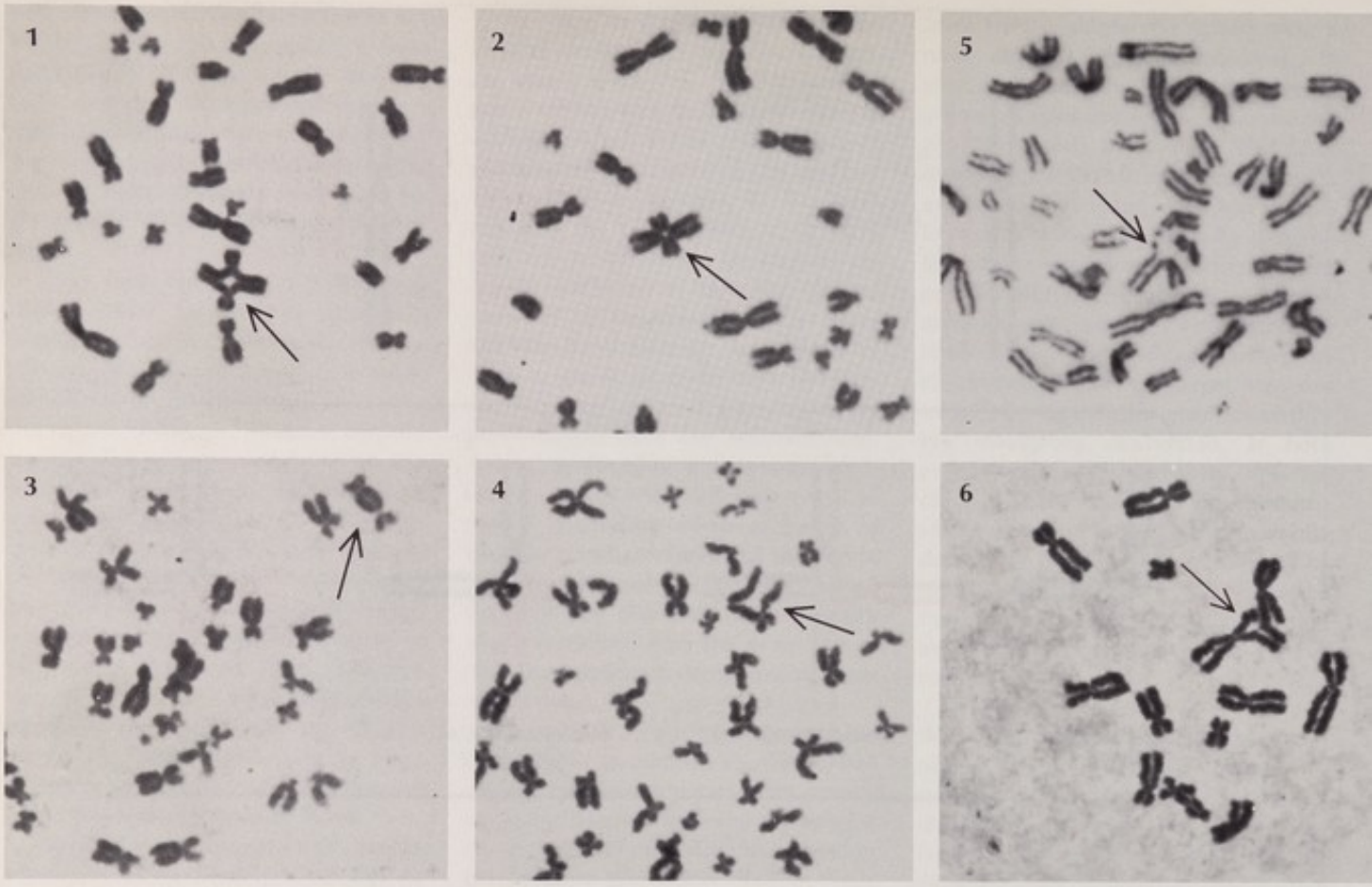
Other data of a noncytogenetic nature appear to confirm the single cell origin of at least some cancers. I refer to studies aimed at tracing the origins of neoplasms in individuals having two or more genetically distinct types of cells. As is known, women normally have two cell types by virtue of fixed genetic inactivation (occurring early in embryogenesis) of one of the two X chromosomes in each somatic cell. A similar inactivation has been demonstrated for several X-linked gene loci, including the one for the enzyme glucose-6-phosphate dehydrogenase. In a woman heterozygous at this locus with one normal and one variant allele, both the normal G-6-PD allele and the variant will be expressed in a mixture of her cells in tissue culture. However, a single cell or clone of cells will show only one of the two enzyme types. Linder and Gartler found this to be the case with uterine fibroid tumors: any one tumor exhibits only one type of G-6-PD in a heterozygous woman, but her different tumors will have different types. When this thesis was tested by Fialkow and coworkers in women with chronic myelogenous leukemia, a sampling of blood cells showed only one enzyme type, suggesting that all the blood cells were descendants of one cell. The same proved true for Burkitt's lymphoma and most other types of tumors tested.

But if cancer frequently appears to arise from a single cell manifesting a new chromosomal complement, is such a chromosomal rearrangement a primary factor in the conversion of the progenitor cell to cancer? The findings suggest that it may be; however, available knowledge does not yet allow us to state unequivocally that the first cell to qualify as cancer is the first cell in which the chromosomal rearrangement occurs. Furthermore, if it should be, it is unclear whether the chromosomal rearrangement precedes or follows the cellular event constituting malignant conversion. If both events occur simultaneously, it is unclear whether they are one and the same, as some suggest, or mutually dependent but separate, as favored by others. For example, the mutation might alter the cell in such a way that another factor could take over genetic control.

While the sequence of events and their significance remain to be determined, the point essential to our discussion is that *a chromosomal mutation is often present in the single cell* from which all cells that can be sampled in an established cancer seem to have been derived. Instead of speculating further on the possible relationships involved, let me direct your attention to the rare genetic disorders mentioned above, in which cancer is seen to develop with unusual frequency – disorders also characterized by increased chromosomal instability and a tendency to the formation of new chromosomal arrangements. They are Bloom's syndrome, Fanconi's anemia, the Louis-Bar syndrome, and, prob-

The possibility that chromosomal mutation may have a major role in emergence of some human cancers is supported by the presence of chromosomal disruptions or rearrangements, such as those depicted schematically at right, in the complements of many cells from patients with genetic disorders in which cancer develops with unusually high frequency. For example, in Bloom's syndrome a characteristic cytogenetic lesion seen in blood cells dividing in vitro is a symmetrical quadriradial (Qr) figure apparently produced as result of chromatid interchange following breakage at corresponding sites on homologous chromosomes (upper left). Qr's involving nonhomologous chromosomes (upper right) have been associated with Fanconi's anemia. Other types of chromosomal anomaly or rearrangement are shown in lower panel.


Homologous
Nonhomologous
SYMMETRICAL ALTERNATE
SYMMETRICAL ADJACENT
ASYMMETRICAL ALTERNATE
ASYMMETRICAL ADJACENT
QUADRIRADIALS
TRIRADIAL
DICENTRIC
RING
ACENTRIC FRAGMENT
CHROMATID GAP
CHROMATID BREAK
ISOCHROMATID GAP
SISTER CHROMATID REUNION
SEPARATION AT CENTROMERE



Examples of the types of chromosomal breakage and rearrangement seen in three of the diseases discussed are given above. The first four photos show the Qr figures characteristic of Bloom's syndrome. Figures 1 and 2 were made with cultured blood lymphocytes, 3 and 4, with cells taken directly from a patient's bone marrow. The resemblance between the two pairs indicates that the abnormalities observed in culture preparations are present in vivo. In figures 5, 6, and 7, cultured lymphocytes from Fanconi's

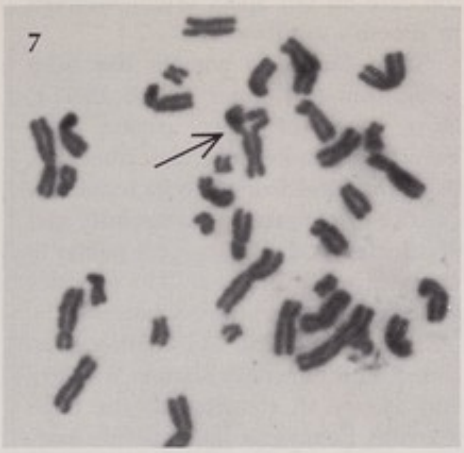
ably in the same class, xeroderma pigmentosum. What light can a consideration of these rare disorders shed on the cancer problem, and especially on the role, if any, of chromosomal mutation in conversion of a normal cell to cancer?

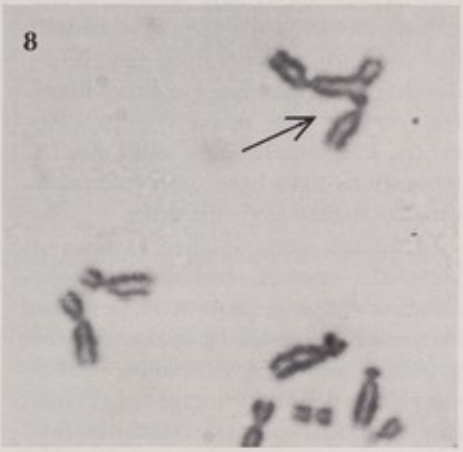
Bloom's Syndrome

Individually, these genetic conditions admittedly are rare; collectively, however, they may prove to be members of a group of conditions more common than realized, a possibility to be elaborated on later. In our laboratory, the most extensive studies in this connection have been on the genetic disorder referred to as Bloom's syndrome, named after David Bloom who in 1954 recognized it as a clinical entity and described its cardinal features – stunted growth and sun-sensitive telangiectatic lesions on the face. Growth retardation, the most conspicuous feature of Bloom's syndrome, begins in utero and continues throughout the affected person's life. Maximum height has exceeded five feet in only one of the recorded cases. Affected individuals bear a striking physical resemblance to each other because of the characteristic facial appearance (small and narrow) and head shape (dolichocephalic) as well as the telangiectatic lesions usually present in butterfly distribution on the face.

Although in some patients the facial lesion may be relatively mild, represented by little more than telangiectasia about the lips, more often it is severe and ugly, with a brilliant erythema extending over the malar areas, cheeks, and nose. Sun sensitivity results in an accentuation of the lesions in the summertime. Minor anatomic defects may be present but intelligence is usually not impaired.

My interest in this disorder dates from 1960 when cytogenetic analysis was performed in a 10-year-old girl referred to me by Dr. Bloom. Examination of the girl's cultured lymphocytes showed the normal diploid chromosome complement; however, chromosomal disruption and rearrangements were evident in an abnormally large proportion of cells. This unexpected discovery spurred a search for other individuals showing this syndrome; a similar cytogenetic disturbance, to be described more fully later, was demonstrated in them also. At present we are following close to 40 affected families, representing virtually all known cases of the disorder. According to the findings concerning the number of affected sibs and the frequency of parental consanguinity, transmission of Bloom's syndrome is best explained in terms of a single autosomal recessive gene, which I call *bl*. The affected homozygotes may thus be described *bl*/*bl*, carriers of the gene *bl*/+, and all others +/+. The disorder has been identified predominantly in individuals who are of





anemia show chromatid disruptions, complex interchanges, and triradials. Figure 8, with its triradial, is from a patient with Louis-Bar syndrome.

Eastern European Jewish ancestry.

As clinical features of Bloom's syndrome were analyzed, several new and significant features were recognized. For one thing, during childhood there is a predisposition to infection, which probably explains the condition's having gone unrecognized in the preantibiotic era. Subsequently this predisposition subsides, but a severe disturbance in immunity remains demonstrable. In all patients studied so far, the concentration of one or more of the circulating immunoglobulins (usually IgA and IgM) has been found decreased; in those few tested, delayed hypersensitivity has not been demonstrated. Further immunologic study is clearly indicated, to include heterozygotes as well as homozygotes, because in a number of the former also, concentrations of circulating immunoglobulins are depressed.

It is not known how – or if – the disturbance of immunity in patients with Bloom's syndrome relates to another observation of paramount importance, the affected person's predisposition to develop cancer. Two of the three patients in Dr. Bloom's original report were to succumb to acute leukemia, one at age 13 and the other at 25. Among the first 45 cases about whom we have learned and who have survived early childhood, nine instances of cancer have occurred. One 39-year-old man developed two primary cancers, a squamous cell carcinoma of the upper esophagus and an adenocarcinoma of the sigmoid colon. Two other men have also developed a cancer of the alimentary tract after the age of 30.

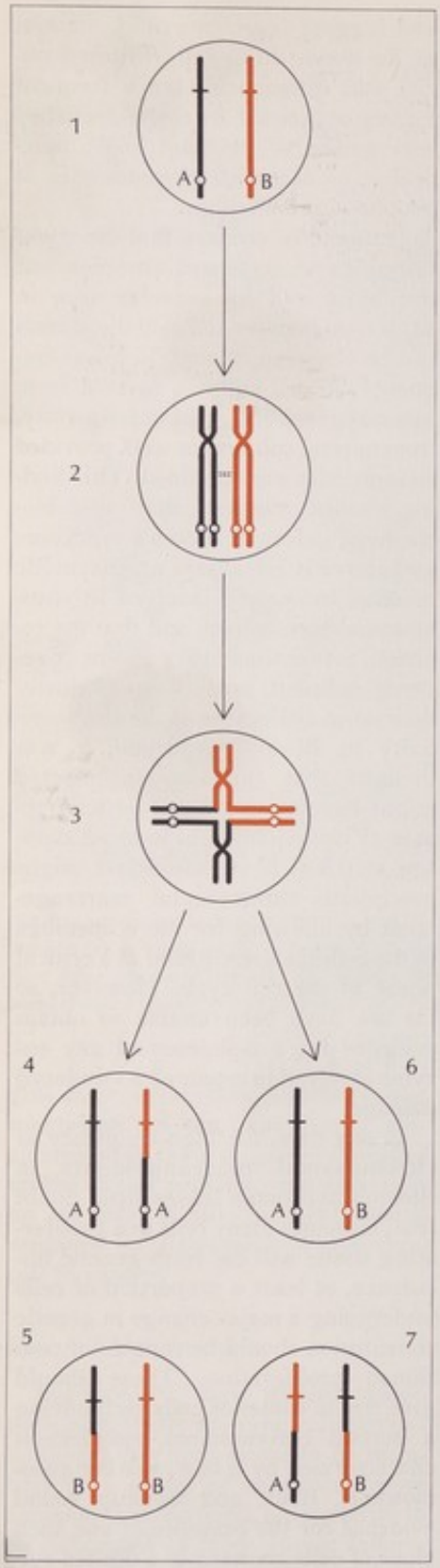
It is now clear that the occurrence of cancer at this frequency cannot be a coincidence, and it is tempting to suggest that the tendency of cells homozygous for the *bl* gene to undergo chromosomal abnormalities may be of importance in this cancerous predisposition. The pattern of these cytogenetic abnormalities has been repeatedly confirmed. In both blood lymphocytes and skin fibroblasts in tissue culture, an increased number (0.5%–11%) of cells display chromatid interchanges, the most characteristic cytogenetic lesion being a quadriradial configuration (Qr), derived from two chromosomes and having four arms, each arm consisting of portions of the sister chromatids of one of the two chromosomes.

Usually the Qr seen in Bloom's syndrome is composed of two homologous chromosomes; the four-armed figure formed is strikingly symmetrical, with its two centromeres in opposite arms and equidistant from the point of crossing and with the opposite arms equal in length. The chromosomal sites affected by these lesions are not randomly distributed throughout the cell complement but instead affect certain chromosomes and certain regions preferentially. In addition to this Qr, *bl/bl* cells at metaphase usually manifest other, nonsymmetrical Qr's at a frequency definitely greater than in cell cultures derived from normal individuals. This applies as well to triradial and more complex interchanges, dicentric chromosomes and acentric chromosomal fragments, and to various chromatid lesions such as gaps, breaks, and sister chromatid reunions. At anaphase there are increased numbers of cells with chromatin bridges and lagging fragments of chromatin; in the nondividing cell, distorted nuclei and micronuclei are a frequent finding and result from the disturbed segregation of aberrant (i.e., polycentric or acentric) chromosomes at anaphase and telophase.

It should be stressed that the manifestations of increased chromosomal instability and its sequelae seen in *bl/bl* lymphocytes and skin fibroblasts can be observed, though in lower frequency, in cell cultures derived from heterozygotes and, very infrequently, from normal subjects as well, provided enough cells are examined. This finding would suggest that the biochemical defect in Bloom's syndrome (whatever it is) affects an enzymatic reaction intimately involved in chromosomal metabolism, and that the reaction, rather than being absent or severely deficient, proceeds more slowly than normal. Because of the sun sensitivity in Bloom's syndrome it was thought that the enzyme affected might have something to do with repair of ultraviolet-light-induced damage to DNA; if so, the defect might precipitate chromosomal rearrangement by allowing for more openings in the cellular DNA strand at a critical phase in the cell cycle. However, so far we have been unable to obtain evidence for a deficiency of any enzyme involved in repair of UV-induced damage.

In any case, the high frequency of chromosomal rearrangements in Bloom's syndrome leaves little doubt that, although many cells in a proliferating tissue will die from genetic imbalance, at least a proportion of cells undergoing a major change in genetic constitution should be capable of continued proliferation. These should give rise to clones of cells each having a mutant chromosomal complement characterized by a new, marker chromosome. Rauh and Soukup found evidence for the presence of one such clone of cells growing in a diploid culture derived from *bl/bl* skin fibroblasts.

Following division of a cell that had contained an equal and symmetrical Qr, theoretically the two daughter cells could possess genomes different from each other and from all other somatic cells in the body; that is, by a process functionally equivalent to crossing-over, the cells would have



Diagrams suggest sequence whereby chromatid interchange (2, dashed line) between homologous chromosomes (1) resulting in formation of symmetrical Qr figure (3) could yield progeny genetically different from parent cell strain and homozygous for all genes distal to the point of exchange (4, 5). Alternatively, the progeny might be like the parent cell (6) or merely show new position of alleles (7).

become homozygous for all genes of the affected chromosome distal to the point of exchange. Though retaining a balanced and complete human genome, each daughter cell, and its progeny, would lack certain alleles present in all the other cells. If this is actually the case in vivo in Bloom's syndrome, evidence for antigenic or enzymatic mosaicism should be demonstrable, if an adequate search were made.

The cytogenetic findings described in reference to Bloom's syndrome cells in tissue culture do not necessarily apply to cells proliferating in vivo, but if they do, this could account for some of its clinical features. For example, poor survival among daughter cells receiving the more abnormal chromosomal complements could relate to the severe growth retardation characteristic of Bloom's syndrome. Evidence that chromosomal instability does occur in vivo was obtained when we examined cells aspirated from the bone marrow of a *bl/bl* seven-year-old boy and found increased numbers of cells with dicentric chromosomes, triradial configurations, and other complex exchange figures. (The one other marrow we have been able to examine appeared normal.)

In the course of our cytogenetic studies of Bloom's syndrome, similar chromosomal aberrations in dividing lymphocytes have been identified in dividing lymphocytes from all those examined, i.e., from most of the known affected individuals, including three of those who went on to develop cancer. This probably means that the chromosomal instability has little to do with the cancer itself, but does it not provide a predisposing background?

Fanconi's Anemia

The situation is, in some ways, similar in Fanconi's anemia, another rare genetic disorder, which combines a functional defect in the bone marrow with major anomalies of solid structures including the skeleton, heart, and kidneys. Actually there are not many clinical similarities between Bloom's syndrome and Fanconi's anemia. The latter may be associated with growth retardation, but not to the same degree. Brownish skin pigmentation is prominent but more extensive and of different distribution than the café au lait spots often observed in Bloom's syndrome.

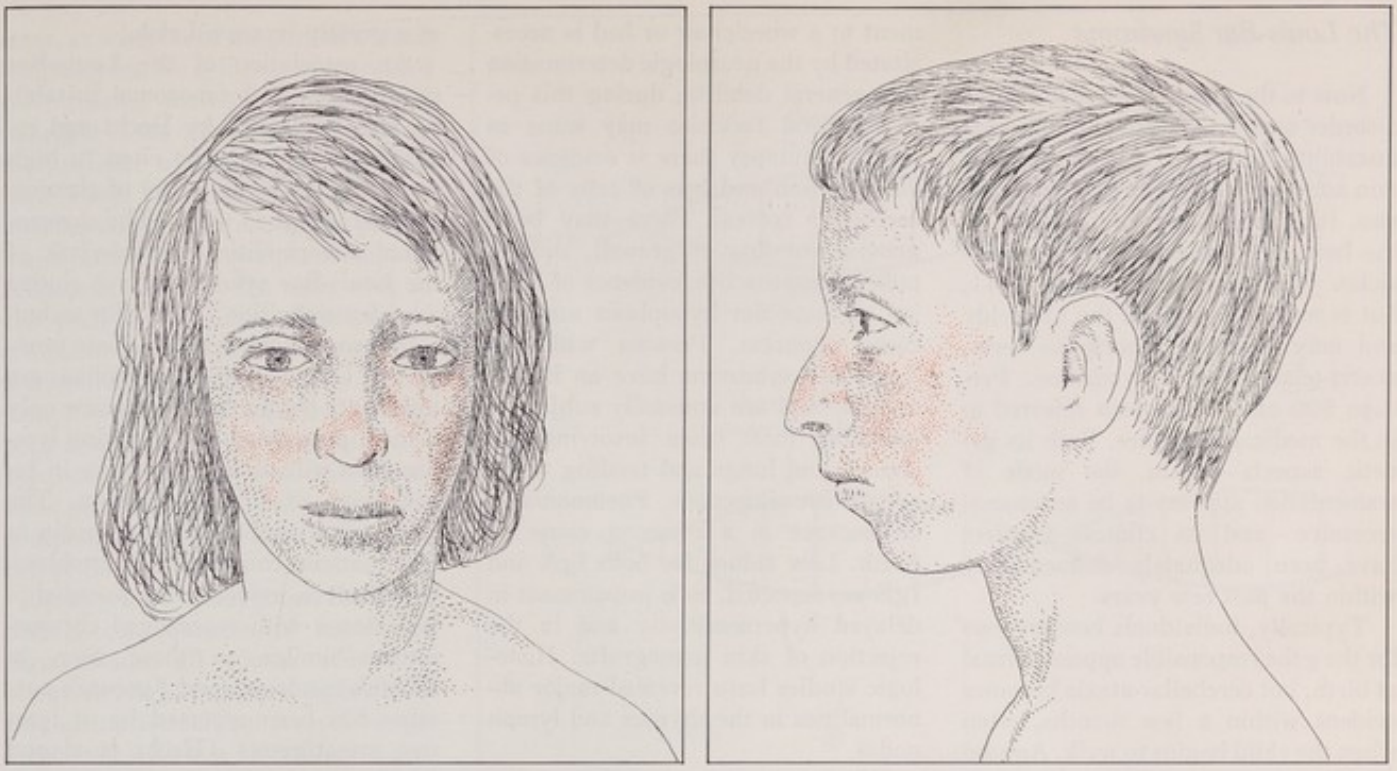
Since Fanconi's anemia has been recognized far longer (since 1927), more cases, probably around 200, have been studied and described. As in Bloom's syndrome, an increased incidence of parental consanguinity and of affected sibs of both sexes points to a recessive transmission. The genetics has not been well studied, however.

The hematologic dysfunction associated with Fanconi's anemia is severe and affects all elements of the bone marrow. Pancytopenia and bone marrow hypoplasia usually become symptomatic between the ages of 4 and 12 and tend to progress until the child's death from hemorrhage or other manifestations of bone marrow failure. Recently, hormone therapy and other interventions have been used to ameliorate the hematologic disorder.

Although some detailed reviews of Fanconi's anemia, including one by Fanconi himself, do not refer to an increased incidence of leukemia, the generally accepted view now is that the risk is definitely increased. While this view is based on relatively few cases, considering the infrequency both of leukemia and of the syndrome itself, the number of affected persons with leukemia appears sufficient to indicate an association. An important possibility that has long intrigued clinicians observing this disease is that cancer risk may also be increased in relatives of affected individuals, presumably heterozygous for a Fanconi's anemia gene, and supporting evidence was recently provided by Swift and coworkers.

Given the severity of the hematologic abnormality in Fanconi's anemia, any predisposition to leukemia might not be fully expressed because, at least in the past, few patients have survived for long after the appearance of panmyelopathy. If a neoplastic cell line is to emerge and replenish the marrow, obviously it would have to do so before the patient succumbs to hemorrhage or overwhelming infection. Will prolongation of life in Fanconi's anemia by hormonal therapy be associated with increasing numbers of leukemics?

Some cytogenetic comparisons bear mentioning in reference to Fanconi's anemia and Bloom's syndrome. The increased chromosomal instability in



Because of their characteristic facial appearance as well as the presence of telangiectatic lesions, unrelated individuals with Bloom's syndrome bear a close physical resemblance to one another. A key feature shared is their small stature, resulting from growth retardation beginning in utero and continuing throughout life. Maximum height almost never exceeds five feet.

Fanconi's anemia was first reported some eight years ago by Schroeder, on the basis of her studies of cultured blood lymphocytes. We have observed this in skin fibroblast cultures as well. On first inspection the chromosomal abnormalities seem similar to those in Bloom's syndrome; there certainly are many breaks and rearrangements. Recently, Schroeder and I jointly evaluated the cytogenetic disturbance in the two disorders and concluded that the patterns of chromosomal disruption and rearrangement are different. Whereas chromatid interchanges and rearrangements, chiefly the equal symmetrical Qr, predominate in Bloom's syndrome, in Fanconi's anemia a relatively higher proportion of cultured cells show chromatid gaps and breaks. When Qr's occur in Fanconi's anemia, nonhomologous chromosomal regions are usually affected; moreover, sites of lesions are fairly randomly distributed throughout the chromosomal complement. Genes responsible for the two disorders obviously have different effects clinically, and a difference in the pattern of chromosomal instability is not unexpected.

Using bone marrow aspiration studies to evaluate the degree of chromosomal instability occurring in vivo, Schroeder found an increase in mitotic and nuclear aberrations during metaphase in a few patients as well as anaphase bridges, fragmented chromosomes, and micronuclei. Using serial cytogenetic analysis of aspirated bone marrow, she also detected and followed the proliferation of a pseudodiploid clone, i.e., a cell population with 46 chromosomes but with some chromosomal rearrangement. In examining skin fibroblast cultures derived from a case of Fanconi's anemia, Young detected two clonal subpopulations, each with a mutant chromosomal complement, growing among cells with normal complements. In these patients there was no association with leukemia. All that can be said at present is that clones of cells containing a clearly aberrant chromosomal complement but exhibiting no malignant characteristics are associated with a disorder carrying an increased risk of cancer. Moreover, like the stem lines of established cancers, these clones evidently originate from a single cell containing a stable chromosomal rearrangement.

In attempting to elucidate the relationship between Fanconi's anemia and leukemia, Todaro used simian virus (s v) 40 to infect diploid fibroblastic cell lines either homozygous or heterozygous for a Fanconi's anemia gene. In such cell lines, as compared with control cultures, monolayers of cells planted on the surface of a culture dish showed a greater number of localized areas undergoing rapid and disarrayed proliferation. Whereas in control cultures 1.6 to 5.1 areas of "transformation" appeared for every 10,000 cells, the number increased to 20.1 to 28.2 per 10,000 among cells carrying one gene for Fanconi's anemia; in cells containing two Fanconi's anemia genes it increased to 41.4 to 79.7 per 10,000. Although these proliferative responses termed "transformation" may not always be assumed to represent malignant conversion, what is of interest is that while the changes occur most readily in cells from the affected homozygote, they are also seen in cells of normal individuals. Response in the heterozygote is intermediate between the two. Thus, the difference is not in the kind of interaction between the virus and the cell but in the frequency with which it occurs.

The Louis-Bar Syndrome

Now to the consideration of a third disorder associated with chromosomal instability and cancer and sharing certain additional features with the other two. It is conveniently referred to as the Louis-Bar syndrome, for the physician who wrote about it in 1941, but is more often given the unwieldy and only partially descriptive term, ataxia-telangiectasia syndrome. Perhaps 200 cases have been referred to in the medical literature. Both its genetic aspects – again, the mode of transmission appears to be autosomal recessive – and its clinical features have been adequately defined only within the past few years.

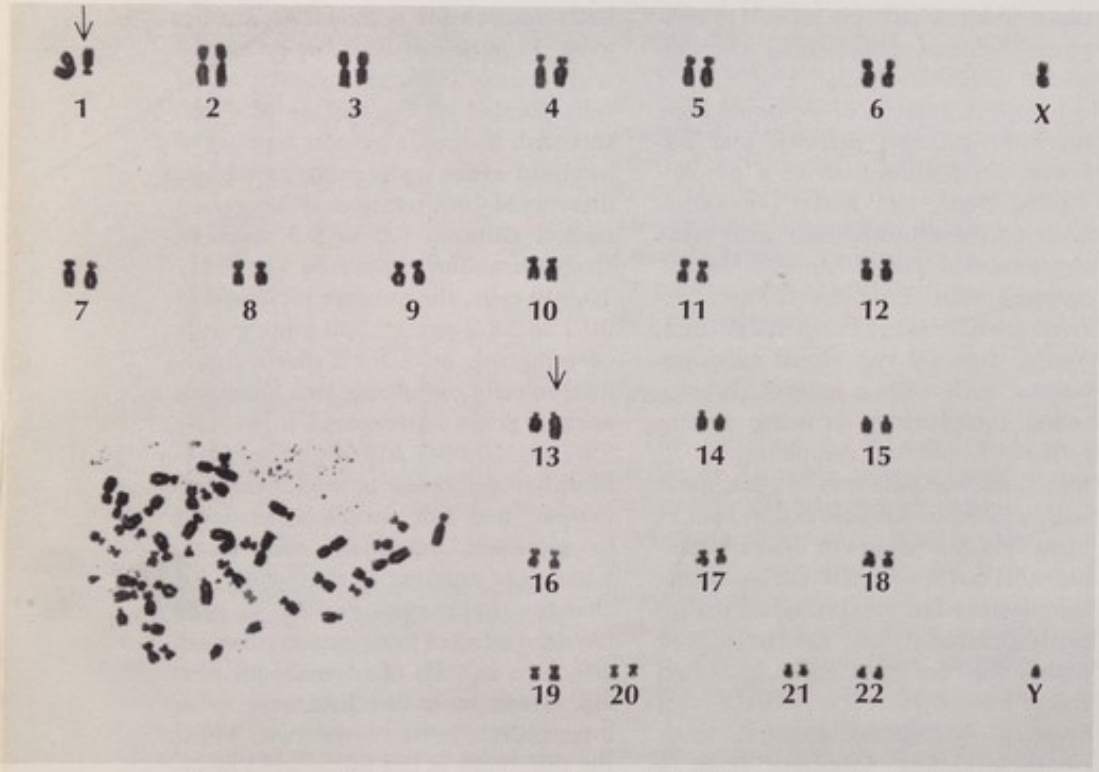
Typically, individuals homozygous for the gene responsible appear normal at birth, but cerebellar ataxia becomes evident within a few months, often when the child begins to walk. Around the same time, or later, telangiectases appear in the bulbar conjunctivas. Both the ataxia and telangiectasia tend to be progressive. The latter affects other regions including the face, neck, antecubital fossas, hands, and knees. Within a few years confinement to a wheelchair or bed is necessitated by the neurologic deterioration and general debility; during this period mental faculties may wane as well. At autopsy there is evidence of deterioration and loss of cells of the cerebellar cortex. There may be a general stunting of growth, and, as puberty approaches, evidence of ovarian or testicular hypoplasia may become apparent. Persons with the Louis-Bar syndrome have an immunopathy and are unusually subject to infection, most often involving the sinuses and lungs and tending to result in bronchiectasis. Pneumonia in adolescence is a common cause of death. Low values for both IgA and IgE are reported, as is impairment in delayed hypersensitivity and in the rejection of skin homografts. Histologic studies have revealed major abnormalities in the thymus and lymph nodes.

When cancer develops in association with the Louis-Bar syndrome it is likely to originate in lymphoid reticular tissue. An accurate estimate of its frequency is not possible on the basis of available data, but there remains no doubt that these patients are at a greatly increased risk.

An association of the Louis-Bar syndrome and chromosomal instability was first noted by Hecht and coworkers in 1966, who cited "a high frequency (20% to 30%) of chromosome breakage in vitro." The chromosomal abnormalities characteristic of the Louis-Bar syndrome have eluded complete definition, largely for technical reasons, however. Adequate numbers of lymphocytes in metaphase are difficult to obtain, in part because only a small proportion of circulating lymphocytes will divide in culture in response to phytohemagglutinin. The presence of absolute lymphopenia in these patients compounds the problem. Nonetheless, evidence for pseudodiploid clones with rearranged chromosomes – similar to those seen in Bloom's syndrome and Fanconi's anemia – has been reported by at least two investigators (Hecht et al and Pfeiffer).

Xeroderma Pigmentosum

In the three genetic disorders considered thus far, chromosomal instability leading to chromosomal mutations has been documented by direct cytogenetic observation in individuals homozygous for the responsible genes. In Bloom's syndrome and Fanconi's anemia similar abnormalities have been shown in heterozygotes as well. In a fourth genetic disorder, xeroderma pigmentosum, a tendency to chromosomal instability has not been shown directly. However, clones of cells with pseudodiploid complements have been observed growing among normal skin fibroblasts in tissue culture, suggesting that at some earlier time a tendency to chromosomal disruption and rearrangement may have existed, perhaps when the cells were exposed to ultraviolet light. The detection of such mutant clones was made in our laboratory using a fibroblast cell line derived from non-neoplastic skin of a young man homozygous for the abnormal gene. (Like the other genetic disorders we have been discussing, classical xeroderma pigmentosum is transmitted as an autosomal recessive trait, but recent findings indicate the existence of genetic heterogeneity.) Among some 200 cells examined, most appeared to have a normal diploid chromosomal comple-

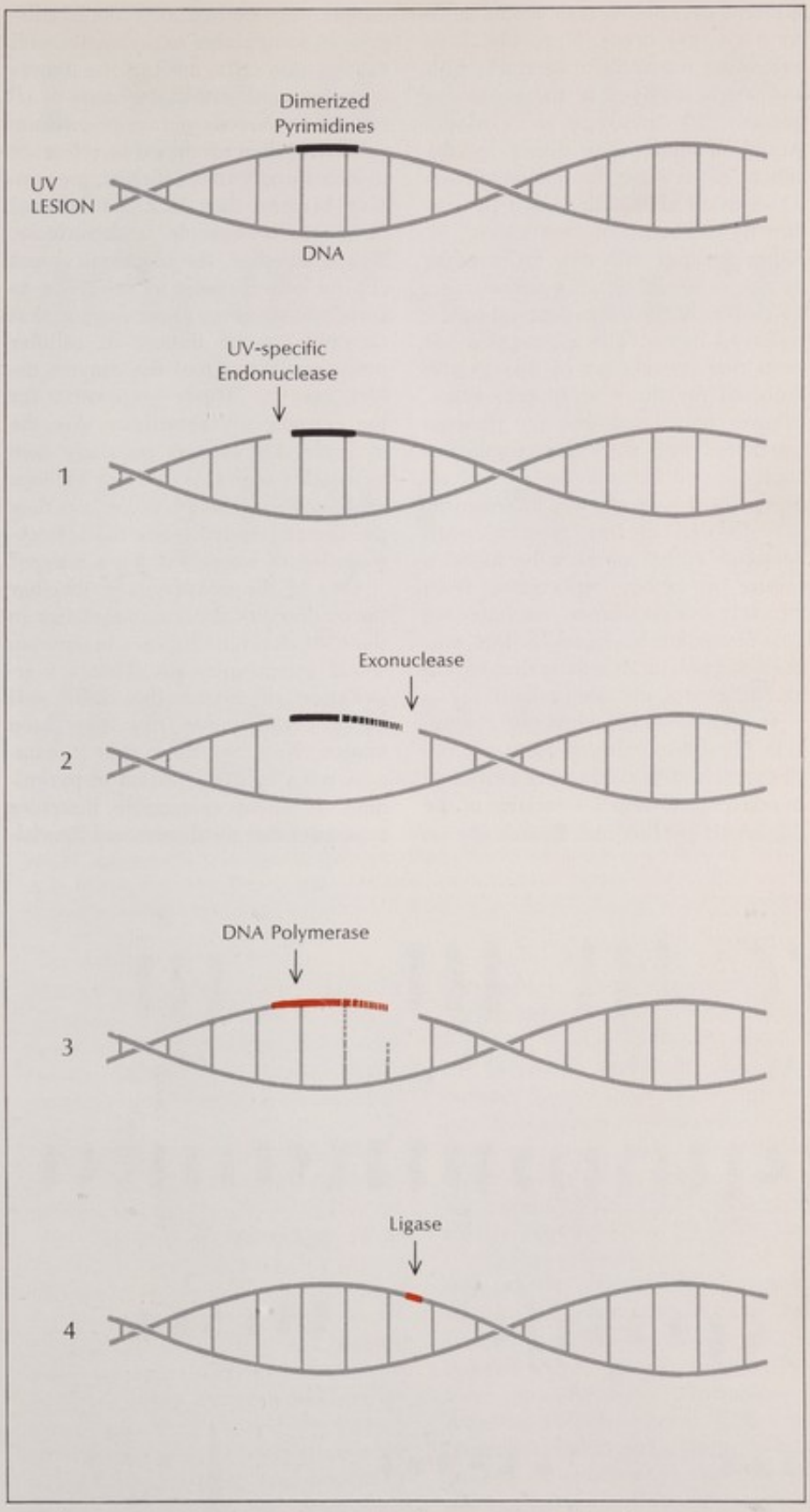


Identification of this reciprocal translocation (1 – 13) in several cells of fibroblast line from a patient homozygous for xeroderma pigmentosum gene indicates derivation of mutant clone from single cell in which translocation had occurred. The other cells in culture had a normal complement of chromosomes.

ment, as was true of the patient's blood lymphocytes. No evidence of increased chromosomal instability was found, but eight of the cells examined showed the presence of reciprocal translocation. These translocations were both unusual and similar enough to permit the conclusion that two pseudodiploid subpopulations of cells, each with an abnormal and unique chromosomal complement, were proliferating in the predominantly diploid culture. Since then, five new fibroblast cell lines have been derived from new biopsies from apparently noncancerous skin of the same patient; distinctly aberrant pseudodiploid clones have been detected in two of them.

In affected homozygotes, the major clinical consequences of xeroderma pigmentosum all appear related to ultraviolet light exposure. At birth the skin seems normal, but severe changes associated with sun exposure are usually evident by age three and are progressive thereafter. The cornea also is susceptible to severe damage. Freckles of varying size and color appear in exposed areas of skin and are accompanied by telangiectasia, increasing dryness and atrophy, and numerous keratoses. Histopathologic study documents the presence of hyperkeratosis, marked atrophy of the dermis with irregular proliferation of some layers, dilatation of vessels, and increased pigment accumulation. Skin cancers, sometimes of multiple types, usually cause death before age 30. Frequently, basal cell and squamous cell epitheliomas appear in such large numbers that it is surprising the patients survive as long as they do. Both benign and malignant tumors of ectodermal and mesodermal origin also emerge with increased frequency.

As noted above, a tendency to chromosomal disruption and rearrangement is not characteristic of xeroderma pigmentosum in cells grown under ordinary laboratory conditions. However, certain biochemical findings in this disorder, to be summarized now, suggest that more than the normal number of rearrangements would be expected if the cells should receive UV damage. (The experiments to demonstrate this conclusively remain to be done.) On exposure of cells (from normal persons as well as from persons with xeroderma pigmentosum) to UV light, dimerization of



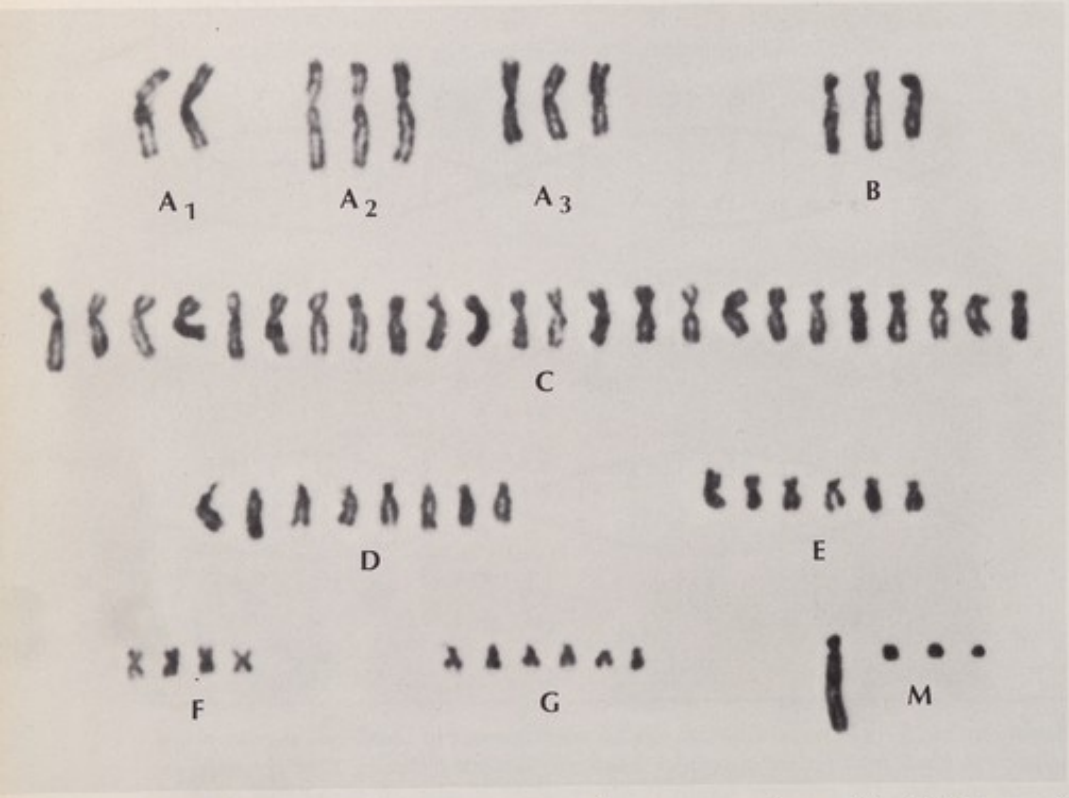
Sensitivity to UV light responsible for clinical consequences of xeroderma pigmentosum appears to result from failure in cellular repair mechanisms following light exposure. In normal situation, UV exposure results in dimerization of adjacent pyrimidine residues in cellular DNA, followed by action of a UV-specific endonuclease and an exonuclease to break DNA strand and excise dimer (1 and 2). Formation of new bases and closure of strand are effected by a DNA polymerase and by ligase (3 and 4). Functional form of UV-specific endonuclease is lacking in xeroderma pigmentosum.

adjacent pyrimidine residues in cellular DNA may occur. Normally these lesions are reparable in the dark, with four steps involved in the repair sequence: 1) breakage in a single strand alongside the dimer by the action of UV-specific endonuclease; 2) excision of the dimerized pyrimidine by the action of exonuclease; 3) filling the gap with new intact bases by the action of a DNA polymerase; 4) closure of the opened strand by the action of ligase. The assumption has been that completion of one step is required for the next to take place. Cleaver postulated that UV damage associated with xeroderma pigmentosum resulted because the normal sequence of these steps was interrupted; in cells derived from patients with xeroderma pigmentosum he found a failure in repair replication, from which it was possible to conclude that step 3 was not being taken. But, was the DNA polymerase defective, or had an earlier step not been taken?

In studies with Setlow and Regan, skin fibroblasts derived from our patient referred to above were examined in reference to step 2 – excision of the dimerized pyrimidine. Essentially no dimers were excised after UV irradiation; in comparable experiments with normal skin cells, most of the dimers were removed within 24 hours of irradiation. Xeroderma pigmentosum cells were then examined in reference to step 1 and found deficient, presumably because they lack a functional form of UV-specific endonuclease. Taken together, the biochemical and clinical observations in reference to xeroderma pigmentosum suggest that accumulation of dimers in cellular DNA, resulting from the enzyme defect, may be largely responsible for the extreme UV sensitivity. Are the multiple skin cancers regularly seen in patients with this disorder perhaps explainable on this basis, i.e., are they due directly or indirectly to the accumulation of unexcised DNA lesions?

One of the unknowns is whether the tendency to dimer accumulation in the DNA chain predisposes to chromosomal rearrangements. That UV irradiation of mammalian cells will break chromosomes has long been known. Now we know that it damages DNA by dimerization of pyrimidine. It seems reasonable therefore to suspect that the dimers and the visible breaks are related to one another in some way. Persistent (unrepaired) aberrant bonding of bases in a DNA strand would appear to be a reasonable starting point for an abnormality related to faulty chromosome duplication during the next S-period of the cell cycle; without better understanding of the molecular events accompanying alteration of chromosome structure, further details of the production of chromatid aberrations are difficult to describe. What can be said is that observations thus far are consistent with the interpretation that unexcised UV-induced pyrimidine dimers may result in increased numbers of chromosomal rearrangements, and that some of these rearrangements are stable and result in a nonlethal alteration in the genome so that they can be retained as marker chromosomes in the complement of a cell that becomes the progenitor of a mutant clone. This should happen more often in xeroderma pigmentosum than in normal cells; it would explain the ease with which we detected pseudodiploid clones in our cell lines.

The critical question then becomes whether a cell with such a chromosomal rearrangement could become the progenitor of a cancer stem line. It is a question that must be asked in reference to all four of the genetic disorders discussed, for although in each disorder a different biochemical mechanism predisposes to cancer, it may also be true that as a group these genes predispose to development of chromosomal mutations and that this is the important feature they have in common. All may thus increase the likelihood of a specific mutation that in itself may constitute conversion to cancer or serve as an integral part of the conversion process.



The presence of marker chromosomes in human cancers is exemplified (M) in case of seminoma of the testis studied by Atkin and Baker. Markers tend to be unique to a given patient's cancer, despite similar clinical expression and pathology. However, chronic myelogenous leukemia is associated with the same marker in almost all cases.

Environmental Factors

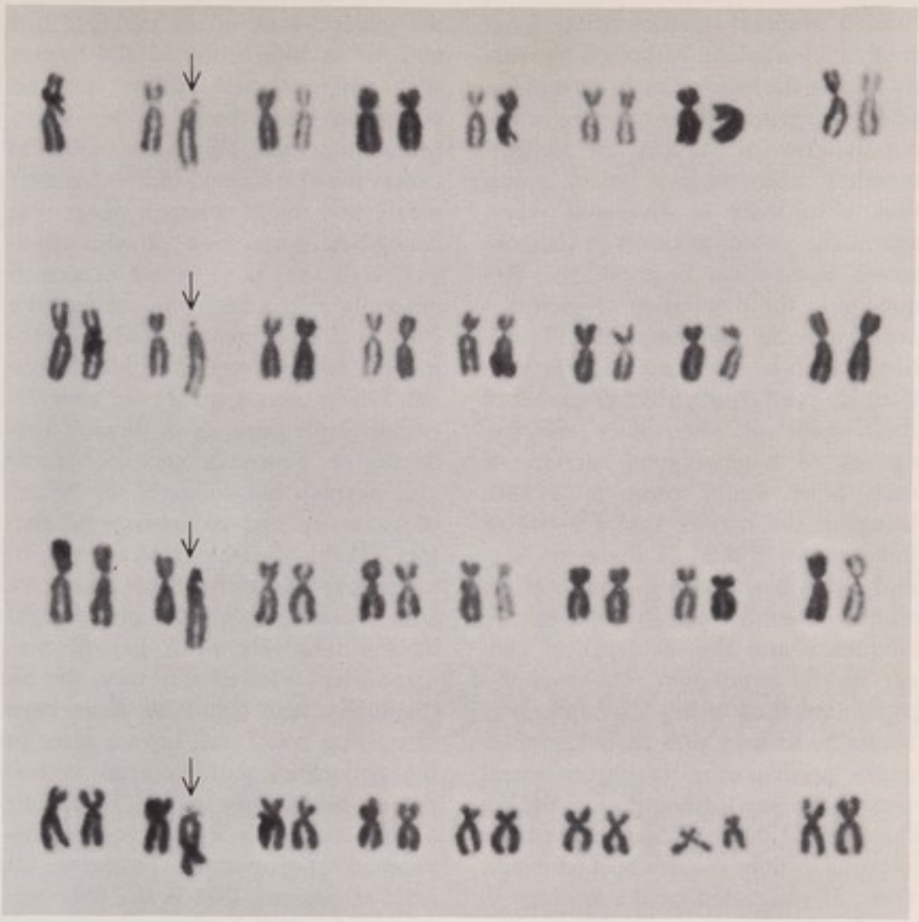
Having said this much about these genetic disorders in which chromosomal anomalies are accompanied by increased cancer risk, it becomes easy to perceive the parallel they offer to environmental situations in which an association has been firmly established between mutation and oncogenesis.

The feature in common? Increased chromosomal instability, displayed by cells taken from many persons belonging to groups with a high risk of

cancer by virtue of their exposure to such environmental factors as certain chemicals (e.g., benzene) or ionizing radiation. Undoubtedly the best known example is provided by the Japanese survivors of the atomic blasts in 1945, in whom mortality from leukemia is significantly higher than expected statistically (see Adelstein, HOSPITAL PRACTICE, January 1973). Among this population, chromosomal disruption and rearrangement have been observed to occur with increased frequency by two independent investigating teams (A. D. Bloom et al and Ishihara and Kumatori). Conducting cytogenetic studies among survivors of atomic irradiation in Japan and near Bikini, they found evidence for the presence of pseudodiploid clones of blood lymphocytes proliferating among normal diploid cells in some of these individuals. On the basis of the uniqueness of the marker chromosomes in such clones, it was concluded that in each case the affected cells were descendants of a single cell in which a rearrangement had taken place.

Among individuals exposed to intensive ionizing radiation for either diagnostic or therapeutic purposes, a similar pattern has prevailed. For example, an increased incidence of leukemia has been recorded among arthritic patients who at one time in their treatment received x-irradiation of the spine. A marked increase in chromosomal abnormalities – chromatid gaps, breaks, rearrangements, new abnormal chromosomes, and clones of circulating lymphocytes with a stable chromosomal rearrangement – was recorded among nonleukemic patients once irradiated for this purpose. Children who during infancy received x-ray therapy to shrink the thymus have also been shown to have an increased incidence of leukemia. Older radiologists appear to have had a shorter life span, as compared with other medical specialty groups, with a higher incidence of leukemia and cancer at least partly responsible.

From such findings the significance of the chromosomal instability and the ensuing new and stable chromosomes remains unclear, because, as indicated, they may be present without cancer. At the same time there must be some explanation for the fact that the same populations show chromosomal abnormalities and an increased expectancy of cancer. The well-known fact that many viruses, including many associated with cancer, are capable of disrupting chromosomes also bears mentioning here.



Environmental situations predisposing to cancer are associated with chromosomal aberration, apparent here in four lymphocytes from Hiroshima A-bomb survivor studied by A. D. Bloom et al. Presence of same structural inversion in all suggests that cells are members of a mutant clone descended from a single cell containing a stable rearrangement.

As already noted, part of the dilemma in relating chromosomal abnormalities to cancer – in all these situations – is lack of knowledge as to whether a mutation present in the established cancer's cells was present at the time conversion occurred in the original cancer cell. If it was, theoretically the cell undergoing conversion by virtue of the mutation could be the progenitor of all cells of the cancer, all of which in turn would contain the mutated chromosome. One possibility is that the original cell could have been part of a nonmalignant clone bearing a rearrangement, one predisposed to conversion because of the chromosomal abnormality. Another is that the cell converting to cancer might have had a normal chromosome complement but that mutation bestowing a proliferative advantage occurred in one of its descendants because of the increased karyotypic instability known to characterize established cancers. The chromosomal mutation would then represent a step in the evolution of the cancer but not have a part in the initial conversion. The latter possibility may seem less plausible since it requires changes to occur in two steps; on the other hand, the possibility of multiple steps in cancerogenesis, at least one of them being mutation, appears to be a tenable hypothesis to some investigators (for example, Knudson and coworkers). C. E. Ford, Grouchy, Ohno, and others have advanced as possible explanations for the emergence and changing clinical features of cancer an evolution of the cancer's karyotype through repeated structural and numerical changes.

Regardless of *how* chromosomal abnormalities and cancer prove to be related, the genetic disorders under dis-

cussion warrant consideration from another viewpoint: Although the rarity of the disorders has been emphasized, the genes responsible are really not that rare when they are lumped together and viewed as a group that, admittedly in divergent ways, shares the ability to result in chromosomal instability. Suppose the frequency in the population of each disorder were in the range of 1:1 million, which is surely an underestimation. Even so, from what is known of their mode of inheritance the frequency of heterozygous carriers of each gene would come to 1:500, bringing the carrier rate for one of the four to 1:125. If there are several more genes with a similar effect on chromosomal stability, the carrier frequency and the incidence of cancer in the population may well approximate each other. There is every reason to assume that there are other genes predisposing to chromosomal rearrangement, although identifying them might prove difficult since their expression may require special conditions, as suggested by the findings in xeroderma pigmentosum. To borrow a term from the microbial geneticist, such genes may be "conditional mutants."

Further studies must be oriented toward heterozygotes as well as homozygotes in reference to disorders associated with chromosomal instability. Our cytogenetic observations in heterozygotes for the gene for Bloom's syndrome (*bl*/+) indicate this, as do Todaro's findings in s v-40-infected cultures of cells heterozygous for the Fanconi's anemia gene. Also, according to Cleaver, some heterozygotes for the xeroderma pigmentosum gene manifest diminished endonuclease activity. Further studies of the heterozygotes are also required to determine whether they as well as homozygotes are at increased risk of cancer.

In considering these and other findings it should be borne in mind that as a rule neither human nor other animal cancer is associated with single gene inheritance; rather, the generality of cancers appears dependent on the genetic makeup at multiple loci and on an interaction of the tissues with environmental factors; i.e., on polygenic inheritance. If this interpretation is valid, the predisposition to cancer may be likened to that for cleft palate and many common congenital anomalies: it may be a "quasi-continuous" trait that is or is not expressed clinically depending on whether the combination of genetic and environmental factors exceeds a critical threshold. Given homozygosity for an overriding single gene, as in Bloom's syndrome or Fanconi's anemia, cancer may develop regardless of the "dose" of cancer-related environmental factors. On the other hand, an x-ray technician exposed to irradiation may be at increased risk even though he might have a relatively weak genetic predisposition. Viewed this way, the individually rare genes we have been discussing could well have a place in the polygenic, multifactorial system determining cancer in man. If so, their effect presumably would be to predispose to chromosomal mutation; at least at present, that is the only feature we know they share.

In summing up, I should like to re emphasize the diversity – and disparity – of the chromosomal changes seen in cancer. Chromosomal instability appears characteristic of cellular systems from which cancer is likely to emerge; but only sometimes does a clinical cancer actually develop. That this unusual tendency to chromosomal disruption and rearrangement is to be found in populations both environmentally and genetically predisposed to cancer suggests strongly that it forms a background from which cancer will emerge with increased frequency. In many human cancers there is a stem line with an abnormal chromosomal complement; in others, there is no evidence of such a stem line or of any chromosomal aberration. These markers appear to differ from one patient to another (except in the case of chronic myelogenous leukemia) although both patients may have the same type of cancer. The great frequency with which marker chromosomes are found suggests they are important; in contrast, their regular absence from some types brings up the possibility that two classes of cancer exist, in only one of which is chromosomal mutation of significance. In established cancers there may be a varied assortment of numerical and structural abnormalities in different cells, and at the same time there very often can be demonstrated one or more unique marker chromosomes present from cell to cell. A tendency to abnormal mitosis has long been recognized as characteristic of cancer by pathologists. It doubtless explains the varying number of chromosomes per cell in a population of cancer cells. I suggest that this process is a different one from the tendency to chromosomal disruption and rearrangement spoken of above, and also from the transmission of the marker chromosomes that characterizes cancerous stem lines, but that the chromosomes are involved in all three may explain the degree of confusion extant in our understanding of the role of chromosomal changes in cancer.

It would appear to me to be a mistake to discount an essential role for chromosomal abnormality just because the cytogenetic findings are complex and sometimes confusing. One could imagine a parallel confusion in reference to mongolism at a time before its chromosomal basis (trisomy 21) was recognized. What if such a patient had developed leukemia, as mongols frequently do, and been treated by x-irradiation? A cytogeneticist studying the case would have identified not only an extra chromosome in all cells (the trisomy) but probably a different aneuploid subpopulation in cells of the bone marrow and blood (the leukemic cells). He might also have found evidence of chromosomal disruption and rearrangement in lymphocytes in culture (from the radiation). If intent on a single explanation for the findings, he would have had difficulty in finding one. If he had offered one, would it have been correct?

Section Two

Basic Biochemical Genetics

5

Inherited Hemoglobinopathies

C. LOCKARD CONLEY *and* SAMUEL CHARACHE
Johns Hopkins University

In examining the large area of genetic disease, the hemoglobinopathies are a logical place to start. They present significant public health problems in some areas of the world, and from the standpoint of pathophysiology they are of unique importance. In most of these disorders, not only can we identify the specific abnormality of hemoglobin synthesis or of the hemoglobin molecule itself, but we can inferentially pinpoint the precise "lesion" in the genetic mechanism which produces it. In a number of instances abnormal molecular structure can be directly related to abnormal function and in turn to the production of clinical symptoms. This clarity of understanding demonstrates that such genetic concepts as "dominant" and "recessive" are descriptive abstractions that must be interpreted in the light of quantitative measurements of the biochemical processes involved in inherited traits.

Genetics, as the biochemist sees it, is largely concerned with the synthesis of proteins, whether these be the quantitatively important body constituents such as hemoglobin, the equally vital regulatory proteins – enzymes and hormones – or others including immunoglobulins and binding proteins such as transferrin. Genetically determined differences among individuals are thought to stem from differences in the structure of specific proteins or in the rates at which they are produced.

The hemoglobin of the normal adult (Hb A) contains four amino acid chains – two "alpha" chains, each with 141 amino acid residues, and two "beta" chains, each with 146 residues; descriptively it is identified as $\alpha_2 \beta_2$. (The red cells of normal adults also contain a minor component, Hb A_2, consisting of alpha and delta chains, whereas the hemoglobin of the fetus and newborn is predominantly Hb F, consisting of alpha and gamma chains.) The complete primary structure of each of the polypeptide chains has been established.

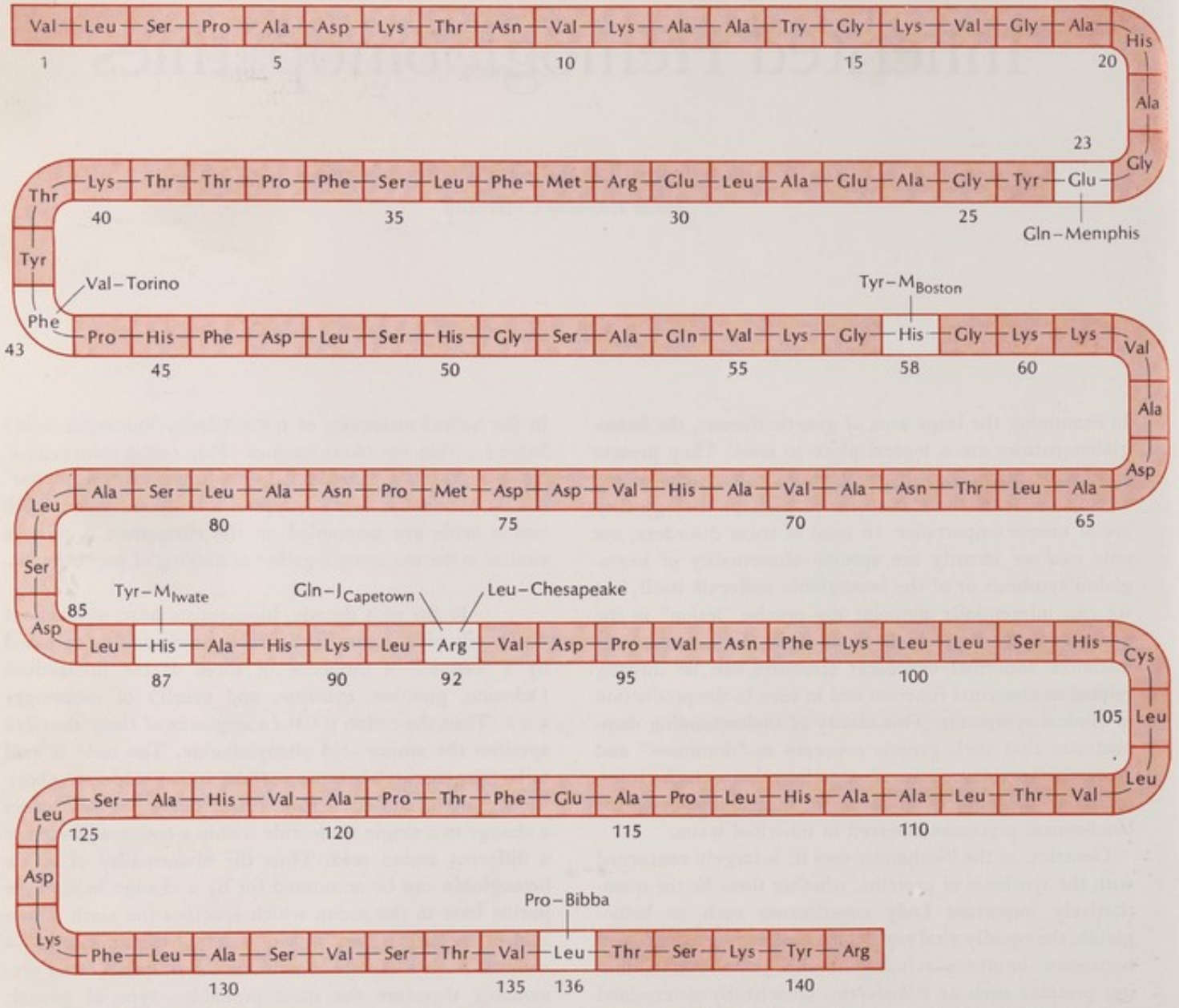
The properties of any protein are determined by the number, identity, and sequences of its amino acid residues. The information which specifies these features is encoded in the helical molecules of DNA (deoxyribonucleic acid) lodged within the chromosomes. This coded information, transferred to the cytoplasm by molecules of "messenger" RNA (ribonucleic acid), supplies a template upon which amino acids are assembled on the ribosomes, a process similar to the snapping together of a string of poppit beads.

In the past decade, biochemists have worked out the details of the code itself. Each amino acid is specified by a "codon" – a sequence of three of the nucleotides (adenine, guanine, cytosine, and uracil) of messenger RNA. Thus the codon UUU (a sequence of three uracils) specifies the amino acid phenylalanine. The code is said to be "degenerate" because certain amino acids are specified by more than one codon. Nevertheless, in most cases a change in a single nucleotide within a codon will specify a different amino acid. Thus the abnormality of sickle hemoglobin can be accounted for by a change in a single purine base in the codon which specifies the sixth amino acid of the beta chain (GAA → GUA, glutamic acid → valine). Such a change constitutes the simplest – and presumably therefore the most probable – type of genetic mutation and accounts for most of the inherited abnormalities of hemoglobin. Rarely, two amino acid substitutions occur in a single polypeptide chain, as in Hb C_{Harlem}; presumably two independent mutations occurred within the same gene (a gene is that portion of a strand of DNA that determines the structure of a single polypeptide.)

In a few abnormal hemoglobins (e.g., Hb Freiburg and Hb Gun Hill) amino acid residues are actually missing from a polypeptide chain. This type of abnormality can be accounted for by loss of a segment of DNA ("deletion" of part of the gene), presumably as a result of nonhomologous crossing-over of chromosomes. A hemoglobinopathy of no clinical importance but of great theoretical interest is "hereditary persistence of fetal hemoglobin," an anomaly that can be explained by deletion of two closely linked genes (those for the beta and delta chains). Persons homo-

Sequence of Amino Acid Residues in

Alpha Chain



Ala – Alanine **Arg** – Arginine **Asn** – Asparagine **Asp** – Aspartic Acid **Cys** – Cysteine **Gln** – Glutamine **Glu** – Glutamic Acid **Gly** – Glycine **His** – Histidine **Ileu** – Isoleucine

Sequences of the 141 amino acid residues in the alpha and the 146 in the beta chains of normal human hemoglobin (Hb A) are shown schematically on these pages. (The normal hemoglobin molecule, of course, includes two alpha and two beta chains, or a total of 574 amino acid units.) Indicated at their appropriate positions are the substitutions or deletions responsible for the ab-

zygous for this condition synthesize neither of the two normal hemoglobins of the adult, and their hemoglobin is exclusively fetal in type. The hemoglobins Lepore, Miyada, and P_{Congo} are fascinating because they contain portions of both the normal delta and beta chains; their production can be explained by loss of DNA from segments of adjacent genes, with formation of a "fusion" gene.

The thalassemias are disorders of hemoglobin synthesis in which one or the other polypeptide chain of hemoglobin is produced at a retarded rate. Those chains which are produced have a normal structure. The cause of the abnormality has not been identified with certainty but is associated with insufficient messenger RNA.

The clinical manifestations of the hemoglobinopathies are determined by both the quantities and properties of the molecules produced. The properties in turn are related to the structure of the polypeptide chains. To each of the four globin chains is attached a "heme" group, a porphyrin ring structure containing an atom of iron. In one view, the major function of the globin chains and the heme rings is to adjust the environment of these four iron atoms so precisely that

Polypeptide Chains of Human Hemoglobin

Leu—Leucine **Lys**—Lysine **Met**—Methionine **Phe**—Phenylalanine **Pro**—Proline **Ser**—Serine **Thr**—Threonine **Try**—Tryptophan **Tyr**—Tyrosine **Val**—Valine

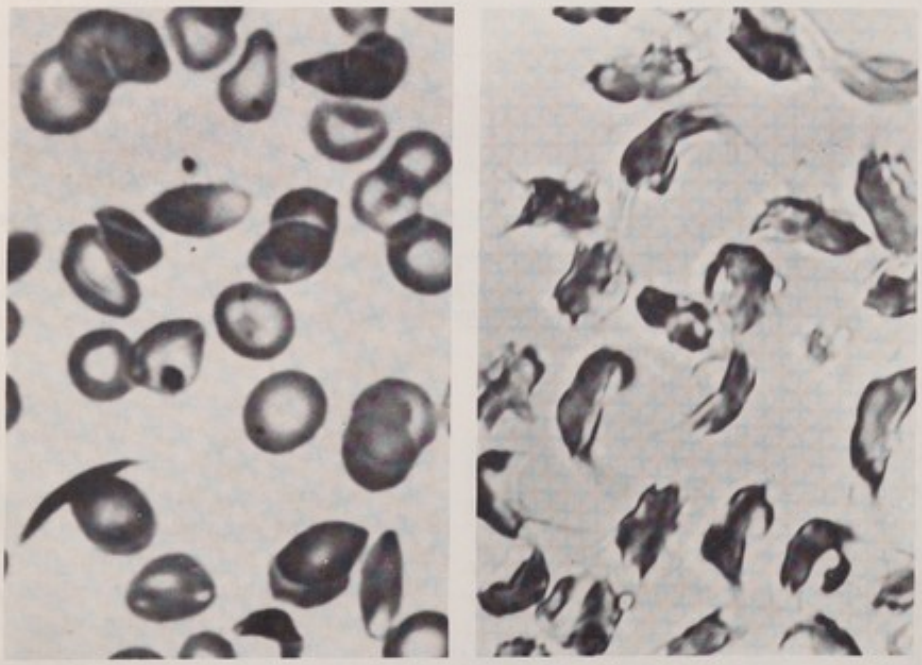
normal hemoglobins named. Many other abnormal hemoglobins have been identified, most of which do not affect the functioning of the hemoglobin molecule. This rendering emphasizes (but is not necessarily a complete listing of) those in which the mutations produce clinical symptoms. A table of syndromes associated with such hemoglobin abnormalities appears on page 58.

they can reversibly combine with oxygen under normal physiologic conditions. In addition, the particular fashion in which the globin chains and their heme rings are bonded together lends unusual stability to the molecule. Oxygenated hemoglobin has one stereochemical conformation called "R," or "relaxed." Deoxyhemoglobin has another called "T," or "tense," because of extra intramolecular bonds which form and bind the alpha chains together more tightly than in the "R" state. The "T" conformation is further stabilized by 2,3 diphosphoglycerate, an organic ion present in high concentration within red cells, which can bind together the beta chains of deoxyhemoglobin.

If these precise stereochemical relationships are disturbed, abnormal oxygen binding and decreased stability may result. Three examples may be cited: a) In the variety of thalassemia in which insufficient alpha chains are produced, a relative excess of beta chains accumulates and the chains combine to form Hb H (β_4). Although the beta chains are individually normal, the structure of the molecule is deranged. Hb H has an abnormally high affinity for oxygen, which cannot be released to the tis-

Tactoids shown at left are composed of deoxygenated Hb S in a stroma-free preparation; resemblance to sickled red cells (same enlargement, right) confirms it is the hemoglobin, not the cell, that causes the deformation of the erythrocytes in sickle cell anemia. These illustrations are from a study by Dr. John W. Harris.

sues. The hemoglobin is not only physiologically useless but is very unstable, tending to precipitate within the red cells and cause their premature destruction. b) In hemoglobins Rainier and Bethesda, bonds which normally stabilize the "T" conformation of deoxyhemoglobin can no longer be formed. As a result, the "R" conformation becomes a more favored state: at any given oxygen pressure Hb Rainier will be more oxygenated than Hb A. These same bonds are in part responsible for the pH dependence of oxygen affinity, and as a result Hb Rainier and Hb Bethesda exhibit a diminished Bohr effect. c) Binding of 2,3 diphosphoglycerate to deoxygenated gamma chains of Hb F is less avid than to the beta chains of normal Hb A. The "R" (relaxed) conformation of fetal hemoglobin is favored, since the intramolecular bonds which bind the "T" conformation together are weaker. Because fetal blood has a greater oxygen affinity, persons homozygous for "hereditary persistence of fetal hemoglobin," whose blood contains only the fetal form, obviously will manifest enhanced blood-oxygen affinity. Release of oxygen to the tissues is decreased, and polycythemia ensues, in such individuals and in carriers of Hb Rainier and Bethesda, as a sequel to the compensatory increase in erythropoietin production.



Blood from Hb S homozygote has sickled forms amid normal (left); when same blood was treated with reducing agent, deoxygenation caused all the erythrocytes to sickle (right).

Such striking abnormalities, however, make up only a small proportion of the hemoglobinopathies. At first glance, one might suppose that a change in any one of the 574 amino acid residues of the hemoglobin molecule would render it incapable of functioning normally. Actually, substitutions on the external surface have little effect on molecular function, presumably because the altered side chains protrude outwards, rather than toward the interior of the molecule where the heme rings are buried. Many amino acids are similar in size and charge, and these can be substituted one for the other even within the interior without disastrous consequences resulting.

That certain amino acids are essential for normal molecular function is suggested by comparing the hemoglobins of different species of animals: seven residues are "invariant" in all species examined, indicating that these may be of critical importance. Hence it is noteworthy that significant clinical abnormalities have resulted when mutation occurred at some of these sites, as in the Hb M syndromes.

Four of the hemoglobins M are produced through substitutions at the invariant histidine residues on either side of the iron atom in the heme ring. Ordinarily this atom is in the reduced or ferrous state, which permits reversible binding of oxygen. The presence of oxygen tends to change the iron to the ferric form, producing methemoglobin, a compound that cannot combine reversibly with oxygen. But the erythrocytes contain enzymes designed precisely to reverse that process, so

that nearly all hemoglobin iron is normally in the ferrous state. In the hemoglobins M, substitution of tyrosine for histidine renders ferric iron so stable that those enzymes cannot reduce it. Methemoglobinemia in heterozygotes does not seriously compromise oxygen carrying capacity of the blood, but it does produce striking cyanosis. The homozygous condition has not been encountered and presumably would be incompatible with life.

In another hemoglobinopathy, methemoglobinemia is conditional rather than absolute. Hb Zürich involves replacement of one of the invariant histidine residues by arginine rather than tyrosine. Under normal conditions, heterozygotes show no unusual symptoms, aside from a slightly shortened life-span of the erythrocytes. When such patients are given sulfonamides or other oxidant drugs, methemoglobinemia develops; moreover, the hemoglobin denatures and precipitates, which produces fulminant hemolytic anemia.

Hemoglobins Chesapeake and Malmö have increased affinity for oxygen. Their discovery led to the recognition of another critical area of the molecule, the interface between the alpha and beta chains. Substitution here may lead to increased affinity for oxygen, but the stereochemical mechanism of this effect is still unclear. Carriers of both hemoglobins are polycythemic. Since hemoglobin Malmö is very difficult to detect by electrophoresis, it becomes clear that measurement of oxygen affinity is a necessary part of the investigation of all patients with unexplained polycythemia.

Study of patients heterozygous for Hb Köln and certain other abnormal hemoglobins led to recognition of the entire "heme pocket" as another critical region of the globin chain. Substitutions of amino acids lining this crevice by others that are too large, or which alter the distribution of electric charges, or which permit formation of abnormal cross links within the molecule, decrease its stability. Precipitation of hemoglobin within the red cell is associated with hemolytic anemia.

In general, mutations on the surface of the molecule are by all odds the least serious. Of the more than 60 that are known, only a rare one is ordinarily harmful in the heterozygous state and only three are known to be harmful to homozygotes. Oddly enough, it is one of these, Hb S, that is both the most common and in some respects the best understood of all hemoglobinopathies. Rather than producing an *intra*molecular dysfunction, abnormal reactions *between* molecules of Hb S in the red cells of homozygotes result in a distinct disease – sickle cell anemia.

When deoxygenated, molecules of Hb S stack together to form rod-like structures (fibers) that tend to cluster together in parallel arrays. This molecular aggregation is responsible for sickling. Whether the fibers form in cell-free solution (as shown in this electron micrograph at 46,000-fold magnification – the bar is one micron long) or in sickle cells themselves, the morphologic appearance is the same. Each fiber is approximately 170 A in diameter and interfiber distances are constant. (Courtesy John F. Bertles and Johanna Döbler, Columbia University)

This disease was discovered by the late J. B. Herrick, who in 1904 made an office examination of a Negro student recently arrived from the West Indies. Under the microscope, the young man's blood showed a number of peculiar crescent-shaped erythrocytes. Herrick kept his patient under observation for the next six years, by which time the man had displayed many of the manifestations we now recognize as typical of the disease. He presented his findings at the 1910 meeting of the Association of American Physicians, but his paper stimu-

Clinical Manifestations Associated with Some Abnormal Hemoglobins

Disorder	*Abnormal Hb*	*Structural Change*	*Comments*
HEMOLYTIC ANEMIA	H	$alpha_2$ $beta_2 \rightarrow beta_4$	Unstable hemoglobin occurring in some forms of alpha-thalassemia; precipitation of hemoglobin and hemolysis are accelerated by certain drugs
	S	beta 6 glu → val	Forms molecular aggregates when deoxygenated, producing sickle cell anemia in homozygotes
	C	beta 6 glu → lys	Low solubility lessens plasticity of red cells, causing hemolytic anemia in homozygotes
	E	beta 26 glu → lys	Mechanisms unknown
	D_{Punjab}	beta 121 glu → gln	
	Zürich	beta 63 his → arg	Unstable hemoglobin precipitated by certain drugs, producing hemolytic anemia in heterozygotes
	Torino	alpha 43 phe → val	Unstable hemoglobin causes congenital nonspherocytic hemolytic anemia in heterozygotes; precipitated hemoglobin tends to form inclusion bodies within red cells, under certain conditions
	Bibba	alpha 136 leu → pro	
	Savannah	beta 24 gly → val	
	Genova	beta 28 leu → pro	
	Hammersmith	beta 42 phe → ser	
	Bristol	beta 67 val → asp	
	Christchurch	beta 71 phe → ser	
	Shepherds Bush	beta 74 gly → asp	
	Borås	beta 88 leu → arg	
	Santa Ana	beta 88 leu → pro	
	Sabine	beta 91 leu → pro	
	Köln	beta 98 val → met	
	Casper	beta 106 leu → pro	
	Wien	beta 130 tyr → asp	
CYANOSIS due to methemoglobinemia	M_{Boston}	alpha 58 his → tyr	Methemoglobin causes cyanosis in heterozygotes; some also have evidence of hemolytic anemia
	M_{Iwate}	alpha 87 his → tyr	
	$M_{Saskatoon}$	beta 63 his → tyr	
	$M_{Milwaukee}$	beta 67 val → glu	
	$M_{Hyde\ Park}$	beta 92 his → tyr	
	Freiburg	beta 23 val *deleted*	
CYANOSIS due to increased deoxyhemoglobin	Kansas	beta 102 asn → thr	Decreased oxygen affinity of hemoglobin causes cyanosis in heterozygotes
POLYCYTHEMIA	$J_{Capetown}$	alpha 92 arg → gln	Increased oxygen affinity of hemoglobin hinders release of oxygen to tissues, causing compensatory polycythemia in heterozygotes
	Chesapeake	alpha 92 arg → leu	
	Olympia	beta 20 val → met	
	Malmö	beta 97 his → gln	
	Kempsey	beta 99 asp → asn	
	Yakima	beta 99 asp → his	
	Hiroshima	beta 143 his → asp	
	Rainier	beta 145 tyr → cys	
	Bethesda	beta 145 tyr → his	
ANEMIA due to suppressed erythropoiesis	Seattle	beta 76 ala → glu	Decreased oxygen affinity of hemoglobin enhances release of oxygen to tissues and inhibits production of erythropoietin
	Yoshizuka	beta 108 asn → asp	
HYDROPS FETALIS	Barts	$alpha_2$ $gamma_2 \rightarrow gamma_4$	Unstable hemoglobin with high oxygen affinity occurring in high concentration in stillborn fetuses with homozygous alpha-thalassemia

lated no discussion. A few years later, a St. Louis anatomist found that sickling could be induced by placing a patient's blood sample under a sealed cover slip. This procedure became the standard test for the sickling phenomenon, but for 10 years nobody speculated on why it occurred under these conditions.

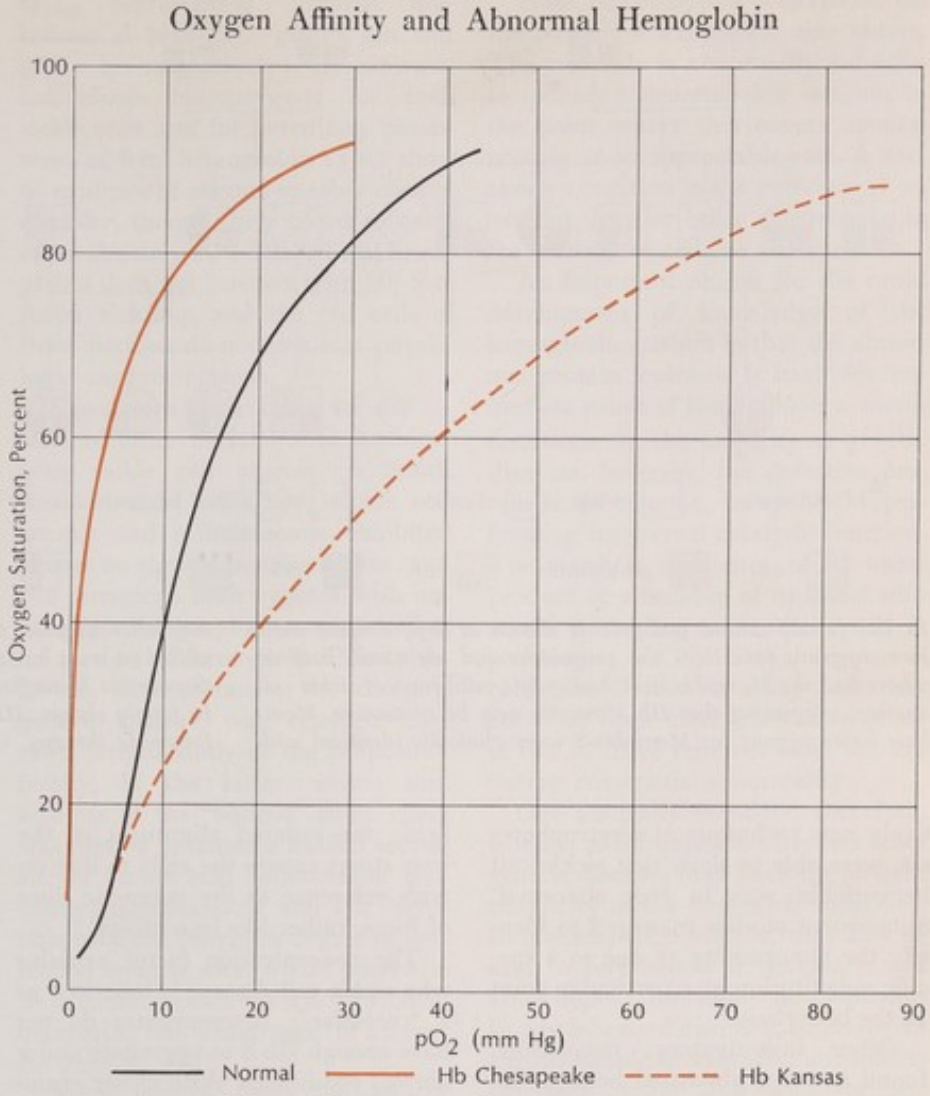
In 1927, Hahn and Gillespie, two Indianapolis physicians, observed a child with sickle cell anemia, the first case in which splenectomy was reported. Stimulated by this case, they conducted, with homemade equipment, an extraordinarily sophisticated study that showed sickling was caused by reduction in oxygen tension and pH in the blood. Moreover, it was the hemoglobin that was responsible; erythrocyte "ghosts," from which the hemoglobin had been extracted, would not sickle.

Since sickling depends on low oxygen tension, it occurs predominantly on the venous side of the circulation, throughout the body but especially in the spleen, where the slow movement of low-oxygen blood gives sickling time to develop. The result can be a vicious cycle: the relatively rigid sickle cells block some capillary channels, which further slows the flow of blood, which permits further reduction of the oxygen tension, which produces more sickle cells, and so on. The rigid and deformed erythrocytes are rapidly removed from the circulation, giving rise to chronic hemolytic anemia. Even more important, sickled erythrocytes block small blood vessels, producing infarcts in various organs and giving rise to painful "crises," destructive lesions, and premature death.

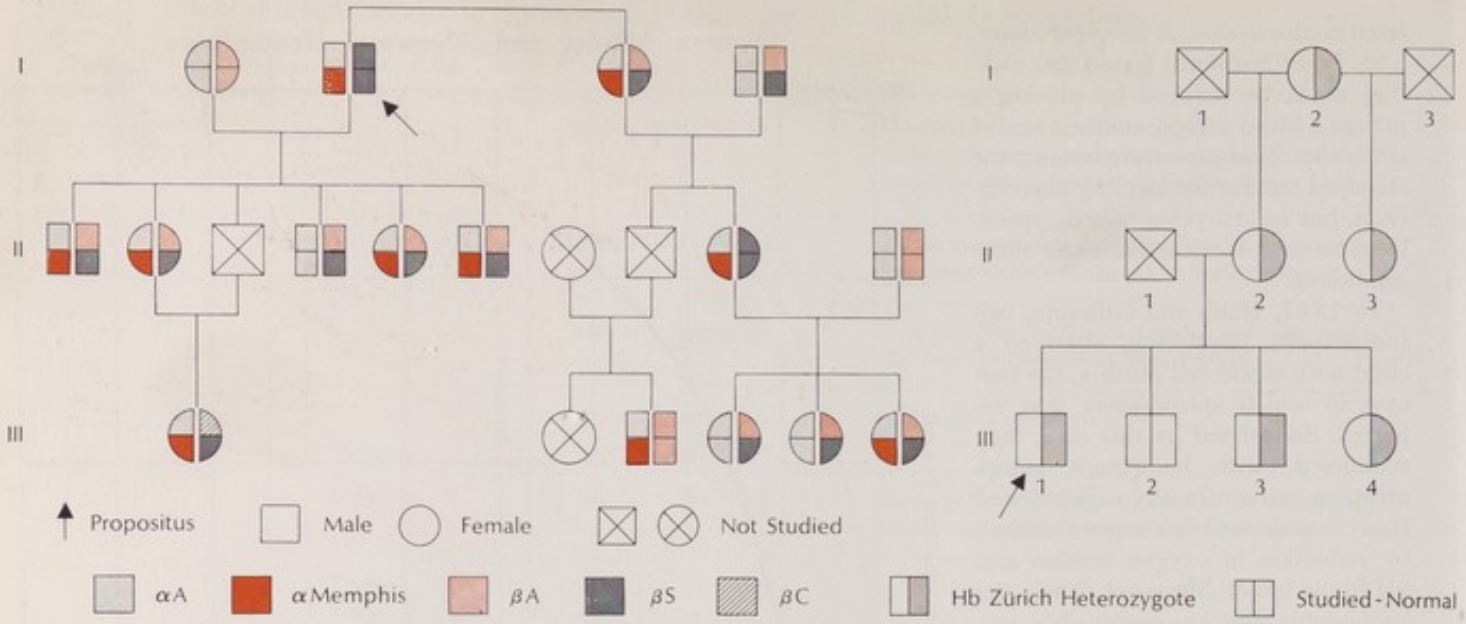
Epidemiological studies revealed why the trait is so common among Negroes. It was found to be peculiarly prevalent in the moist areas of west and central Africa – precisely the areas that were the center of the American slave trade. Specifically, the incidence of the trait (as high as 50% in some communities) showed a high correlation with the presence of the malignant falciparum form of malaria. It has since become apparent that heterozygous sickle cell trait is a positive advantage in these regions, because it partially protects the erythrocytes against invasion by the falciparum organism. How this occurs is unknown. Sickle cell trait does not reduce the incidence nor the severity of malaria in adults, but it evidently gives some protection to infants, improving their chances of surviving their first attack of falciparum malaria – and thereby developing antibodies that can partially protect them against further infections with the endemic disease. The protective effect of sickle cell trait explains the persistence of an otherwise disadvantageous mutation that in malaria-free regions would have been eliminated quite rapidly by natural selection. It also explains how Negro communities survived and thrived in an environment European colonizers described, with some justice, as "the white man's graveyard."

Long considered a "racial characteristic" of Negroes, sickle cell trait has also turned up among some Mediterranean peoples – though its incidence is considerably less. It seems to be particularly prevalent in areas such as Greece, Sicily, and Turkey, where malaria is or was endemic; alternatively, however, it may simply reflect a gradual drift of genes northward from tropical Africa. It even turns up occasionally in people of English stock, presumably as a result of intercourse between the Mediterranean world and northern Europe.

Elucidation of the molecular basis of sickle cell disease began in 1945 with a Pullman-car conversation between the great clinical investigator William B. Castle and the distinguished chemist Linus Pauling, in the course of which Castle remarked on the evidence for some sort of molecular alignment or orientation in sickled erythrocytes. Subsequently, Pauling and his associates, using the then rela-



O_2 affinity of normal hemoglobin is compared with that of Hb Chesapeake, in which it is increased – hindering release to tissues and stimulating erythropoiesis – and with that of Hb Kansas, in which O_2 affinity is decreased, causing cyanosis in heterozygotes.



In the family whose pedigree is shown at left, despite their homozygosity for Hb S the propositus and his niece, both of whom had the Memphis trait, had only a mild form of sickle cell anemia, suggesting that Hb Memphis may be protective. Members heterozygous for Memphis-S were clinically identical with A-S heterozygotes. C trait appeared in third generation, transmitted (it must be assumed) by unstudied father. Hb Zürich predisposes to hemoglobin instability in presence of certain drugs. In family shown, III-1 had acute hemolytic anemia during sulfonamide therapy, III-2 during urinary tract infection. I-2 was

tively new technique of electrophoresis, were able to show that sickle cell hemoglobin was in fact abnormal. Subsequent studies managed to identify the abnormality as due to a specific substitution at a particular point on the beta chain.

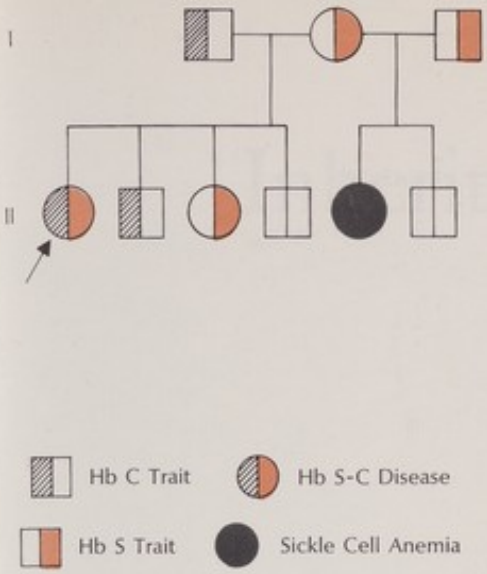
Other investigators, meanwhile, found how the abnormal hemoglobin (Hb S) produces abnormal cells. When deoxygenated, Hb S is insoluable, and in the high concentration in which it occurs in red cells forms boat-shaped liquid crystals. These "tactoids," which can be seen with the light microscope, deform the red cell and produce the rigid sickle-shaped erythrocyte. We can now put together a picture of how the abnormal molecules behave. The substitution of valine on the surface of the molecule provides two reactive sites, one for each beta chain; when the molecule is deoxygenated these sites react with complementary sites on adjacent molecules producing linear strands. Secondary reactions bind these fibers together to form twisted cables. These linear arrangements, which are in a sense polymerized hemoglobin, can actually be seen when sickle cells are examined under the electron microscope. The alignment of the molecules can also be shown when sickle cells are subjected to a strong magnetic field: the ordered alignment of the iron atoms causes the cells to line up with reference to the magnetic lines of force, rather like iron filings.

The concentration factor explains why sickle cell disease is inherited as a "recessive": heterozygotes do not have enough Hb S to aggregate under normal conditions. With closer examination, however, one finds that "dominant" and "recessive" really depend on how one looks at things. From the molecular standpoint, sickle cell trait is a "co-dominant," since heterozygotes produce both Hb S and Hb A (the proportion of Hb S is somewhat less than half, because – for unknown reasons – it is synthesized at a slightly slower rate). From the cellular standpoint, sickle trait is a "dominant," since all of the red cells of the heterozygote will sickle when anoxic conditions are present.

And from the pathologic standpoint, the trait is essentially "recessive," since sickle cell disease ordinarily occurs only in homozygotes. But under rare circumstances even the red cells of sickle trait heterozygotes undergo intravascular sickling and produce clinical manifestations. For instance, a number of Negro soldiers, while being airlifted over the Rocky Mountains, developed acute abdominal pain which was traced to infarction of the spleen. The precipitating cause was mild anoxia (the planes were not pressurized) which, though harmless to normal individuals, was sufficient to induce sickling in these heterozygotes, though they had previously shown no symptoms of the disease.

Untoward effects have been observed in heterozygotes suffering from cyanotic heart disease (eliminated when the defect was surgically corrected), in acute alcoholism or narcotic overdose, and under anesthesia, where they presumably result from decreased ventilation and acidosis. Since sickling is dependent upon deoxygenation of hemoglobin, the "expression" of the gene (another genetic abstraction) is influenced by environmental factors that affect intracellular oxygen tension, including, for example, local conditions of blood flow.

Further complexities arise when we find that the "dominance" or "recessiveness" of sickle cell trait can also be modified by the presence of other mutations. There are, for instance, individuals who are heterozygous also for Hb C, Hb D_{Punjab}, or Hb O_{Arab}. In homozygotes, the first two of these produce a rather mild hemolytic anemia (O_{Arab} homozygotes have not been encountered); heterozygotes (A-C, A-D, or A-O) show no symptoms what-



jaundiced in youth, III-3 and III-4 had neonatal jaundice, II-3 neither jaundice nor hemolytic anemia. Heterozygosity for both Hb S and Hb C (right) is more detrimental than heterozygosity for S alone.

ever. But individuals heterozygous for *both* sickling and one of these other mutations (S-C, S-D, or S-O) are sicker than heterozygotes of any one of them, or even than the C-C or D-D homozygotes; in the case of the S-D and S-O combinations, the results are almost as serious as in the homozygous sickle cell anemia (S-S). In these instances, the second abnormal hemoglobin interacts with Hb S and enhances the tendency of red cells to sickle; the clinical manifestations are those of sickle cell anemia.

So here we have three instances in which, pathologically speaking, two different "recessive" genes add up to one "dominant" trait. The reasons are still obscure. Probably significant is that Hb C, D_{Punjab}, and O_{Arab} involve surface modifications of the beta chain, like Hb S itself (D_{Punjab} and O_{Arab} representing different mutations at beta-121: glu → gln and glu → lys respectively). By contrast, individuals heterozygous for both sickle trait and for hereditary persistence of fetal hemoglobin (S-F) show no evidence of anemia or other clinical disorder, though their blood contains approximately 70% Hb S. Fetal hemoglobin does not interact with Hb S to foster sickling, and the red cells of these persons do not sickle at physiologic oxygen tensions.

Even more remarkable are the effects of other mutations on homozygous sickle cell anemia. A Saudi Arabian child with both sickle cell anemia and α-thalassemia exhibited almost no clinical manifestations, and two American homozygotes with unusually mild symptoms were found to be carriers of hemoglobin Memphis. In the first instance, the complex genetic picture was clarified only after careful study of the proposita's family; in the latter, amino acid analysis of the "normal" alpha chain was needed to detect a hidden second site of amino acid substitution. Evidently, α-thalassemia and Hb Memphis trait can provide a degree of protection against what would otherwise be a more serious disease – in effect decreasing the "penetrance" of a harmful gene.

More simple, from the genetic standpoint, are such conditions as Hb Zürich and Hb Chesapeake. Like Hb S, these abnormal molecules are inherited as co-dominants, meaning that in heterozygotes both the abnormal and normal hemoglobin are present. But since the abnormal reactions to which they are subject are purely intramolecular, high concentrations of the abnormal molecules are not required to exert pathologic effects. As a result the conditions are inherited as dominants.

Note, however, that with Hb Zürich the matter of expression also enters. The molecule is abnormally sensitive to oxidative denaturation but not to the point where this occurs spontaneously at an appreciable rate. A necessary condition is the presence of an oxidant drug or other disturbance in the immediate cellular environment.

An important reason for the rapid development of knowledge of the hemoglobinopathies is that the abnormal protein molecule is itself the immediate cause of the pathologic manifestations. In the majority of genetic diseases, however, the defective protein is an enzyme, incapable of performing its normal catalytic function. The result is deficiency of its usual product or a buildup of its usual substrate. In either event the result is pathologic, but the actual biochemical lesion (which in many cases has not even been identified) may occur at two or three removes from the initiating enzymatic abnormality.

One naturally wonders, therefore, whether as information increases other genetic diseases may have the elegant simplicity of the hemoglobinopathies. In one respect, however, they can be said to foreshadow a necessary and (one hopes) inevitable development. In any inherited disorder, be its mechanisms simple or complex, our understanding cannot stop with the flat statement that the condition is "dominant" or "recessive," that it does or does not "express" itself fully, and so on. We need to determine precisely what processes are involved in dominance or recessiveness; precisely why, and under what conditions, expression does or does not take place. Such knowledge can open the way to therapeutically blocking the expression of an otherwise harmful gene and conceivably even to counteracting it genetically.

6

Inherited Enzyme Defects

W. EUGENE KNOX
Harvard University

Taken in its literal sense, the phrase "inherited enzyme defects" defines an area of pathology that is both too large and too diverse for useful discussion. The reason lies in the biochemical nature of heredity itself, which is based on the control of protein synthesis by DNA in the cell nucleus. So far as we now know, proteins are the only body constituents whose structures are hereditarily determined.

It follows, then, that all hereditary diseases can be grouped into two broad categories. In the first, the pathology is the direct consequence of some error or inadequacy in protein synthesis; presently the best known examples are the hemoglobinopathies, in which the body synthesizes hemoglobin that is qualitatively or (in some cases) quantitatively inadequate to perform its normal job of oxygen transport (see Conley and Charache, Chapter Five). The second category includes all other genetic diseases, i.e., those in which the pathologic process occurs at least one remove from the primary protein abnormality. The latter is presumed to involve one or more of the body's enzymes or enzyme-like substances – the proteins that catalyze the myriad biochemical processes involved in normal physiologic functioning.

But this presumption, though logical in terms of what we now know about heredity, is still only a presumption. In many genetic diseases the specific enzyme defect, assuming it exists, is not yet known; in some, which result from the deletion or duplication of an entire chromosome (e.g., Down's syndrome), the enzyme defects involved and their biochemical consequences must surely be of extraordinary complexity. For simplicity, therefore, I shall devote this article chiefly to conditions in which the enzyme deficiency has been identified and in which the biochemical sequelae of that deficiency are at least reasonably well understood.

An enzyme performs its catalytic function by binding a molecule of some substrate compound, plus on occasion another atom or group of atoms, to "active sites" on its own intricately structured protein chain. The binding, it is thought, distorts the substrate molecule in such a way that it loses part of its substance; alternatively, by forcing it into a particular spatial relationship with another atom or radical, the enzyme causes the two to join together. The presence of electrical charges at the active sites may also play a part in these transformations. In this manner, the enzyme quickly mediates a chemical reaction which in its absence, and at the moderate temperatures, pressures, and concentrations typical of living organisms, would take place very slowly or not at all.

A change in even one of the amino acid residues that make up the enzyme molecule could render it structurally incapable of performing its catalytic function. But not every change is a catastrophic one. Like the hemoglobin molecule, enzymes can undergo considerable variation without completely losing their functional properties; with enzymes, indeed, a degree of variation appears to be the rule rather than the exception. Different organs of the body are known to produce "isozymes" – enzymes that are chemically different but functionally identical, or at any rate identical enough to do the same chemical job in the same way.

If, however, the enzyme molecule is changed too radically by a genetic mutation (or, of course, if it is not produced at all), catalysis cannot take place. In most hereditary enzyme defects we do not as a matter of fact know whether the enzyme is structurally inadequate or missing altogether. For relatively few enzymes do we even know the complete molecular formula; nearly all our enzyme tests are purely functional ones, showing only that the enzymatic reaction in question is or is not proceeding at a normal rate, but not why. From a pathologic standpoint, however, the distinction between a defective and a missing enzyme is trivial, since the result in either case is the same: no catalysis. The outcome is either a buildup in body levels of the substrate or a deficiency in the product of the enzymatic reaction, or a combination of the two.

This picture of enzyme function helps to explain why, of the known hereditary enzyme disorders, very nearly all are recessive conditions, manifesting themselves pathologically only in homozygous individuals. Heterozygotes, with one defective gene, do in fact show a reduction in body enzyme levels to something like 50% of normal, but so efficient are these biocatalysts (a single enzyme molecule can process up to 100,000 molecules of substrate a minute) that even the reduced level is adequate under most conditions to avert pathologic consequences.

To put some flesh on these generalities, let us now consider some specific enzyme disorders. A good one to start with is phenylketonuria (PKU). On the one hand, we know a great deal about it compared with many enzyme diseases; on the other, there is still a lot we would like to know (see Guthrie, Chapter 22, Hsia and Holtzman, Chapter 23, and Howell, Chapter 27.)

The condition was first described in 1934, when the Norwegian clinician A. Følling published a report on 10 patients, several of them siblings who had the characteristic combination of severe mental retardation and excretion of phenylpyruvic acid in the urine. Over the next 20 years, further research (notably by G. A. Jervis) established the following fundamental facts about the disease:

1) It is recessively inherited through a single autosomal gene.

2) It is characterized by the accumulation in the body of large quantities of the amino acid phenylalanine.

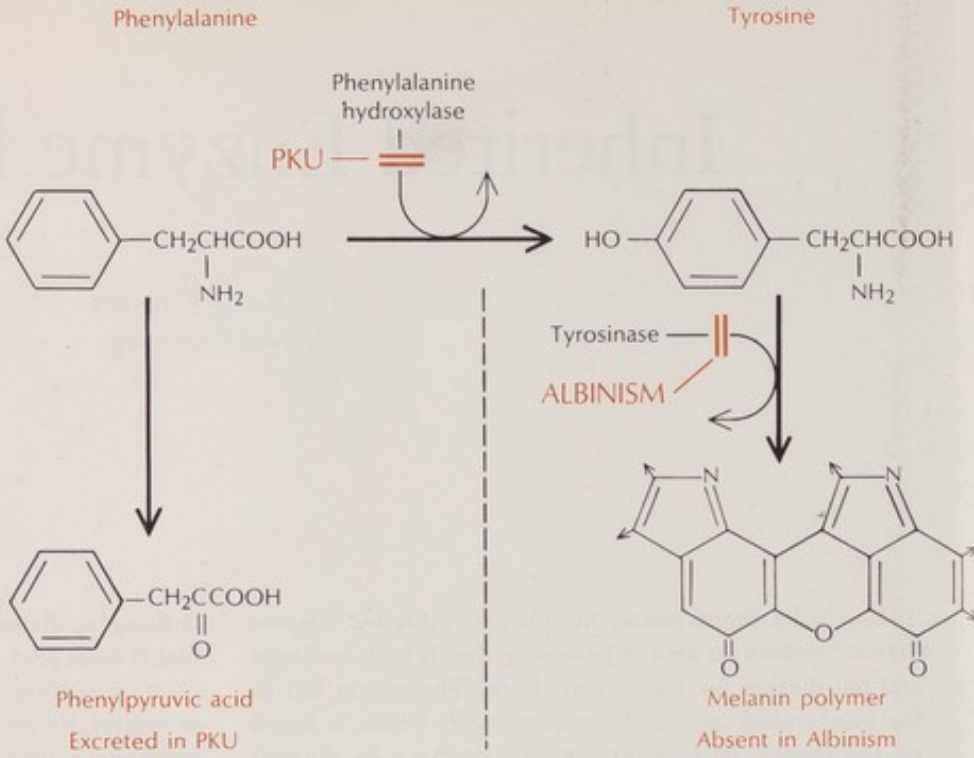
3) The accumulation results from the body's failure to oxidize phenylalanine taken in the diet to another amino acid, tyrosine.

4) The failure, in turn, is caused by inactivity of the liver enzyme phenylalanine hydroxylase, which in normal individuals presides over this oxidative reaction.

5) The major pathologic effect of these metabolic abnormalities is interference with the development of the central nervous system during early childhood. Cytologically, this is manifested only very subtly, perhaps by disordered myelination of nervous tissue; neurologically, it is manifested by the obvious and severe retardation already mentioned.

Given these facts, there was one obvious possibility for rational treatment of the disease: strict limitation of phenylalanine intake, at least during the early years. This approach has in fact yielded good results, especially with the development of simple techniques for early case finding (testing of newborn infants' blood for excess phenylalanine is now mandatory in some hospitals). Early and consistent treatment can prevent or at least markedly ameliorate the irreversible retardation that is the hallmark of PKU. To fully understand the rationale for this therapy as well as to provide background for some of its still problematical aspects, we must now look for a moment at some details of phenylalanine metabolism.

Metabolism of Phenylalanine-Tyrosine System



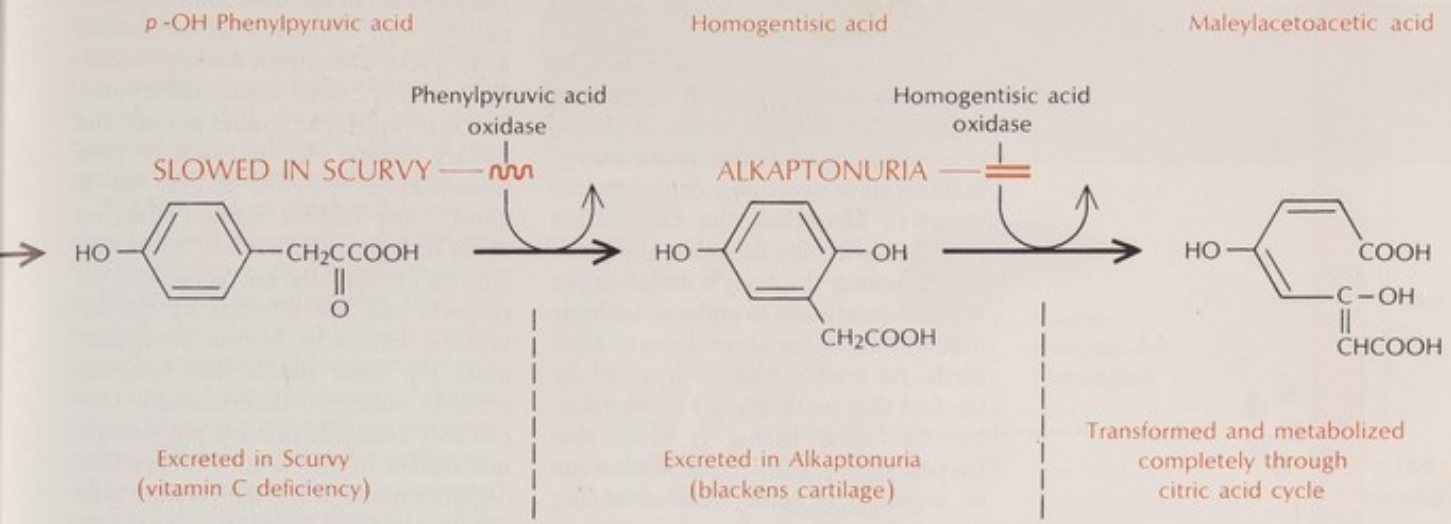
Phenylalanine is one of the "essential" dietary amino acids, i.e., the human body cannot synthesize it. In normal individuals, a modicum of the phenylalanine intake is incorporated into various body proteins, but the bulk of it is oxidized to tyrosine. Since it can be manufactured from phenylalanine, tyrosine is not in the "essential" category but is in any case a normal constituent of dietary protein.

The tyrosine, whether derived from phenylalanine or directly from food, is further processed in the body into the pigment melanin, into the neurohumors epinephrine, norepinephrine, and dopamine, and into the hormone thyroxine. It also, of course, plays a role in protein formation. Only a fraction of total tyrosine is utilized in these ways, however; the remainder is biologically degraded and ultimately enters into the energy-yielding citric acid cycle.

In the patient with PKU, phenylalanine cannot follow its normal oxidative path into these varied reactions involving tyrosine. As phenylalanine accumulates in the blood it is gradually metabolized through various subsidiary processes – notably, that leading to phenylpyruvic acid – but these "leaks" are too slow to hold blood concentrations down to anything like normal levels.

Since phenylalanine is essential for the formation of body protein its complete elimination from the diet would clearly interfere with growth, and this is precisely what happened in some early experiments with diet restriction. Such "overtreatment" even led in some instances to the breakdown of already formed proteins. The system became flooded with phenylalanine, producing severe epileptiform seizures such as are not uncommon in untreated PKU patients. Present practice, therefore, is to lower intake of

with Some Genetically Determined Pathologic Alterations



This simplified and abridged map of the main metabolic pathways of phenylalanine and tyrosine illustrates how enzymatic abnormalities appear to be involved in three genetic diseases (PKU, albinism, and alkaptonuria) and possibly in a vitamin deficiency state (scurvy). A functional deficit in phenylalanine hydroxylase is responsible for PKU (see page 67). Melanin is a metabolite of tyrosine and the pigment's absence in albinism may be a result of a genetically determined "lack" of tyrosinase. The excretion of p-OH phenylpyruvic acid in scurvy patients is believed sequential to accumulation of the substrate due to enzymatic inhibition. In alkaptonuria, the "missing" enzyme is homogentisic acid oxidase; the result is accumulation of the acid. Color bands indicate the enzymatic abnormalities, resultant substrate derangements, subsequent diseases and their manifestations.

the amino acid to the amounts needed for normal growth. The actual criterion is provided by blood assays, which are held to slightly above the normal phenylalanine level (about 2 mg%) to provide a margin for error. The diet also includes extra tyrosine to replace what would normally be provided by phenylalanine oxidation. I know of no conclusive evidence that this is in fact necessary, but the rationale – to keep the body chemistry as near normal as possible in all respects – is certainly sound enough.

Despite all we have learned about PKU, including development of a reasonably effective treatment, we still do not know precisely how its metabolic defects actually damage the body.

As already indicated, there is no very persuasive evidence that the lowering of body tyrosine levels is directly harmful, especially since in the normal as well as the affected individual much of the tyrosine in the diet is simply utilized for its energy content. There is a good deal of evidence, however, that both excess phenylalanine and perhaps some of its abnormal metabolites can derange a number of physiologic processes by inhibiting the action of the enzymes that catalyze these processes. It is quite certain, for example, that PKU lowers the body's production of melanin by interfering with one or more of the metabolic steps by which that substance is manufactured from tyrosine. Untreated PKU patients typically show low melanin levels in skin, eyes, and hair; their coloration, though never reaching frank albinism, is noticeably lighter than what would be expected from comparison with their families.

Similar interference with other paths of tyrosine metabolism may perhaps lower production of epinephrine; rather more certainly, the body's production of another neurohumor, serotonin, is reduced through interference with the metabolism of tryptophan, a compound somewhat like phenylalanine itself. These neurohumoral disorders may help to account for some of the secondary neuropsychiatric symptoms seen in the disease: agitated behavior, muscular hyperactive reflexes, hyperkinesis, tremors, and (as already noted) seizures. An additional or alternative cause of these manifestations may be certain metabolites of phenylalanine itself; in the concentrations characteristic of untreated PKU, these appear to possess pharmacodynamic activity. Phenylethylamine, for example, stimulates the sympathetic nervous system.

But all of these abnormalities, since they are dependent on high phenylalanine levels, should be reversed when phenylalanine intake is reduced, and in fact many are. Most noticeable, perhaps, is a jump in melanin output, shown by darker hair growth when treatment is instituted. The main feature of the disease – CNS damage and its resultant retardation – is not reversible, however. This fact points to the existence of some additional pathologic mechanism.

The most plausible explanation of what this mechanism might be derives from a rather simple fact about PKU patients: despite the relatively enormous quantities of phenylalanine in their blood (up to 30 times normal), *total* serum amino acids are within normal limits. Necessarily, then, the quantities of other amino acids must be subnormal, and indeed assays have shown this to be true, with levels ranging from 85% normal for valine to only 50% normal for arginine.

It has been hypothesized, therefore, that brain damage occurs because the cells of the infant's immature and rapidly growing brain are presented with a highly abnormal pattern of amino acids, with the result that the enormous proportion of phenylalanine crowds out other essential nutrients in whole or in part. The brain thus can-

not grow normally; in particular, myelin is not formed or if formed is defective and breaks down.

What is not wholly clear is why these abnormal growth processes should affect only the brain. It should be stressed that no other tissue abnormalities have ever been demonstrated in PKU. Doubtless the explanation is tied in with the fact that the brain alone, among the body's major organ systems, continues to undergo biologic differentiation for several years after birth. At least a hint is provided by the fact that tooth enamel formed during the same period is often thin (hypoplastic). A fuller elucidation can be expected to come with advancing knowledge of brain growth in general and myelin formation in particular.

The hypothesis, correct or not, is consistent with some other facts about PKU. For one thing, the disease, while unquestionably genetic, is not congenital; brain damage in the untreated patient definitely develops after birth and not before. We know that in *all* infants the phenylalanine hydroxylase system in the liver matures postnatally; only when it fails to mature, as in PKU, does brain damage occur. For another, once brain differentiation is complete, at around age six, the dietary regime can be eased or even discontinued without, so far as is known, any further serious effect on brain structure or function. Such individuals are usually not normal in all respects and are affected by the reversible effects of excess phenylalanine, but their intellectual faculties seem to suffer no deterioration. One can only conclude that the physiologic maturation of the brain renders that organ immune to the irreversible type of phenylalanine damage.

Having summed up the important facts we have learned or suspect about PKU, let us now turn to some of the facts we do not yet understand. Most of these have to do with treatment for the disease and with its therapeutic results. At one time there was considerable dispute over whether dietary

The large excess of phenylalanine in phenylketonuria depresses the levels of most other amino acids, creating a highly abnormal chemical milieu for the cells. Graph shows amino acid levels as percent of normal (analyses by F. Lenneweh and M. Ehrlich).

restriction of phenylalanine was even worthwhile. Further experience, however, has convinced nearly all clinicians that it is. There seems no doubt that treatment can turn children who would have led an essentially vegetable existence into individuals of normal or near normal intelligence.

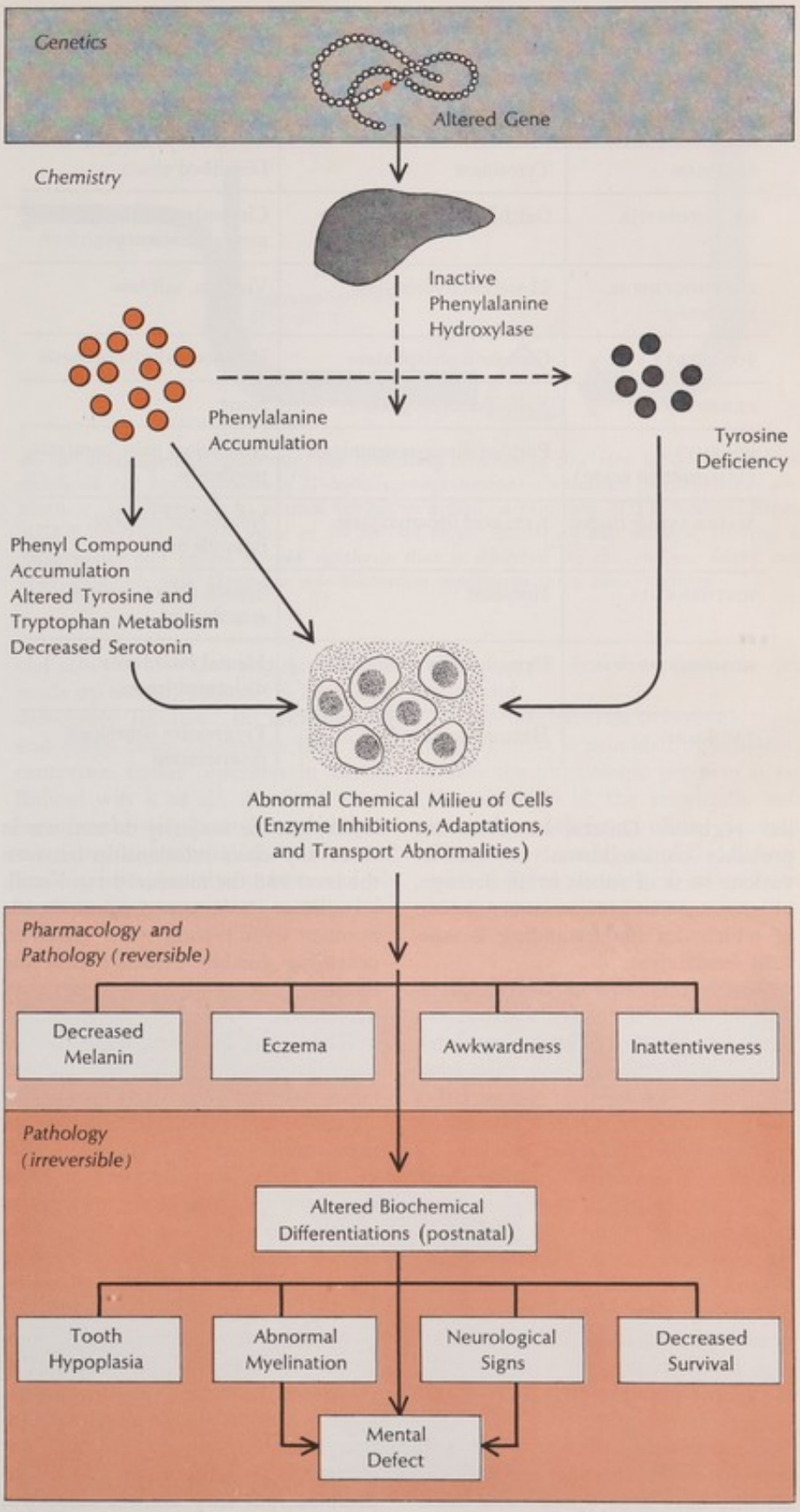
Yet I should stress that this conclusion is based essentially on averages. Treated children can be expected to be noticeably more intelligent than untreated ones, but within both groups there is a considerable range of variation and a degree of overlap between them. Treated PKU patients are sometimes well below normal intelligence, while very occasionally some untreated ones are intellectually normal. And we still have not clearly identified all the factors responsible for the variations.

One obvious variable is the promptness with which treatment has been instituted, and here there seems to be no doubt that the longer treatment is delayed the more irreversible intellectual impairment patients undergo. But the correlation is far from perfect, so that clearly there are other factors. Another obvious possibility is the level and consistency of dietary control as measured by blood phenylalanine levels – but the relationship between serum levels and intellectual impairment has thus far been demonstrated much less clearly than one could wish.

Again, one might think that, other things being equal, treated PKU children of intelligent parents would be more intelligent than those of dull parents – yet again the relationship is far from clear. I must say myself, however, that I would not expect to see such a relationship where brain damage is considerable; it seems unlikely that an imbecile child of bright parents would necessarily be a "bright imbecile."

Even among the children in whom treatment has produced normal intellectual functioning (at least so far as this can be measured by standard IQ tests), not all are "normal." Frequently they show clumsiness, anxiety, sleeplessness, hyperactivity, and sometimes the behavior patterns classified under the rather loose term "autistic." Some of these phenomena are doubtless purely psychologic reactions to family tensions produced by the presence of a seriously ill or endangered child or to the necessity for living on a pecu-

Pedigree of Causation in Phenylketonuria



The defect in phenylketonuria is in the gene that codes for phenylalanine hydroxylase. The direct result is failure of the enzyme's substrate, phenylalanine, to be metabolized to tyrosine. Depicted are the varied results of the accumulation of phenylalanine, the alterations in its metabolic pathways, and the decreased production of tyrosine.

Partial List of Hereditary Enzyme Deficiencies

Disease	*Enzyme*	*Major Consequences*
PHENYLKETONURIA	Phenylalanine hydroxylase	Mental deficiency
ALKAPTONURIA	Homogentisate oxygenase	Arthritis
ALBINISM	Tyrosinase	Disturbed vision
GALACTOSEMIA	Gal-1-P uridyl transferase	Cirrhosis, cataracts, mental deficiency
ADRENOGENITAL SYNDROME	21-steroid hydroxylase	Virilism, salt loss
VON GIERKE'S	Glucose-6-phosphatase	Hepatorenal glycogenosis
PENTOSURIA	Xylitol dehydrogenase	None
PORPHYRIA (intermittent acute)	Porphobilinogen deaminase	Abdominal pain, paralysis, psychosis
MAPLE SYRUP URINE	Keto acid decarboxylase	Neurologic damage, infantile death
HISTIDINEMIA	Histidase	Speech defect, mental retardation
HOMOCYSTINURIA	Cystathionine synthetase	Mental retardation, dislocated lenses
TAY-SACHS	Hexosaminidase A	Progressive neurologic deterioration

liar regimen. Others, however, are probably nonintellectual sequelae of various sorts of subtle brain damage, phenomena that are known to exist but of which our understanding is woefully insufficient.

Most remarkable of all are individuals in whom untreated PKU apparently has produced little or no intellectual impairment. There exist some 30 attested cases (all, of course, either predating the development of effective treatment or discovered too late for treatment) with IQ's ranging from 60 to 107; of these, nearly one third fall within the normal range of 90+. By contrast, the median IQ of untreated PKU patients is around 20. Psychologically, none of these "high-grade" phenylketonurics are normal, but intellectually they are normal enough to pose something of an etiologic problem.

Enzymatically, few of them have been found to be in any way atypical; all who have been tested show essentially no phenylalanine hydroxylase activity. Nonetheless, some may represent a genetic variant and have unexpectedly low plasma phenylalanine levels, but the majority do not, nor is there any clear relationship between the level and the measured IQ. Familial studies have turned up a sizable number with typical PKU relatives, providing further evidence that the mutant gene involved is in no way anomalous.

There are, however, a number of cases in which high-grade PKU patients have high-grade PKU relatives – enough to suggest the possibility, in these cases at least, of some familial genetic factor that in some way moderates the normally destructive effects of the enzyme defect. A less severe type of defect in the same gene may also exist. Or it may be, for example, that some other genetic anomaly has given them other, more effective, mechanisms for disposing of excess phenylalanine. Instances of two genetic wrongs making a right – of one mutation cancelling or ameliorating the effects of another – are anything but common, but they are not unknown; examples exist among the hemoglobinopathies.

The final problematic factor with PKU is when, if ever, to terminate the dietary regimen. In general, at around age six one reaches a point of diminishing returns so far as intellectual function is concerned; by that time, as previously noted, brain development is essentially complete, and the patient's intellectual capacities can be expected to maintain themselves even in the face of high serum phenylalanine levels. From this standpoint, one might well consider sparing the patient and his family the tension, inconvenience, and expense of continuing a diet from which essentially all natural proteins must be eliminated.

There remains the seemingly incontestable fact, however, that pathologic serum levels of phenylalanine possess psychotropic activity, either directly or indirectly, via the compound's abnormal metabolites. This activity and its behavioral manifestations (restlessness, short attention span, and the like) nonetheless shows much variation from patient to patient – again, for unknown reasons. Thus in each individual case one must experiment to balance the gains from normalizing the social environment against the possible losses from "drugging" the patient with "endogenous" phenylalanine or phenylalanine metabolites.

Turning now to some other enzyme diseases, we note that one of them shows a distinct biochemical parallelism to PKU. In alkaptonuria, the problem is again the abnormal build-up of a particular substrate due to the inactivity of the enzyme that would normally process it. Moreover, the substrate itself, homogentisic acid, is chemically akin to phenylalanine and tyrosine, and is in fact a metabolite of the latter. Homogentisic acid is an intermediate compound in the metabolic sequence whereby tyrosine is converted to acetoacetic acid, which in turn is oxidized in the citric acid cycle to yield energy. In alkaptonuria, however, the enzyme, homogentisic acid oxidase, is inactive, so that serum concentrations of the acid are markedly increased.

Much of the acid is eliminated in the urine, and its presence there is diagnostic even in infancy. In some cases diagnosis is actually made at that time, since under certain circumstances the urine is abnormally dark, or turns dark on standing, or diapers are stained. But frequently the condition is detected only much later in life

as the result of more certain indications. One of these is ochronosis, the deposition of a blackish pigment in connective tissues, notably the sclera, ear cartilages, and especially the joints and tendons. This substance is believed to be a polymer resembling melanin, which is itself a polymer of another tyrosine metabolite.

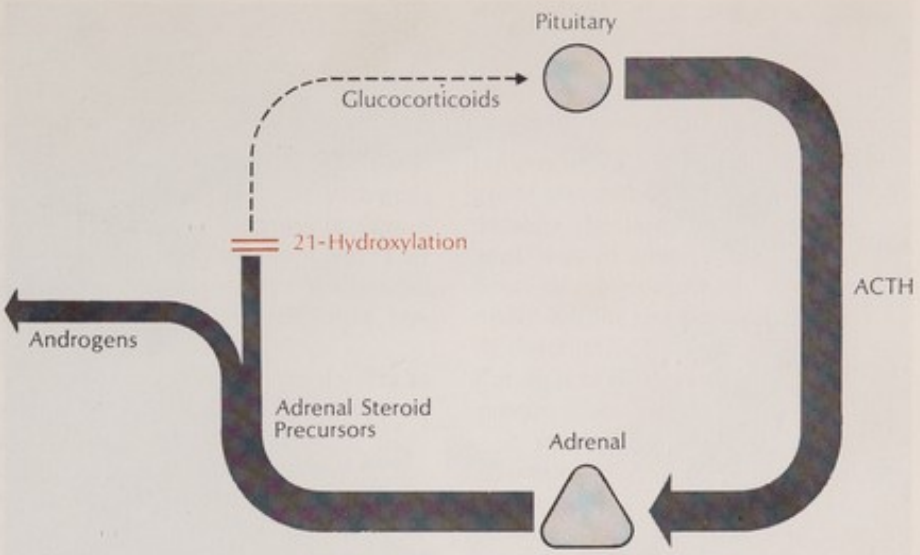
Much more serious than ochronosis per se, which at worst amounts to a cosmetic blemish, is the gradual development of alkaptonuric arthritis. This comes about through damage to the joint tissues, either from the ochronic pigment or the homogentisic acid itself; the exact mechanisms are not known.

Interestingly, alkaptonuria was one of the diseases cited by Garrod in his historic series of lectures on inborn errors of metabolism. As early as 1902, he suggested that the disease was transmitted as a single, recessive Mendelian trait – a hypothesis that has since been thoroughly proved.

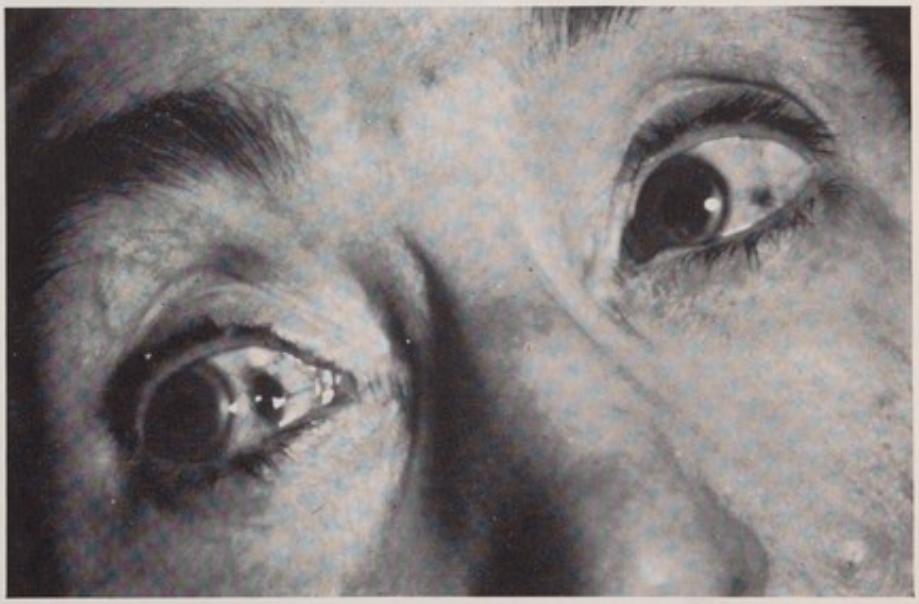
Thus far, no treatment has been developed for alkaptonuria. In theory its manifestations could be prevented by lifelong limitation of tyrosine and its precursor, phenylalanine, from the diet. But the metabolic paths leading from tyrosine are so diverse and so physiologically important that such a step would almost certainly cause more trouble than it prevented. A few physicians have reported good results by simply limiting dietary protein, but their cases are too incomplete to be conclusive. (Supervision for a minimum of 40 years after childhood would be required.) And of course many cases are not diagnosed at all until the arthritis manifests itself, and sometimes not even then.

This fact points up a serious problem involving all the hereditary enzyme diseases, along with many other genetic disorders: their great rarity. PKU, though one of those more frequently seen, has an incidence of something like five per 100,000; alkaptonuria is only about one tenth as common. Thus the average physician may well see no more than one PKU patient in a lifetime and not a single alkaptonuric. The medical researcher is also handicapped; our relatively complete knowledge of PKU reflects chiefly the fact that before treatment was developed its victims were largely concentrated in institutions and special schools, of whose population they made up a small (about 1%) but still detectable fraction. In alkaptonuria and many other conditions this "concentration factor" operates in a very limited way if at all. A certain number of arthritics do tend eventually to wind up in custodial institutions of various sorts, but the incidence of alkaptonurics among them is still probably a good deal less than one per thousand.

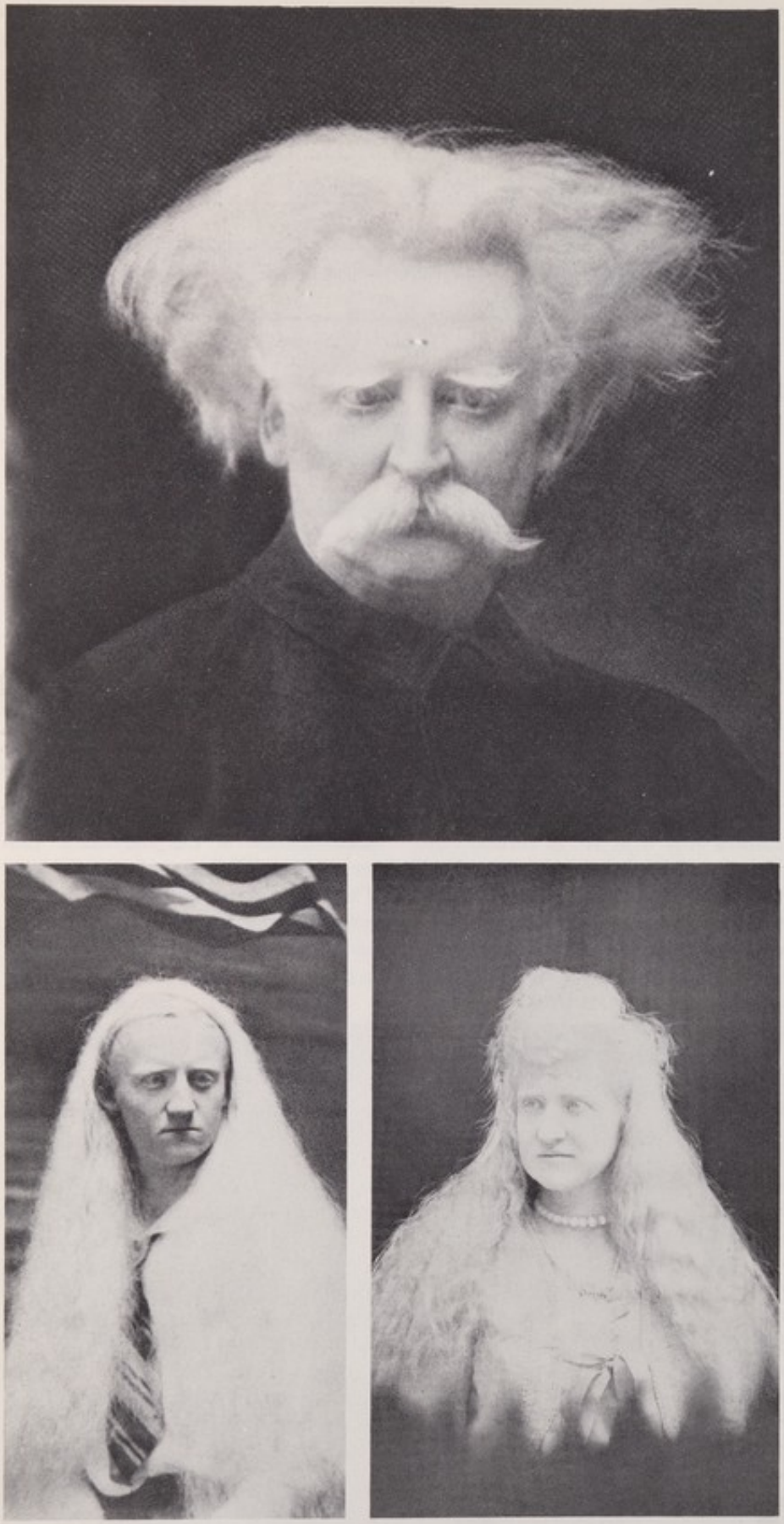
A certain contrast to both PKU and alkaptonuria is provided by albinism, where the physiologic problem is not accumulation of the enzymatic substrate but deficiency of the product. There are several types of this disease, with different modes of inheritance, but the most severe form (complete or



In the adrenogenital syndrome, the hereditary absence of one step in adrenal steroid synthesis – 21-hydroxylation of 17-hydroxyprogesterone – diminishes glucocorticoid production and eliminates its normal inhibitory action on pituitary ACTH release. Excess ACTH therefore exerts more of its normal tropic action on the adrenal, leading to hyperplasia and excess steroid synthesis that is diverted by the enzyme block into androgenic steroids. Cortisone administration suppresses these abnormalities.



Accumulated blackening that occurs in all cartilages and tendons of alkaptonuric patients can be seen in 60-year-old woman at the insertions of the rectus muscles on both sides of the eyeball. Her nose, ears, and tendons also appeared blue through the overlying skin, and she had the typical arthritis in the spine and large joints.



This unusual family of albinos – father (top), son, and daughter – were circus performers in the early part of the century; the apparently normal wife and mother presumably carried the gene for albinism. The family was included in one of the earliest definitive studies of the heredity of human albinism (C. B. Davenport, J. Hered. *7:221, 1916).*

oculocutaneous albinism) is inherited as an autosomal recessive. In patients with this condition the body is incapable of synthesizing the pigment melanin, which, as noted earlier, is a metabolite of tyrosine. The exact mechanism is still uncertain; some possibilities are inactivity of tyrosinase, the enzyme that mediates two of the metabolic steps between tyrosine and melanin, inhibition of melanin formation itself, or a defect in the "active transport" mechanism that carries tyrosine from the serum into the melanophores, where melanin is synthesized. [The active transport genetic disorders, a quasi-enzymatic group of diseases, are discussed in detail in Chapter Eight.]

Whatever the precise mechanism, it does not appear to interfere with other metabolic processes involving tyrosine. That substance shows no tendency to pile up in the bodies of albinos, nor do they show deficiencies in any of its other metabolites (dopa, epinephrine, thyroxine, etc.).

It should be noted in passing that the presence of the phenylalanine-tyrosine metabolites in all three of these diseases does not mean these metabolites are particularly prone to such disturbances. Their prominence is something of an artifact. Garrod pointed out that the aromatic compounds involved here are ones that give rise to colored products in the test tube. Hence they possess a high "scientific visibility." There is no reason to suppose that the body's aromatic compounds are peculiarly prone to metabolic derangements, but good reason why such derangements would be more likely to be noticed early on. With the development of more refined analytical and detection methods, the "aromatic predominance" in genetic disease has greatly lessened.

As a final example of enzyme disorders, let me cite one of what might be called a composite type, i.e., one in which the pathology involves both a pileup of substrate and a deficiency of product. The adrenogenital syndrome is a disease (more accurately, a group of closely related diseases) in which the metabolic flaw deranges the synthesis of adrenocorticosteroids. A number of different defects can occur, each one presumably reflecting a different enzyme defi-

ciency. Though the defects have been identified the enzymes have not.

In the commonest form, the metabolic block limits or prevents synthesis of the 21-hydroxysteroids, one of which is aldosterone. A conspicuous physiologic result is salt loss, sometimes to a life-threatening degree; this may be aggravated by a buildup of the immediate precursors of the 21-hydroxysteroids, such as progesterone, some of which are thought to act as aldosterone antagonists. In another form of the disease, interestingly, a different metabolic defect produces an accumulation of a powerful salt-*retaining* steroid, with resulting arterial hypertension.

Rather more serious than accumulation of the progestins (which at worst merely reinforces an existing pathology) is a more complex metabolic derangement affecting the pituitary. Normally, synthesis of the 21-hydroxysteroids, like that of many other hormones, is regulated by a feedback system involving that gland. High serum levels of the steroids inhibit the pituitary's production of ACTH, which reduction in turn cuts back hormone synthesis by the adrenal cortex. But when 21-hydroxysteroids are *not* being synthesized, ACTH production rises. The result is an increase in those adrenocortical activities that are not metabolically blocked. The effect is to stimulate synthesis of the 21-deoxysteroids, some of which serve as precursors for testosterone and other androgens. It is these which engender the precocious and progressive virilization that is a hallmark of the disease.

Given this analysis, treatment should in principle be fairly simple: replacement of the hormone(s) whose deficiency leads, directly and indirectly, to most of the pathologic sequelae. In practice, however, this approach has given only moderately good results. As a group, the corticosteroids possess extraordinarily diverse properties, and the biosynthetic pathways that connect them (normal or abnormal) are of notorious complexity. Our ability to manipulate their output with precision is therefore still much less adequate than we could wish.

Summing up the enzyme defects as a group, we have seen that we can sometimes prevent or ameliorate their consequences either by reducing intake of the enzymatic substrate (as in PKU) or by supplying the missing enzyme product (as in the adrenogenital syndrome). What are the prospects for improving therapy of the "treatable" defects and for extending it to others?

To begin with, at present there is no chance of curing any of these disorders, since the underlying enzyme problem cannot be corrected. Speculations about "gene surgery," whereby defective genes would be replaced by normal ones, must for the foreseeable future be relegated to the realm of science fiction. What might seem to be a more plausible approach would be to supply the missing or inactive enzyme exogenously, but the practical barriers to this are for the present insurmountable. Merely to obtain the needed enzyme in therapeutically useful quantities would be difficult enough; to get it intact to the proper site of action in the body (in PKU, the liver, for example) would be wellnigh impossible. And even if we assume that these two formidable problems could be overcome, there remains the fact that to the individual suffering from an enzyme defect, the normal enzyme is a foreign protein which is rejected by the body.

A more productive approach, I believe, would focus on the other factors, whether of heredity or environment, influencing the degree to which a given metabolic disorder can actually damage the body. We have already seen that in some PKU patients the brain develops normally or near normally despite the presence of ordinarily destructive levels of phenylalanine. Parallels to this anomaly exist in other genetic diseases.

A particularly striking example is diabetes. We know that there is a strong familial component in this condition, yet we also know that there is enormous variation in the time when this genetic predisposition expresses itself in frank diabetes. Some individuals develop the disease in childhood or adolescence, some in middle life, and some – even with the most suspicious sort of heredity – not at all. If we knew why, we might be able to defer development of the disease indefinitely in all potential diabetics.

Progress in this direction, or in other aspects of genetic disease, will require much research. Moreover, the rarity of most enzyme or presumed enzyme defects (diabetes, of course, being an outstanding exception), will necessitate the closest collaboration and pooling of information among different investigators. Finally, the problems of genetic disease, perhaps more than any other area of pathology, require an interdisciplinary approach; their elucidation must draw on the skills of the geneticist, biochemist, and nutritionist no less than that of the clinician. Only with this sort of many-sided approach can we hope to bring these obscure, complex, and fascinating groups of ailments under control.

7

Vitamin-Dependent Genetic Disease

LEON E. ROSENBERG
Yale University

For the practicing physician, extensive reading in the area of genetic disease can become a rather depressing business, since it repeatedly reminds him of the limitations of his art. Whether one considers the visible abnormalities of the chromosomes or the invisible mutations that engender metabolic errors, very few of the resulting disorders are treatable at all, and even in the few that are, treatment can seldom do more than ameliorate the damage. From this standpoint, the vitamin-dependent genetic disorders constitute a rather cheering contrast. In principle, and by definition, all of them are wholly correctable, and principle is increasingly being translated into therapeutic practice. Quite apart from this, however, they are worth reading about because, as with so many genetic abnormalities, study of the abnormal has advanced our knowledge of the normal – in this case, the metabolic and biochemical functioning of vitamins.

Discovery of the vitamin-dependent disorders dates from 1954, when A. D. Hunt and his associates reported on the case of two infants, siblings, who suffered from violent convulsive seizures that responded only to large amounts of vitamin B_6 (5 to 25 mg/day – the normal infant requirement is 0.5 mg or less). This condition has since been documented in nearly 40 other cases, and in all pedigree studies have shown the anomaly was inherited.

In the subsequent 17 years, the list of vitamin-dependent derangements has expanded to more than a dozen, involving six different vitamins (B_6, for reasons I shall mention later, is involved in no fewer than six). All of them are alike in being inherited, in involving one or another specific biochemical abnormality – to be further characterized below – and in responding only to pharmacologic doses of the vitamin in question (doses ranging from 10 to 1,000 times the physiologic requirement). Since the first condition (inheritance) cannot always be demonstrated unambiguously, the other two are of particular importance in distinguishing the vitamin *dependencies* from the acquired vitamin *deficiencies*. The latter typically involve not one but several biochemical abnormalities, and respond to physiologic rather than pharmacologic doses.

Current knowledge of the inherited vitamin dependencies is summed up in the table on page 76. In connection with this, it is worth emphasizing that though in all cases the *biochemical* abnormality is correctable, *clinical* improvement does not necessarily ensue. Thus in homocystinuria, for example, the high serum and urinary levels of homocystine can be normalized with large doses of vitamin B_6, but the ectopia lentis, osteoporosis, and skeletal and CNS abnormalities characteristic of the disease remain.

I should emphasize, however, that the youngest homocystinuric treated thus far had already reached the age of 15 at the time of treatment. It is quite possible that we may have here a parallel with phenylketonuria, in which the biochemical abnormality must be treated within the first few weeks of life if irreversible CNS damage is to be prevented. The hope that similar prompt action could prevent or minimize the irreversible effects of homocystinuria is strengthened by clinical results in another vitamin-dependent condition, methylmalonic aciduria.

The very limited experience with homocystinuria and other vitamin-dependent anomalies limits our ability to answer many questions about both their genetic and their epidemiologic aspects. In general, and with two notable and interesting exceptions, all of them seem to be inherited as autosomal recessives, i.e., an affected individual must be homozygous, inheriting the defective gene from both parents. The exceptions are familial hypophosphatemic rickets and B_6-responsive anemia, which are transmitted as X-linked traits. A woman affected by one of these will be heterozygous, with one normal and one mutant X chromosome, and will therefore transmit the disease to half her sons and half her daughters. An affected man, by contrast, will transmit it with his X chromosome to all his daughters but never to a son, who of course receives

only the Y chromosome from his father.

Here is perhaps the place to mention another curiosity among the vitamin-dependent anomalies: cystathioninuria. The first reports of this condition associated it with severe clinical signs such as acromegaly. As further reports came in, however, it began to appear that the association must have been fortuitous. At present, we know of no clinical syndrome consistently associated with the urinary abnormality, and, in fact, many cystathioninurics show no clinical signs whatever.

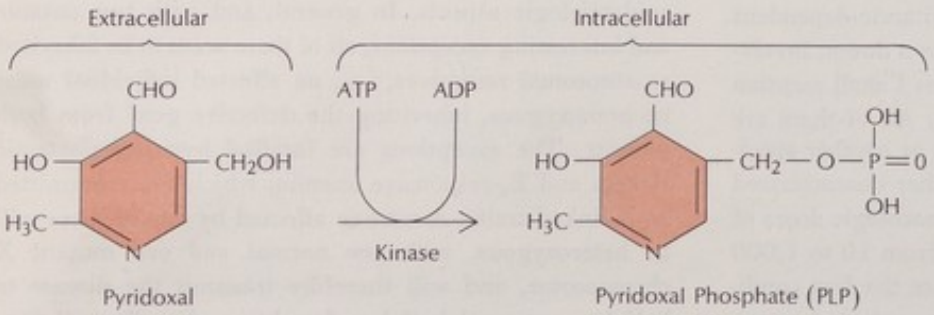
Individually and even collectively, the vitamin-dependent conditions seem to be quite rare, though at present the case material is too limited to permit more than the roughest guess at their incidence. It is conceivable, however, that, as with some other metabolic errors, our improved knowledge of what to look for may eventually reveal that some of them are more common than it now appears.

Much of the research on vitamin-dependent anomalies has concentrated on their mechanisms. Fundamentally, like all other inborn metabolic errors, they are thought to result from defects in protein synthesis, i.e., in the manufacture of enzymes or enzyme-like substances. But because of the peculiar and often complex biochemical interactions in which vitamins participate, identification of the actual defect is often difficult.

Many and perhaps most vitamins, including all those implicated in the anomalies under discussion here, serve as "coenzymes": substances whose presence is required if a particular enzyme is to perform its catalytic function in facilitating metabolic processes. Here is the first source of complexity. Enzymes proper are quite specific in their functions: Most of them are thought to catalyze one and only one reaction. Vitamins, by contrast, are seldom if ever so limited. Vitamin B_{12}, for example, has been shown to be involved in at least five different reactions with five different enzymes, though three of these have been demonstrated only in microorganisms. Vitamin B_6 may be involved in as many as 50.

The second complexity arises out of the fact that most and perhaps all vitamins that act as cocatalysts do so only after being chemically modified from the form in which they enter the body. Thus vitamin B_6 is taken into the body in the form of pyridoxal, pyridoxine, and pyridoxamine, all of which have been isolated from foodstuffs and from extracellular fluids. In the cell, however, it must be transformed into pyridoxal phosphate (PLP). In this process it picks up the phosphate group from adenosine triphosphate (ATP) with the help of a specific enzyme quite different from the enzymes with which the PLP subsequently "cooperates."

Having gotten into the cell and there been transformed into its "active" form, the vitamin must still join with the proteinaceous part of "its" enzyme (called the apoenzyme) to form the metabolically efficient holoenzyme. Exactly what happens here is poorly understood. What seems certain is that the coenzyme becomes bound to the apoenzyme, thereby modifying its structure. It may be that the coenzyme molecule forms part of the "active site" that reacts with the enzyme substrate; alternatively, the binding of the coenzyme may alter the shape of the apoenzyme's protein chain, forming the active site elsewhere. A third possibility has emerged from work on the enzyme tryptophan synthetase in *Escherichia coli*, in which PLP serves as a coenzyme. The enzyme proper is known to be composed of two polypeptide chains while the PLP, it is thought, may serve as a sort of "glue," binding the two chains together.

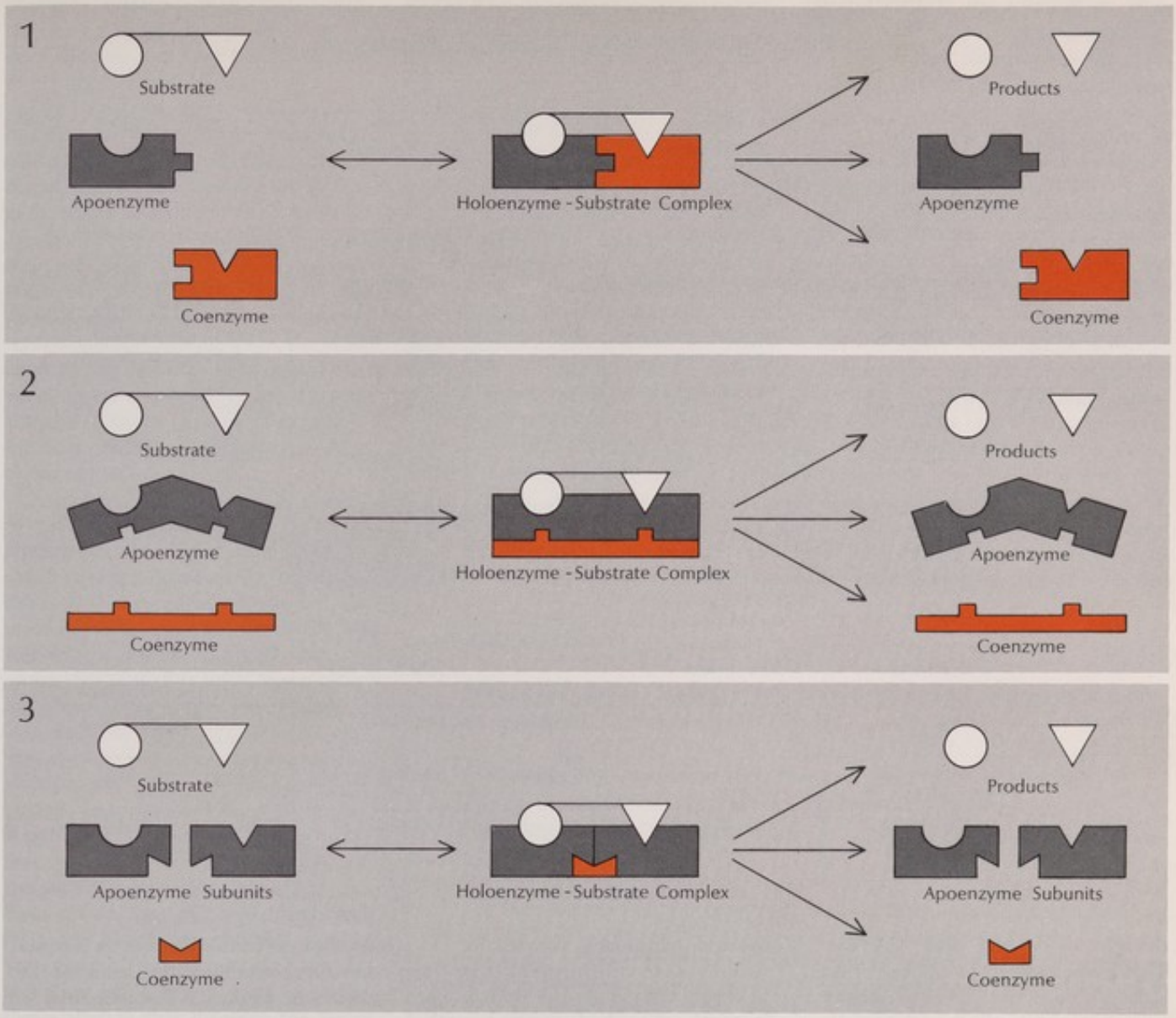


Ingested form of vitamin B_6 (pyridoxal, shown, pyridoxine, or pyridoxamine) is metabolically inert until it is converted to its coenzyme form, PLP, within the cell.

From the preceding discussion, it will be clear there is a wide variety of ways in which a hereditary malfunction in protein synthesis could hinder or block the activity of the vitamin as coenzyme. First, the malfunction might involve the protein responsible for the active transport of the vitamin into the cell. If the vitamin does not get into the cell at all, or hardly at all, one would expect a breakdown of *all* the reactions involving that vitamin, so that the symptoms would approximate the broader spectrum of vitamin deficiency rather than those of vitamin dependency (though the condition might still respond to high vitamin dosages).

Rather less speculative is a defect in the enzyme that transforms the vitamin from its dietary to its active form, e.g., from pyridoxine to PLP. This particular defect, indeed, is almost certainly *not* a case in point, for the reason just mentioned above: A failure to form PLP would derange many enzymatic reactions rather than the single ones that are characteristic of the B_6-resistant conditions. A more likely candidate is vitamin-D-resistant rickets. This vitamin – more accurately, its modified form – is required to activate the protein that transports calcium from the intestine into the blood. Since patients with the disease show no reduction in vitamin D levels either in the serum or intracellularly, it is possible that the conversion mechanism is at fault. Conversely, the defect could involve the apoenzyme rather than the coenzyme, i.e., because of a mutant apoenzyme structure, the coenzyme does not bind to it readily, necessitating high concentrations of vitamin D if the holoenzyme is to be formed in physiologically adequate quantities.

Our data on vitamin-D-resistant rickets is not yet adequate to differentiate between these two possibilities. In the case of another vitamin-dependent condition, methylmalonic aciduria, a number of recent studies, in our laboratory and elsewhere, have provided a rather more certain picture of



Having been converted into its coenzyme form within the cell, a vitamin may act metabolically by the mechanisms diagrammed above: 1) It may provide a portion of a substrate combining site; 2) it may alter the configuration of "its" apoenzyme and thereby facilitate attachment of the substrate; or 3) it may unite apoenzyme subunits to permit substrate attachment.

the mechanism involved. For this reason, and because the disease is of considerable intrinsic interest, I should like to discuss it at some length.

Beginning in 1967, clinical reports began appearing of children with severe metabolic ketoacidosis and developmental retardation. A clue to the metabolic defect came from the finding of high methylmalonic acid (MMA) levels in the urine – up to 5 gm/24 hr. This suggested that vitamin B_{12} might be involved, since the same urinary abnormality (though to a much less marked degree) had been reported in cases of pernicious anemia. However, none of the children showed low serum levels of B_{12}, nor did they have megaloblastic anemia or any other signs of B_{12} deficiency.

The following year, an infant suffering from this condition was admitted to our hospital. At eight months, he was seriously underweight and had an IQ estimated at 50. Seeking to pinpoint the metabolic disorder, we performed a number of tests. Both the ketoacidosis and the excretion of MMA were aggravated by administration of L-valine, which is a specific metabolic precursor of methylmalonate. This suggested that the metabolic block involved the conversion of MMA into succinic acid, in which form it can pass into the citric acid cycle for normal oxidation. The conversion is known to be important in regulating the body's acid-base balance, which would explain how its blockage could produce acidosis; in addition, the accumulation of MMA might itself produce toxic effects.

Confirmation of our hypothesis came from in vitro tracer studies of the child's leukocytes: They could not oxidize either methylmalonate or its immediate precursor, propionate, to carbon dioxide, yet they oxidized succinate normally. It occurred to us to try the effect of large B_{12} doses. We found that the child responded to parenteral administration of 1,000 μg/day (the normal requirement is

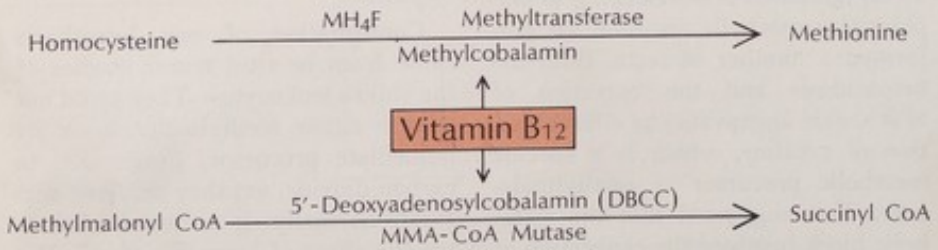
Vitamin-Dependent Inborn Errors of Metabolism

Vitamin Required	*Disorder*	*Clinical Manifestations*	*Vitamin Dose*		*Biochemical Basis*
			Normal Requirement	*Patient Requirement*	
Thiamin (B_1)	Thiamin-responsive anemia	Megaloblastic anemia		20 mg/day	Unknown
	Branched chain ketoaciduria	Slow development	1 mg/day	10 mg/day	Defective branched chain ketoacid decarboxylase
	Pyruvic acidemia	Intermittent cerebellar ataxia		100 mg/day	Defective pyruvate carboxylase
Nicotinamide	Hartnup disease	Intermittent cerebellar ataxia; mental retardation	5-10 mg/day	40-200 mg/day	Defective intestinal absorption of tryptophan
Pyridoxine (B_6)	Infantile convulsions	Clonic and tonic seizures		10-25 mg/day	Unknown
	X-linked anemia	Microcytic, hypochronic anemia		> 10 mg/day	Unknown
	Cystathioninuria	Probably none	1-2 mg/day	200-400 mg/day	Defective cystathionase
	Xanthurenic aciduria	Mental retardation		5-10 mg/day	Defective kynureninase
	Homocystinuria	Ectopia lentis; arterial and venous thromboses; mental retardation		25-500 mg/day	Defective cystathionine synthetase
	Hyperoxaluria	Calcium oxalate renal calculi		75-400 mg/day	Defective α-ketoglutarate: glyoxylate carboligase
Biotin	Propionic acidemia	Metabolic ketoacidosis	10 μg/day	10,000 μg/day	Defective propionyl-CoA carboxylase
Cobalamin (B_{12})	Methylmalonic aciduria	Metabolic ketoacidosis	1 μg/day	200-1000 μg/day	Defective synthesis of B_{12} coenzyme
Calciferol (D)	Familial hypophosphatemic rickets	Rickets; short stature	400 units/day	50,000-200,000 units/day	Defective renal reabsorption of phosphate

thought to be less than 1μg/day). The ketoacidosis was abolished and, indeed, could not be re-precipitated even by oral loading with L-valine. MMA excretion, which had ranged from 800 to 1,200 mg/day, dropped to 200 to 300 mg/day. The ability of the leukocytes to oxidize propionate increased measurably.

We placed the infant on a low-protein diet to reduce valine intake and minimize the need for medication, though additional courses of B_{12} therapy were still required from time to time. After one year on this regimen, the boy had reached normal weight for his age and his IQ had risen to 100. (It is worth noting that similar encouraging results have been reported in two of the B_6-dependent conditions: infantile convulsions and pyridoxine-responsive anemia.)

The outcome in this case, in addition to being therapeutically gratifying, confirmed our belief that B_{12} was implicated in the disease. It did not, however, establish the locus of the metabolic fault. To discover this, we had to give further consideration to the biochemistry of B_{12}.



In man B_{12} is needed as coenzyme in the two reactions shown. Synthesis of methionine requires methyltransferase, methyltetrahydrofolate (MH_4F), and methylcobalamin. Isomerization of methylmalonate to succinate is specifically catalyzed by mutase and DBCC.

The precise form in which B_{12} normally enters the body is unknown. Cyanocobalamin, the commonest commercial form of the vitamin, is almost certainly an artifact of the procedures used to manufacture it; it has never been isolated from animal tissues. Several investigators have established, however, that the cyanide (CN^-) group is but one of several alternative radicals that can attach themselves to the cobalt atom, and that it is the identity of these radicals that determines, very specifically, the vitamin's ability to act as a coenzyme in a given reaction. Thus (to consider the two reactions known to occur in mammalian systems) the remethylation of homocysteine to methionine – a reaction whose biochemical significance is still

obscure – requires methylcobalamin; the isomerization of methylmalonate to succinate, on the other hand, requires 5′-deoxyadenosylcobalamin (DBCC).

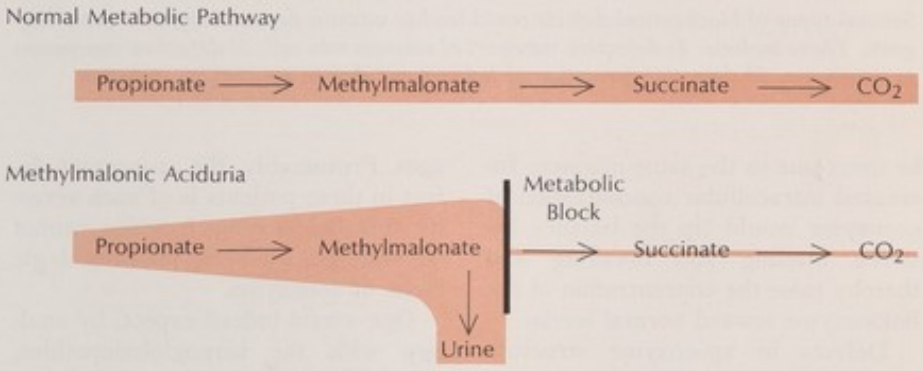
We have been able to show that the biochemical defect in methylmalonic aciduria manifests itself in cultured fibroblasts from affected individuals, and refined analysis of these cells indicates that the concentration of DBCC is less than 10% of normal. High B_{12} concentrations in the growth medium, however, move the level of enzyme activity (i.e., the cells' capacity to oxidize propionate) toward normal, indicating that the apoenzyme is quite capable of functioning if it receives a sufficient amount of coenzyme.

Even this, however, does not amount to conclusive proof of the defective mechanism. In theory this could be obtained by administering DBCC to the cell. If it could be shown that abnormal cells are "normalized" by physiologic levels of DBCC but require 100 to 1,000 times the physiologic level of the vitamin itself (which is in fact the case), we could be certain that it is the conversion of the vitamin to DBCC that has gone wrong. Unfortunately, this particular experiment is impossible, owing to the fact that DBCC will not cross the cell membrane. Recently, however, we have shown that cultured fibroblasts from one patient were unable to convert vitamin B_{12} to DBCC. This observation lends strong support to the thesis that in this disorder the basic defect involves the enzyme that converts the vitamin to its specific coenzyme form.

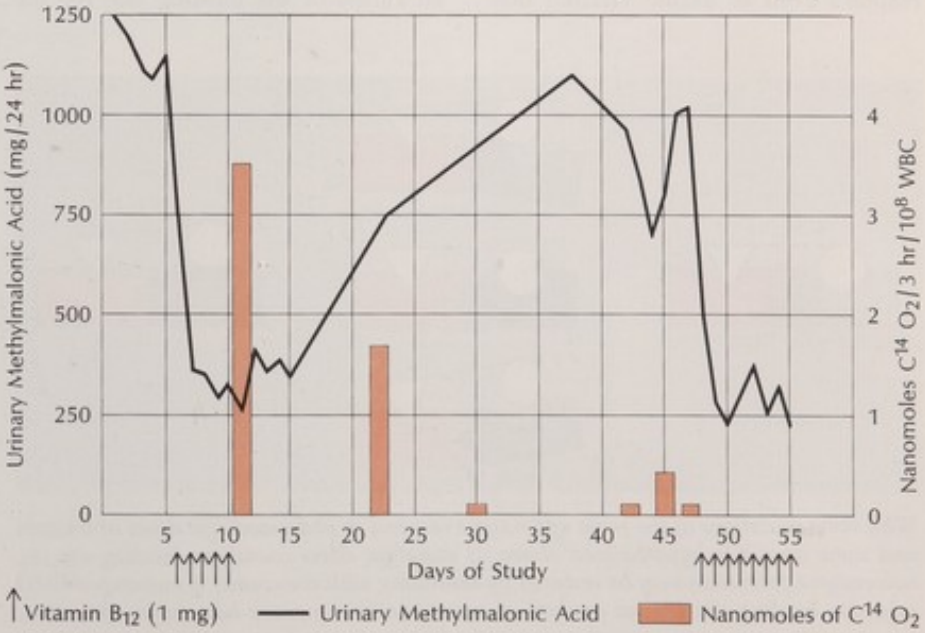
When we turn to the third possible area of metabolic malfunction – reduction of the apoenzyme's capacity to bind its coenzyme – the obvious cases in point are the B_6-dependent disorders. All six of these conditions are clinically and chemically distinct, and in four of them the condition has been traced to the malfunction of a specific enzyme (in the other two, the identity of the enzyme, while suspected, has not yet been proven). In all of these conditions, moreover, B_6 participates in the same active form: pyridoxal phosphate. It would seem a safe inference, then, that in none of these conditions can the PLP-forming system be at fault, since the resulting disorder would reflect malfunction of all the PLP-dependent enzymes rather than a single one. The presumed alternative is that in each case the specific apoenzyme is defective, i.e., its molecule binds the PLP molecule very slowly or very loosely.

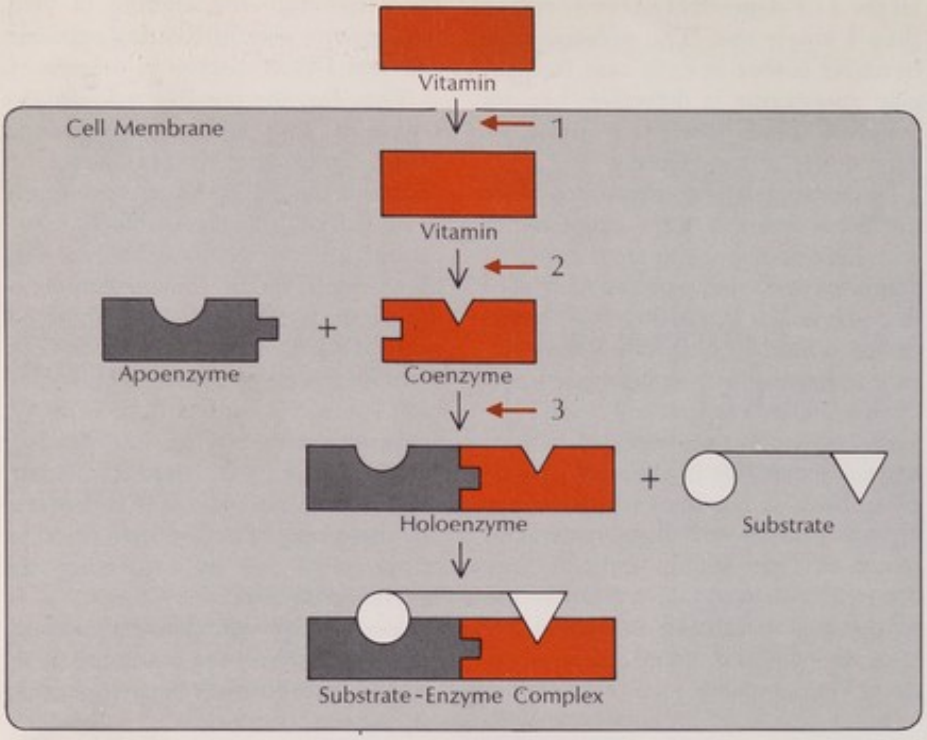
Some support for this hypothesis has come from in vitro experiments with liver homogenates from cystathioninurics and from patients with xanthurenic aciduria, another B_6-dependent condition. In both cases, the relevant enzymes (cystathionase and kynureninase, respectively) are only feebly active, but are restored to near normal activity by in vitro addition of PLP. Such is not the case, however, in some patients with homocystinuria, where in vitro studies indicate that the relevant enzyme, cystathionine synthetase, cannot be activated by even the highest concentrations of PLP. Yet these patients do demonstrably respond to the vitamin in vivo, raising some very interesting possibilities that I shall discuss in a moment.

Thus far, the possible mechanisms I have set forth are all explicable in terms of the law of mass action. A defect in the transport of the vitamin to the cell could be overcome by "saturating" the transport mechanism with high extracellular concentrations of the vitamin (though, as will be recalled, this mechanism seems unlikely in most cases on grounds of insufficient specificity – it would be expected to derange several vitamin-dependent reactions, not one). Similarly, a partial defect in the enzyme that converts the vitamin to its active form could be compensated for by saturating the enzyme with vitamin. Finally, a reduction in the apoenzyme's capacity to bind its coenzyme, assuming it retained some binding capacity, could



If normal major pathway of propionate catabolism in man is blocked by defects in synthesis of mutase apoenzyme or B_{12} coenzyme, methylmalonate accumulates and is excreted in the urine (above). When large doses of B_{12} were given to an infant with B_{12}-dependent methylmalonic aciduria, urinary methylmalonate fell and propionate-C_{14} oxidation by the child's leukocytes increased markedly, as shown in graph below.





Several types of biochemical defects could lead to vitamin dependency, schematic suggests. These include: 1) defective transport of vitamin into cell; 2) defective conversion to coenzyme; 3) defective formation of holoenzyme due to apoenzyme mutation.

be overcome in the same manner: Increased intracellular concentration of coenzyme would tip the balance between binding and breaking and thereby raise the concentration of the holoenzyme toward normal levels.

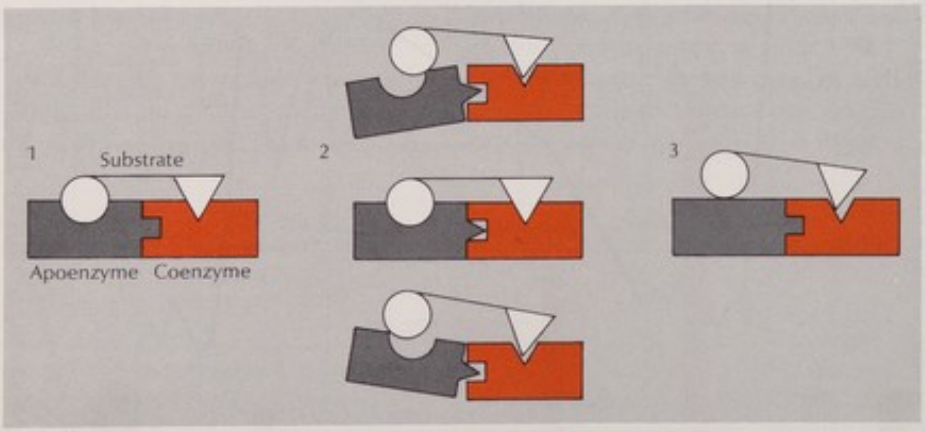
Defects in apoenzyme structure would also account for some reported cases that, though clinically indistinguishable from certain vitamin-dependent conditions (e.g., homocystinuria, methylmalonic aciduria), do not respond even to heroic vitamin dosages. Presumably, the apoenzyme defect in these patients is of such severity that holoenzyme function cannot be restored by even pharmacologic doses of coenzyme.

One would indeed expect, by analogy with the hemoglobinopathies, eventually to discover a whole spectrum of enzyme defects, whose character and severity would depend on the site and nature of the mutant alteration to their protein chains: An alteration of the binding site for the coenzyme could either reduce the binding capacity (vitamin-dependent) or make binding totally impossible (vitamin-resistant). An alteration of the active site could render the enzyme partly or totally inactive but unresponsive to additional vitamin in either case.

We are still left, however, with the puzzling problem of homocystinuria cases that respond to B_6 in vivo although the enzyme in their tissues cannot be activated in vitro. The most plausible explanation, though still only a hypothesis, is that dosage with the vitamin somehow activates an alternate biochemical pathway for "processing" homocystine. We know that such alternatives exist in relation to the sulfur-containing amino acids generally, and given the great versatility of B_6 (or PLP) as a coenzyme, it is not unlikely that enough of it could set one or another of them going. This hypothesis opens up two important considerations.

On the one hand, it should infuse an extra measure of caution into the experimental treatment of homocystinuria. It is one thing to activate a normal enzymatic process that for one reason or another is failing to function (as in methylmalonic aciduria). It is something else again if, the normal enzymatic system being irretrievably damaged, we undertake to activate an alternate and perhaps normally nonfunctional process. While doing so might correct the immediate biochemical abnormality, it might also generate abnormalities on its own.

On the other hand, the "alternate pathway" hypothesis opens up some very exciting, if still very tenuous, therapeutic prospects. If we can correct the biochemical defect in homocystinuria by pharmacologic means (as we apparently can) and do so safely (still to be determined) despite the apparent total inactivity of the enzyme, it is surely conceivable that administration of other drugs – not necessarily vitamins – could correct other inborn metabolic errors that are presently untreatable. Thus our findings in the quite narrowly specialized field of vitamin-dependent anomalies may yet turn out to possess important implications in much broader areas of genetic disease – a prospect that would naturally be enormously gratifying to researchers in the field.



Why some mutations of the same apoenzyme respond to pharmacologic doses of vitamin and some do not is hypothesized above. If mutation alters coenzyme binding site (2), holoenzyme formation may be restored by saturating with coenzyme; if mutation deletes substrate binding site (3), no amount of coenzyme can improve holoenzyme function.

8

Genetic Defects in Membrane Transport Mechanisms

CHARLES R. SCRIVER
McGill University

In all forms of genetic disease, as in most areas of disease generally, medicine maintains a twofold interest. On the one hand, it investigates the etiology and biochemistry of these inherited conditions with a view to treating and/or preventing them; on the other, it examines the same phenomena for clues to the nature of basic physiologic processes, employing the abnormal ("nature's experiments") to elucidate the normal.

As regards the very specialized group of metabolic anomalies called the inborn errors of membrane transport, research has almost inevitably been dominated by this second area of interest. As a group, these conditions are of usually minor clinical importance; all of them are quite rare, and though a few produce severe and lasting damage, others seem fairly easily controllable and some wholly benign. What makes them worth writing about is the light they have thrown on the complex nature of membrane transport and, through it, on the nature of a key aspect of cellular biology, the cell membrane.

To understand the significance of membrane transport in its various forms, both normal and abnormal, we must first say something about the cell membrane itself. The distinguished Soviet biologist A. I. Oparin, in discussing the origin of life on earth, has noted that a key step in that process must have been the physical separation of organisms (or proto-organisms) from their environment, since without such separation it is hardly possible to conceive of even the simplest organism as existing in any stable way. Whatever may have been the original means by which this separation was attained (Oparin himself has speculated extensively about it), present-day organisms characteristically do so, at least in part, by means of the exterior (plasma) cell membrane. This is the case as regards both unicellular organisms and the individual cells of multicelled creatures such as man (though for such cells "the environment" is in most cases not "the outside world" but the body's other cells or its extracellular fluids).

The plasma membrane must necessarily perform a dual function. On the one hand, it must separate the interior constituents of the cell from the exterior environment, thereby providing a locus where those constituents, through their concentration and their proximity, can interact with one another to maintain the processes of life. On the other hand, it cannot set up a total separation, since cells (with a few exceptions, such as bacterial spores and the like) must participate in a constant exchange of chemicals with the environment, taking in oxygen and nutrients and giving off wastes and other products of cellular metabolism. It is here that the membrane transport mechanisms enter the picture.

Parenthetically, it should be noted that consideration of the plasma membrane by no means exhausts the subject of cell membranes. The interior membranes of the cell – nuclear, lysosomal, mitochondrial, and so on – are no less essential to its function than the plasma membrane, though their properties are as yet less well known. Many findings concerning the plasma membrane – specifically including those of its transport mechanisms – appear to be at least partially applicable to other cell membranes.

The notion of a membrane that both separates and does not separate is, of course, not new to biology. In the form of the "semipermeable" membrane, it has been known for at least a century, while the osmotic processes by which the solvent diffuses across such membranes have been extensively studied and measured. Passive diffusion of the solute requires a chemical gradient, with compounds diffusing from regions of high concentration to those of low concentration. However, experiments as early as the turn of the century established that such mechanisms simply could not explain the actual behavior of transport into living cells. Any number of key compounds – sugars, amino acids, sodium and potassium ions, and so on – could be shown to enter cells at rates far faster than could be accounted for by simple diffusion; in most cases, indeed, the movement was actually against a chemical gradient, a process analogous to water flowing uphill.

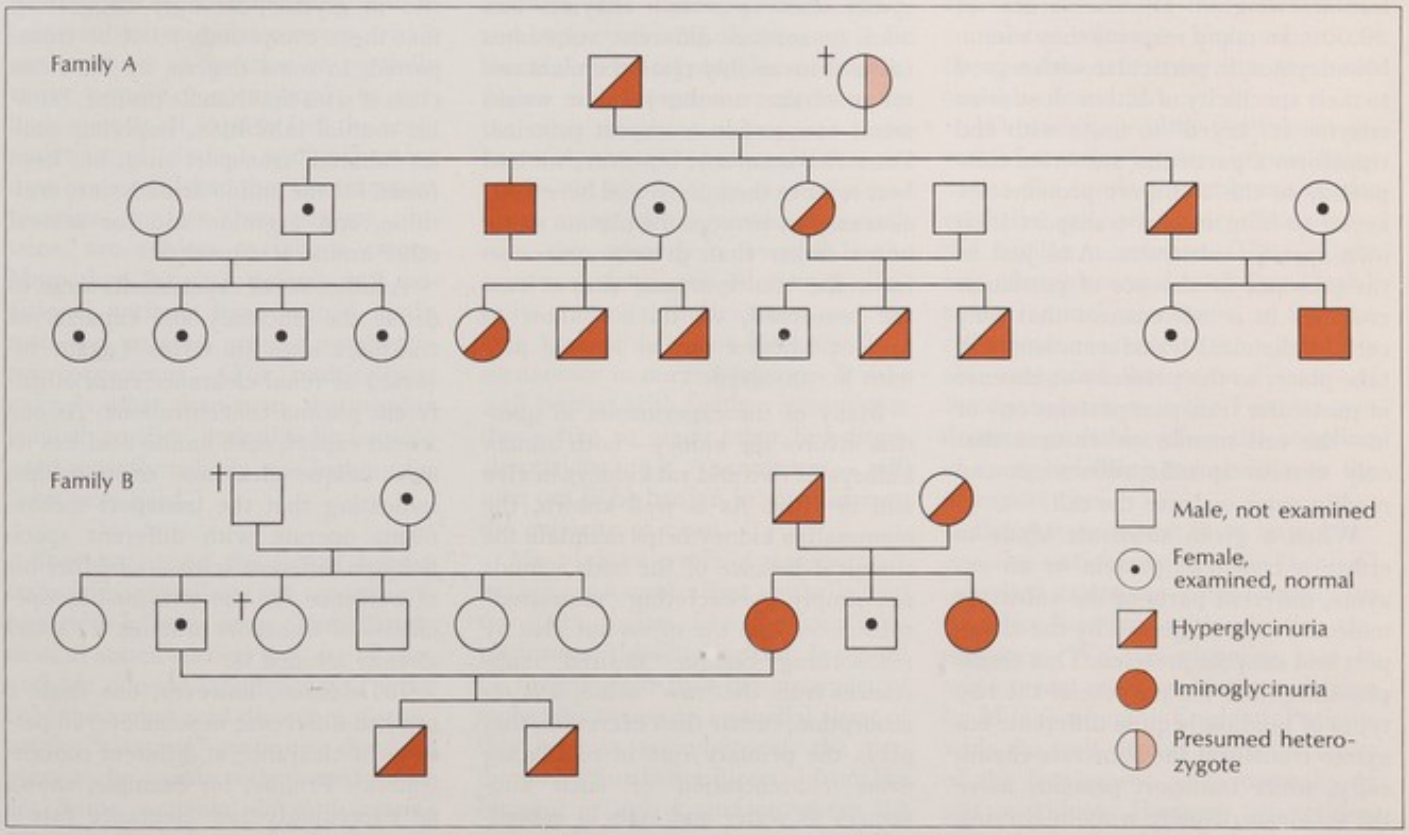
To explain these phenomena it was necessary to postu-

late a "conjugate driving force" (a term used by Schultz) that could "pump" molecules or ions into (or, in some cases, out of) the cell against a gradient. Many years of experimentation have established that there are in fact at least two such mechanisms, each of which involves a carrier or site that first binds the solute. One, called exchange diffusion, is a sort of revolving-door effect whereby the movement of a substance in one direction helps to power the movement of another substance in the opposite direction; this occurs, for example, in certain kinds of ion transport through the membranes of nerve cells. In most cases, however, the transport is driven by a metabolic process of the cell itself which is somehow coupled (conjugated) to the transport site; this mechanism is active transport.

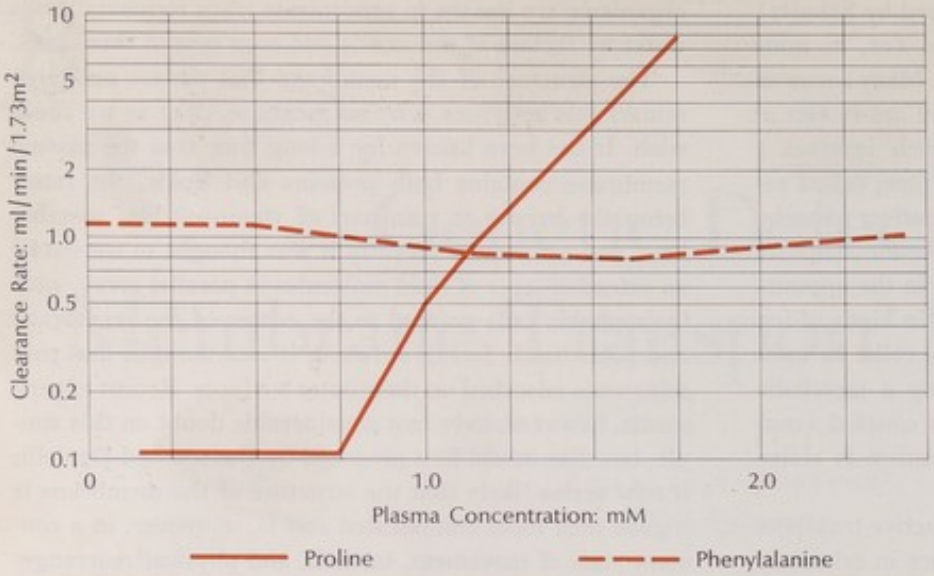
In summary, the metabolic function of active transport is twofold: first, to take substances that are in relatively low concentrations outside the cell and put them into it in high enough concentrations for the cell's chemical processes to operate efficiently; second, to remove certain products of cellular metabolism from the cell rapidly enough so that they do not "back up" and bring the cell's chemical activities to a halt. Mechanisms of this sort possess obvious evolutionary advantages, which doubtless explains why they are found in cells from microorganisms to man. By means of active transport, for example, certain marine organisms are known to concentrate trace elements in sea water by factors of thousands and even tens of thousands.

The structure of the membrane that carries out these remarkable activities is by no means as clear as we could wish. It has been known for a long time that the plasma membrane contains both proteins and lipids, the latter being the barrier to transport of water-soluble metabolites, and until quite recently it was thought to consist of an extended layer of lipid molecules in parallel array, with hydrophobic tails pointed to the center of the membrane and polar heads facing outward; it was thought that proteins were adsorbed on their outer surfaces. Recent experiments, however, have cast considerable doubt on this simple, lamellar model first proposed by Davson and Danielli; it now seems likely that the structure of the membrane is a good deal more complicated and is, moreover, in a constant state of movement, turmoil, and physical rearrangement, in the course of which molecules are in some fashion transported through it.

As we have implied earlier, the actual transport process involves two components: a mediative agent, which "recognizes" and binds specific species of molecules, and a second component, which in some way couples energy to the mediator, altering its physical conformation and thus moving the molecule through the membrane. Of the coupling component relatively little is known; the mediators



Patterns of inheritance of the inborn error of amino acid transport known as "hereditary iminoglycinuria." Two mutant genes cause abnormal excretion of proline, hydroxyproline, and glycine. Subjects with only one mutant allele are either normal or excrete an excess of glycine alone. Two mutant alleles account for the autosomal recessive iminoglycinuria trait, which either is a homozygous phenotype or the result of a "genetic compound" (two different mutant alleles at one gene locus).



Renal clearance patterns at changing plasma concentrations of the two amino acids point to existence of different characteristics of their respective renal transport systems. Proline clearance is very low at low concentrations, indicating highly efficient absorption transport in the tubule. The rapid rise in curve above 0.75 mM shows saturation of the system. Phenylalanine transport is much less efficient initially (clearance rate is higher at low concentrations) but it has a high capacity, since it shows no signs of saturation at any level tested.

in a few cases, however, have been identified quite definitely as transport proteins – macromolecules with molecular weights on the order of 30,000. In many respects they resemble enzymes, in particular with regard to their specificity of action. Just as an enzyme is "keyed" to unite with and transform a particular substrate compound, so the transport proteins are keyed to bind to and transport their own specific substrates. And just as the presence or absence of particular enzymes in a cell ensures that only certain chemical transformations will take place, so the presence or absence of particular transport proteins on – or in – the cell membrane ensures that only certain specific substances can readily enter or leave the cell.

When a given substrate binds to either a transport protein or an enzyme, different parts of the substrate molecule are recognized by the transport and enzyme proteins. This seems plausible since the purpose of the two types of binding is quite different: enzymes transform the substrate chemically, while transport proteins move the substrate, usually without altering it chemically, from one place to another. The two processes thus confer a high degree of selection and specificity to the molecular traffic which is allowed by a cell.

Experiments have also established the existence of certain similarities between transport proteins and enzymes. Just as certain enzymes can bind to several different substrates (almost invariably close chemical relatives of one another) so, it would seem, can certain transport proteins. For strict accuracy, however, we had best refer to them, now and hereafter, as transport *sites* on the plasma membrane rather than proteins, since so far as the "multipurpose" sites at least are concerned, we do not know at present whether one or several proteins are involved.

Many of the experiments in question involve the kidney – both human kidneys in vivo and rat kidneys in vivo and in vitro. As is well known, the mammalian kidney helps maintain the chemical balance of the body's fluids not simply by excreting "undesired" substances into the urine but also by reabsorbing certain "desired" substances from the "raw" urine; it is reabsorption, rather than excretion, that plays the primary role in regulating urine concentration of such substances as water and salts in accordance with the body's changing needs and, perhaps not incidentally, helps to conserve various nutrients and other substances that otherwise might be in short supply. Among these are the amino acids, which, derived from dietary protein, circulate in the plasma and other fluids until they enter cells.

Many of the kidney's reabsorptive processes, in particular those involving the amino acids, involve active transport. It will be seen, therefore, that under certain circumstances the concentration of a given amino acid in the urine (measured by elution ion-exchange chromatography) can provide an inverse indication of how adequately the transport mechanisms that handle that acid are functioning: the higher the urine concentration, the less (proportionately) is being reabsorbed in the renal tubules.

One group of experiments concerns the mutual inhibition of amino acid transport within certain classes of these compounds. If, for example, a normal person is "loaded" with proline intravenously, his urine soon shows a marked rise in proline content, indicating that the plasma load of proline has overtaxed the capacities of the renal reabsorptive transport system. In addition, the urine shows a very marked rise in hydroxyproline and also a significant though smaller rise in glycine, strongly suggesting that these compounds must be transported, to some degree, by the same class of sites that handle proline. Similar mutual inhibition, implying similar "shared" transport sites, has been found for the amino acids lysine, ornithine, and arginine and for several other amino acid groups.

Another set of experiments seeks to define the efficiency and capacity of transport sites in terms (again inverse) of renal clearance rates at different plasma concentrations. As one would expect, each amino acid has its own unique clearance relationships, indicating that the transport mechanisms operate with different specificity on different acids – a further bit of evidence for the enzyme-like specificity of transport proteins (or sites) already alluded to.

In addition, however, one finds a marked difference in some overall patterns of clearance at different concentrations. Proline, for example, shows an exceedingly low clearance rate – i.e., is reabsorbed almost completely by the tubule – at plasma concentrations up to 0.75 mM, which is well beyond the usual physiologic range (0.1 to 0.3 mM). Above this level, however,

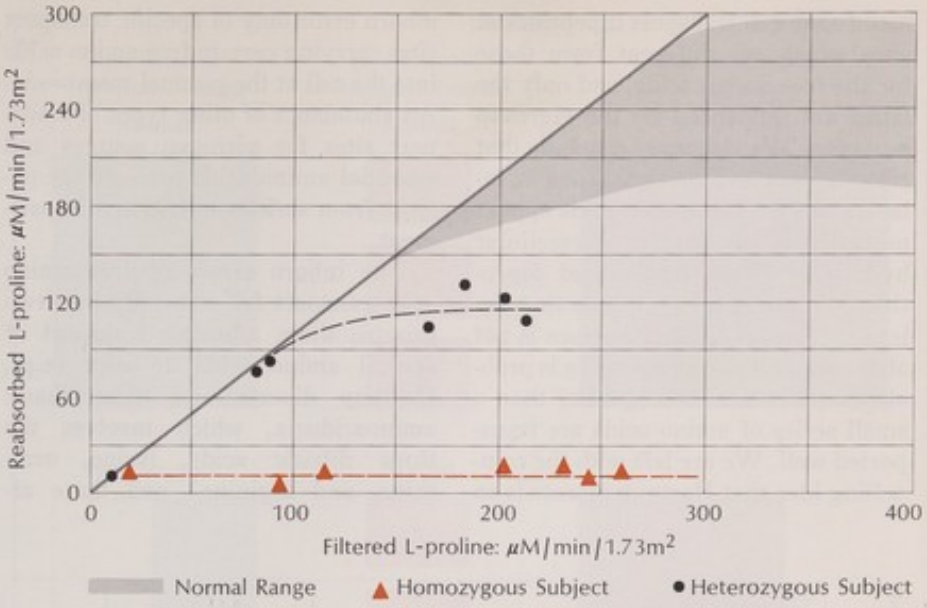
the clearance rate rises precipitously, indicating that the proline transport sites here become effectively saturated.

With phenylalanine, by contrast, the clearance rate at physiologic plasma levels is at least 10 times that of proline, implying a proportionally lower reabsorptive (and, thus, transport) efficiency, but shows no change at levels up to 2.5 mM. Clearly, then, while the efficiency of the phenylalanine transport sites is markedly lower than those of the proline sites, their overall capacity must be much greater, since phenylalanine shows no signs of overloading at more than three times the proline overload level.

These two sets of observations – one concerning the specificity of transport sites (i.e., chemical probes), the other the action of transport sites (i.e., kinetic probes) – have to some extent been complemented by studies on genetic defects of active transfer (genetic probes). But before considering this aspect of the question, we should devote some attention to the defects themselves.

Transport proteins, which resemble enzymes in so many ways, would be expected to resemble them also in being subject to genetic errors, which can produce an inactive or partially active protein (or enzyme), or perhaps none at all. And in fact the accumulation of clinical and laboratory data has established that the transport proteins, no less than enzyme proteins, are subject to inborn errors. More than 20 such defects with additional variants have turned up in man and an even greater number in microorganisms. (For multicellular animals other than man, the number is much smaller, undoubtedly because until recently nobody was looking for them very hard.)

Running down the list of human transport anomalies (see accompanying table), one notes immediately that, as stated earlier, they are of very variable clinical significance. Taking only the amino acid disorders, for example, "classical" cystinuria does not seem to be consistently harmful per se. Some patients develop cystine stones in the kidneys, due to the compound's low solubility, which can of course cause serious problems. Others may be at greater risk of mental illness.



Proline loading experiment reveals the nature of the transport defect in iminoglycinuria. In normal individuals, reabsorption keeps pace with filtration up to quite high levels, reflecting activity of high-capacity system; saturation at lower levels in heterozygotes shows that this system is partially inactive. In homozygotes, residual reabsorption levels are very low. This reflects activity of an unimpaired low-capacity system while the high-capacity system is totally inactive.

Hartnup disease seems harmful only under conditions of marginal nutrition, while one hyperdibasic aminoaciduria, the iminoglycinuria trait, and hypercystinuria do not seem to be harmful. The tryptophan and methionine malabsorption syndromes, on the other hand, would appear to be harmful, though once again the picture is not wholly clear. As in the case of so many other genetic abnormalities, they were first observed in the course of diagnostic studies of clinically abnormal individuals, and we still have little or no data on their occurrence in normal persons. It may well be that with further information they – like so many other hereditary aberrations, e.g., pentosuria – will turn out to be benign in some or even the majority of cases.

Most inborn errors of transport are benign, and most affect epithelialized transport tissues like the small intestine and the renal tubule. It should be remembered that the transport of molecules across an epithelial sheet of this type is asymmetrical, the predominant flux being directed from the luminal or apical surface where the cell surface area is augmented by microvilli (the brush border), across the cell, to exit through the basilar membrane. The latter membrane is not modified by microvilli. Therefore a greater absorptive area on one surface of the cuboidal cell will encourage flux from that surface across to the opposite surface. If the cell also accumulates solute to high concentrations, subsequent entry into the body could be downhill.

We anticipate that the inborn errors of transport either impair entry across the luminal membrane of the epithelial cell, or impair efflux across the basilar membrane. They might also allow abnormal back flux (leak) at the luminal membrane, so that the cell cannot concentrate the solute. We also recognize that some diseases of transport, such as in the Fanconi syndromes, probably affect the coupling of energy-yielding metabolism to the transport sites.

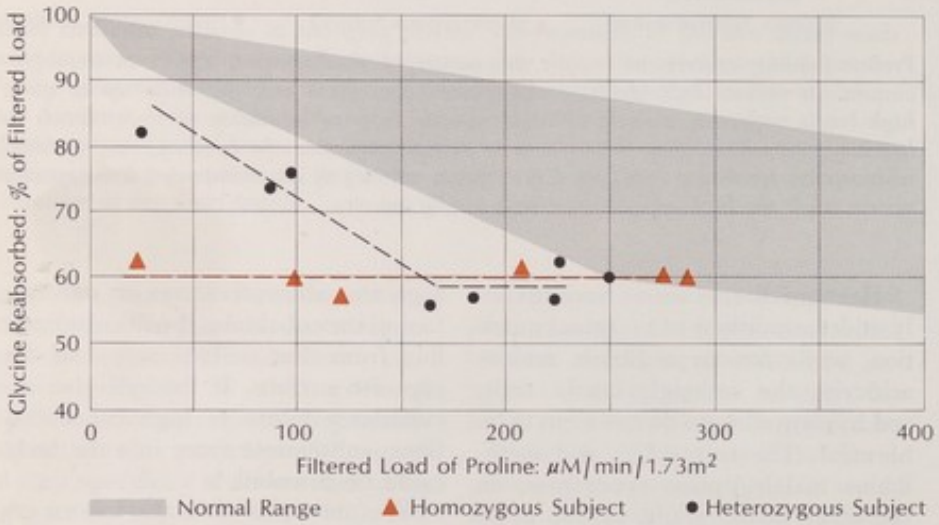
Because transport is so important, one also anticipates that the "benign" inherited traits are compensated for by other functions with adaptive significance. We can visualize this in some recent work on Hartnup disease by Milne's group in England. Patients with the trait do not transport many of the free amino acids normally in gut or kidney. However, dipeptides, comprised of pairs of these various amino acids, are rapidly transported by gut and yield the constituent amino acids in free form in the blood. From this work we deduce that the luminal

membrane can transport dipeptides at sites which are different from those for the free amino acids, and only the latter are influenced by the Hartnup mutation. We also must conclude that efflux at the basilar membrane is intact, since the free amino acids appear normally in blood after intracellular hydrolysis of the transported dipeptide. We might also intimate that the leak rate at the luminal surface is not abnormal and energy coupling is probably intact since solutes other than a small series of amino acids are transported well. We are left with the compelling idea that Hartnup disease is an inborn error only of specific transport sites carrying certain free amino acids into the cell at the luminal membrane. An abundance of other types of transport sites for nitrogen sources and essential amino acids protects the patient from serious nutritional impairment.

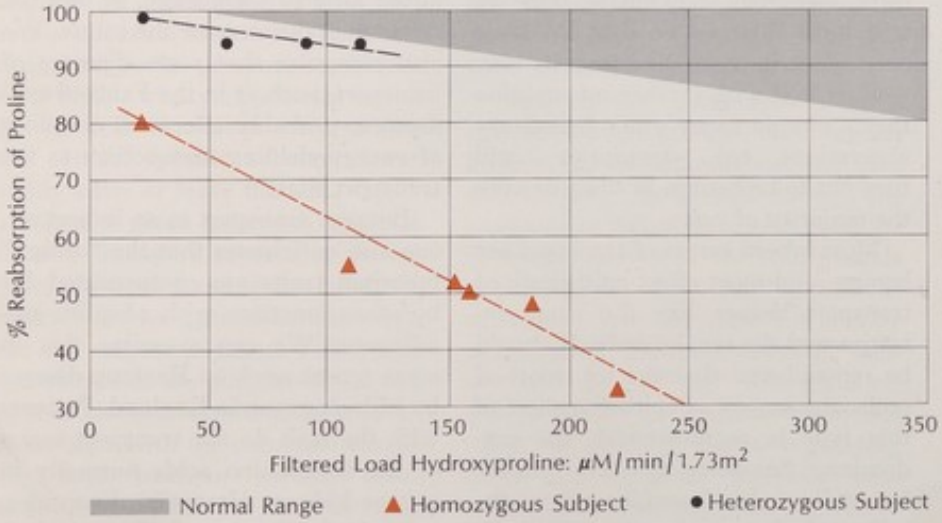
The inborn errors of free amino acid transport fall naturally into two groups; those affecting transport of several amino acids at once (e.g., Hartnup disease and hyperdibasic aminoaciduria, which involves the three dibasic acids, lysine, ornithine, and arginine) and those affecting a single acid (e.g., methionine malabsorption syndrome). The former offer evidence in favor of the "shared" transport sites proposed on the basis of mutual inhibition studies cited earlier. The traits, as a whole, point to the existence of another class of transport systems, each of which handles only one species of amino acid. This supposition is strengthened when we note that in none of the disorders where kidney function can be measured, whether they involve a "shared" or a "single" system, is transport capacity totally missing.

It is possible that the partial loss of function may be due to rather subtle mutations that impair but do not destroy the functioning of the transport proteins (as observed in many enzymatic disorders). But in the light of the other evidence it seems rather more likely that what we are dealing with is two independent transport systems, each under the control of an independent gene, either of which may be totally inactivated without impairment of the activity of the other.

Striking confirmation of this view comes from studies of the families where these genetic anomalies occur. Consider, for example, the case of families carrying the gene for iminoglycinuria – a benign anomaly involving glycine and the two "imino" acids, proline and hydroxyproline (so called because they contain the imino ring made up of four carbons and a nitrogen atom). Pedigree studies indicate a specific pattern of inheritance. Prolinuria and hydroxyprolinuria appear only in individuals who have inherited the abnormal gene from both parents. Glycinuria alone shows up (albeit less markedly) in their parents or offspring, that is in individuals who presumably carry only one mutant allele.



Shared and unshared transport systems are delineated by inhibition experiments in the iminoglycinuria trait. In normals and heterozygotes, both of which have functioning high-capacity systems, glycine reabsorption is inhibited by proline loading, indicating that the two compounds share this system. In homozygotes, however, glycine reabsorption is not inhibited, indicating that the low-capacity glycine system is not shared with proline.



In another experiment in the iminoglycinuria trait, proline reabsorption is inhibited by hydroxyproline loading, not only in normals and heterozygotes but in homozygotes. Since latter possess only the low-capacity system, results show the two imino acids share this as well as the high-capacity system. Conclusion is confirmed by converse experiment in which proline loading inhibits hydroxyproline reabsorption in homozygotes.

Studies of renal clearance rates show that at low normal plasma levels of the imino acids (about 0.2 mM) there is no difference between subjects carrying the mutant allele and normal individuals; in both cases clearance is virtually nil, meaning that tubular reabsorption approximates 100%.

This explanation becomes clear when we load subjects with additional imino acids. In normal subjects, reabsorption keeps pace with rising plasma

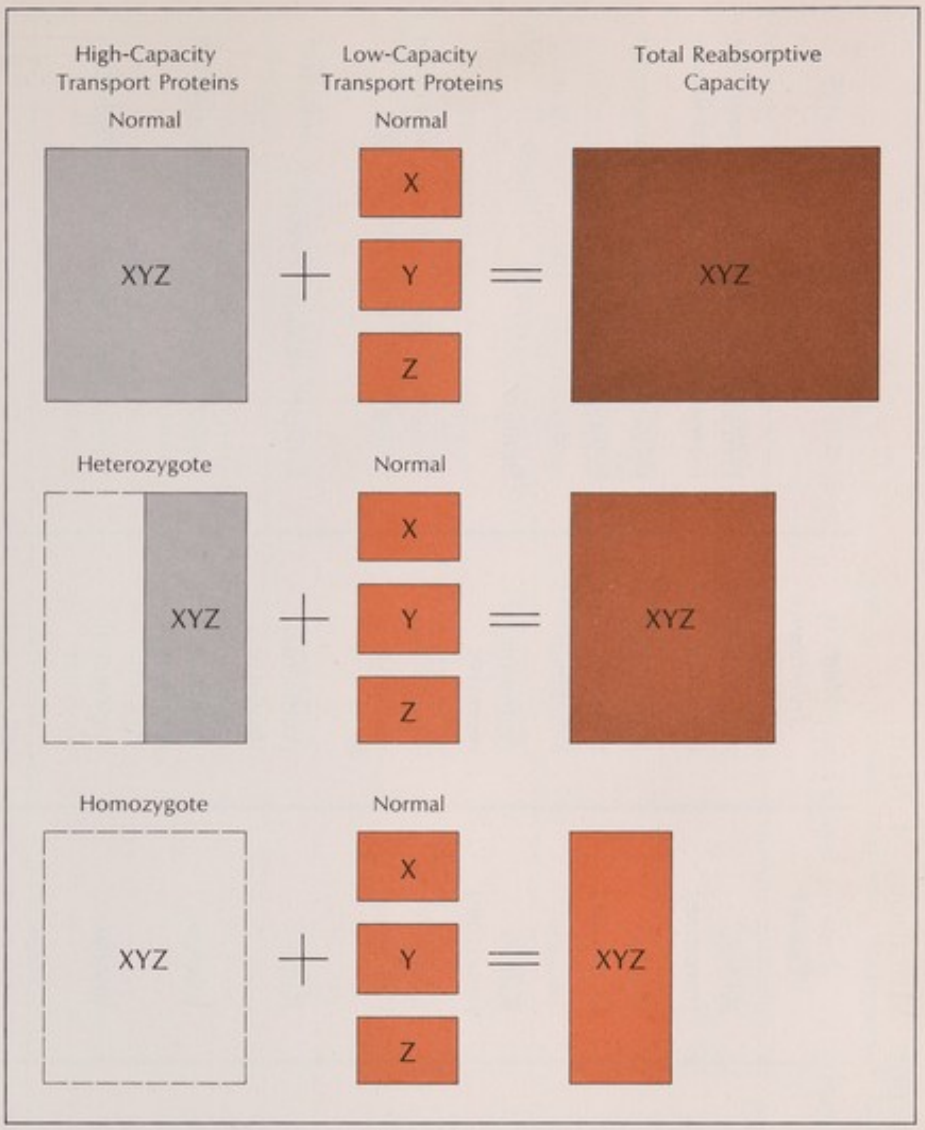
levels, remaining at or close to 100% even at 10 to 15 times the physiologic level; above this point it becomes saturated, with a proportionate rise in iminoaciduria. In heterozygotes, much the same thing happens – but saturation occurs at distinctly lower levels. In homozygotes, no further absorption occurs, indicating that whatever transport function exists was already saturated at the physiologic level (see figure on page 83).

Almost certainly these varying patterns reflect the existence of a low-capacity, easily saturated transport system in the mutant homozygote plus, in the other groups, another, high-capacity, system that is wholly active in normal individuals but only partially active (i.e., shows a lower saturation threshold) in heterozygotes. Presumably this reduced activity results from reduced synthesis of normal transport protein in the heterozygote, due to the presence of only one normal gene controlling that synthesis. (The existence of the low-capacity system in normals and heterozygotes can only be inferred from these data, since its operation is masked by the operations of the high-capacity system. As we shall see, however, there are more direct reasons for assuming its presence.)

Further light on the two systems comes when we measure the capacity of one amino acid to inhibit the reabsorption (transport) of the others. We find that infusions of either proline or hydroxyproline markedly depress the absorption of glycine in both normal subjects and heterozygotes; in homozygotes, by contrast, glycine absorption, albeit low to begin with, remains unchanged. Evidently, then, the high-capacity system, possessed by both normals and (in part) heterozygotes, is also the shared system – which one would have inferred in any case from the nature of the anomaly. In the homozygote, with the shared system entirely missing, glycine reabsorption is handled entirely by a low-capacity, "single" transport system, which remains unaffected by the other members of the shared system.

To complete the picture, one can then measure the inhibitory capacity of proline on hydroxyproline, and vice versa – and here an inconsistency turns up. Either of these amino acids inhibits reabsorption of the other in all three groups – not just the normals and heterozygotes but the homozygotes as well. The only explanation seems to be that these two amino acids are not, as with glycine, handled by two separate low-capacity systems but by one "dual" system. This supposition seems very reasonable when we note, first, that the two are chemically almost identical (differing only by the addition of a hydroxy group), and could thus presumably be easily handled by one system, even a highly specific one, and second, that hydroxyproline is synthesized within the cell from proline so that, teleologically speaking, no separate system for its transport would be necessary.

In the light of these data, the problems with the inheritance pattern of the iminoglycinuria trait can be resolved. In the heterozygote, the diminished capacity of the common transport system means that the three amino acids must compete for its use; evidently, then, the system must have a higher affinity for the imino acids than for glycine. This conclusion, which has been confirmed by other experiments, means that glycine, but



Various experiments suggest model above for some defects in transport of amino acids across renal membranes. Three amino acids, labeled X, Y, and Z, are postulated to be reabsorbed normally through a common high-capacity transport protein as well as by discrete low-capacity proteins specific for each. When genes coding for both systems are normal, reabsorptive capacity is maximal. When heterozygosity exists at locus for common high-capacity system, transport via this system may be substantially reduced. Biggest drop in reabsorptive capacity and greatest likelihood of aminoaciduria will occur with defect resulting from homozygosity at gene locus controlling the high-capacity protein system.

Transport Mutations of Man

	Substrates Affected	*Trait*	*Tissue*[1]	*Mode of Inheritance*	*Clinical Features*
I. AMINO ACIDS Specific groups of amino acids[4]	Cystine, lysine, ornithine, arginine	The "classical" cystinurias	Kidney [intestine]	Autosomal recessive	Urinary calculi (cystine), probable increased frequency of mental illness
	Lysine, ornithine, arginine	Hyperdibasicamino-aciduria(s)	Kidney [intestine]	Autosomal recessive (or dominant)	Protein intolerance, hyperammonemia, vomiting, hepatomegaly
	Imino acids and glycine	Imino-glycinuria(s)	Kidney [intestine]	Autosomal recessive	Benign
	Neutral amino acids (except imino acids and glycine)	Hartnup disease(s)	Kidney [intestine]	Autosomal recessive	None with good nutrition, otherwise pellagra-like rash and ataxia
	Cystine	Hypercystinuria	Kidney	Autosomal recessive	Urinary calculi (cystine)
	Tryptophan	Blue diaper syndrome	Intestine	Autosomal recessive?[2]	Hypercalcemia, nephropathy, failure to thrive
	Methionine	Methionine malabsorp-tion syndrome	Intestine	Autosomal recessive	Odor, diarrhea, mental retardation, convulsions
II. MONO-SACCHARIDES	Glucose	Renal glycosurias	Kidney	Autosomal recessive	Benign
	Glucose and galactose	Glucose-galactose malabsorption	Intestine Kidney?	Autosomal recessive	Profound postnatal diarrhea, failure to thrive
	Phosphate	Familial hypophos-phatemic rickets	Kidney Intestine	X-linked dominant	Rickets/osteomalacia, growth failure
	Calcium (parathyroid hormone binding to membrane)	Pseudohypopara-thyroidism	Kidney Intestine Bone	Autosomal dominant	Symptoms related to hypocalcemia, other signs, e.g., short stature, cataracts, mental retardation
	Calcium (vitamin D binding to nuclear membrane)	Vitamin D dependency	Gut Bone?	Autosomal recessive	Hypocalcemia, hypophosphatemia, hyperaminoaciduric rickets in infancy

III. ELECTROLYTES, WATER, AND OTHERS	Bicarbonate	Renal tubular acidosis	Kidney	Autosomal recessive	Renal tubular (hyperchloremic) acidosis
	Hydrogen ion	Renal tubular acidosis	Kidney	Autosomal dominant	Renal tubular (hyperchloremic) acidosis, nephrocalcinosis
	Chloride	Congenital chloridorrhea	Intestine	Autosomal recessive	Postnatal diarrhea, metabolic alkalosis, hypokalemia
	Sodium	Hereditary spherocytosis	Erythrocyte	Autosomal dominant	Chronic hemolytic anemia, hyperbilirubinemia, splenomegaly spherocytosis
	Sodium and potassium	ATPase deficiency, hemolytic anemia, and	Erythrocyte	Autosomal recessive?	Anemia
		other mechanisms	Erythrocyte	Unknown	
	Water	Diabetes insipidus (vasopressin resistant)	Kidney	X-linked recessive	Dehydration, polyuria, polydipsia, hyposthenuria
	Vitamin B_{12}	Vitamin B_{12} malabsorption	Intestine	Autosomal recessive	Megaloblastic anemia in infancy (intrinsic factor normal)
IV. GENERAL AND OTHERS	Amino acids	Rowley-Rosenberg-Busby syndrome	Kidney Muscle?	Unknown	Growth failure, muscular weakness
	Amino acids, monosaccharides, electrolytes, water, etc.	Fanconi syndrome(s)	Kidney Intestine	Autosomal recessive[3]	Variable, includes hypophosphatemic rickets/osteomalacia, hypokalemia, dehydration, renal tubular acidosis
		Lowe-Terrey-Maclachlan syndrome	Kidney Intestine	X-linked recessive	Oculocerebrorenal syndrome: buphthalmos, hyporeflexia, mental retardation, renal tubular acidosis
	Glucose-glycine	Glucoglycinuria	Kidney	Autosomal dominant	Benign
	Glucose, phosphate, amino acids	Luder-Sheldon syndrome	Kidney	Autosomal dominant	Fanconi syndrome variant

[1] *Bracketed tissues are affected in some forms of the trait and not in others* [2] *Not yet proven conclusively* [3] *Also caused by chemical toxins, e.g., lead*

[4] *Two individuals with impaired renal transport of histidine and lysine, respectively, have been reported to the author*

not the other two, will be "crowded out" and show up in the urine.

From these studies and a number of similar researches into other genetic transport disorders, we can conclude that there is overlap in the function of transport systems in the physiologic range of solute in some human tissues. One system is highly specific, with a transport site singling out only one amino acid (in the case of the imino acids, two). This highly specific system is also highly efficient in transporting the substrates at low plasma concentrations but at the same time possesses a very limited capacity. The other system is far less specific, since a given site will handle several amino acids. It is far less efficient at low plasma levels, but its overall capacity is much higher. Evidently, as in many other contexts, quality and quantity are to some extent mutually exclusive; high efficiency means low capacity and vice versa.

An adaptive role for the dual transport systems is apparent when we consider them in relation to nutrition. When amino acids constitute a very small proportion of the environment, without high-efficiency systems, cells would probably have difficulty in accumulating sufficient amino acids to maintain the necessary protein synthesis. The low-efficiency, but high-capacity systems, on the other hand, enable the organism to take advantage of temporary excesses of amino acids, storing them or utilizing them for energy.

In this connection, it is probably significant that even after a heavy protein meal, measurements of plasma amino-acid levels show only slight increases; evidently the bulk of the protein intake is rapidly passed into the cells.

The "separability" of the two known types of transport systems for free amino acids is further confirmed by studies of their ontogeny during development of the organism or of the cell and of their distribution in various tissues. We know, for example, that reticulocytes in the marrow possess both the low-capacity and high-capacity transport systems, but when they mature into circulating red cells the high-capacity systems disappear. In the kidney, on the other hand, the reverse is true; at birth, the low-capacity single systems handling glycine and the imino acids are missing but appear during succeeding months.

Another factor must now be considered, namely, variability and diversity in the expression of genes regulating a given transport function. Again, genetic probes illustrate this theme well. In family A, depicted in the pedigree chart on page 81, the last child in the third generation has inherited the iminoglycinuric trait. However, only his father shows the hyperglycinuria phenotype which we said was characteristic of the heterozygote with the trait. The mother is completely normal, yet she must carry a mutant allele; her son tells us that this is so. We must conclude that the son is probably a "genetic compound"– the term now recommended for use when two different mutant alleles are inherited at the same gene locus, so that the proband has no normal gene product (in this case, the high-capacity shared transport protein serving binding of the imino acids and glycine). The likelihood that many so-called homozygotes with rare inborn errors are indeed hetero-allelic for the mutant gene, and are thus genetic compounds rather than homozygotes, has become more commonly recognized in recent years. Studies of many iminoglycinuria pedigrees and the extensive work in cystinuria clearly reveal that inborn errors of membrane transport are no different from other types of hereditary metabolic disease in this respect.

The other facet of genetic diversity concerns the expression of mutant transport genes in various tissues of the body. The table indicates that gut and kidney, both rather similar epithelial tissues with regard to their transport role, do not always equivalently express the mutant allele in cystinuria, iminoglycinuria, glucose-galactose malabsorption, and renal glucosuria for example. Nor do diseases of red cell membranes necessarily show themselves in the kidney. We must infer from this that subtle differences in the transport gene products, and in the relevant genes, perhaps exist in the different tissues; or if not in gene products at least in the way they are allowed to express themselves in each tissue. This subtle but important differentiation of transport gene expression in the tissues of the body has been beautifully illustrated by Rosenberg's group; they have now shown that the cystinuria mutation is not expressed in leukocytes, nor is this mutation or that for Hartnup disease expressed in cultured skin fibroblasts obtained from patients who have abnormal kidney and intestinal transport.

Anomalies of this sort – and our as yet very limited information about them – are doubtless a major reason for the confusing and inconsistent pathology of the transport defects. Broadly speaking, one would expect defects of the high-capacity group sites to be less serious than those of the low-capacity specific sites, since at normal physiologic levels (of amino acids, at least) most or all of the transport is handled by the latter. And, broadly speaking, this is what one finds – but there are enough exceptions to pose problems.

Classical cystinuria, for example, which affects a group of amino acids – cystine, lysine, ornithine, and arginine – does not appear to interfere with the metabolism of most cells (though the associated high urinary levels of cystine can, as noted earlier, cause troublesome calculi). Yet there is the further fact that cystinurics are probably at higher than normal risk of mental abnormality, either retardation or psychosis.

As it now is, we are left with a number of hypotheses, none of which can be confirmed at this time. In this respect it is wise to remember that cystinuria is not one disease but two and perhaps three distinct disorders, each the result of a different mutation. Perhaps only one form predisposes to mental illness, perhaps all do.

Hypercystinuria, by contrast, is apparently a defect of a specific low-capacity site handling only this one amino acid; it has, moreover, been at this writing localized only in the kidney. Its apparent benignity may be due to the existence of other transport mechanisms in other tissues that are not affected by the mutation.

One would expect that certain transport defects could be dealt with by loading the body with enough of the substance in question to compensate for the transport deficiency, in a parallel to the treatment of the vitamin-dependent genetic disorders. Just such an approach has proved effective in X-linked hypophosphatemic rickets, in which the body loses phos-

phate as a result of loss of a transport system. Aggressive phosphate replacement, involving sometimes five times the normal intake, appears to compensate for the transport failure, allowing the characteristic bone lesions to heal and normal growth to take place. Logically, one would expect a similar approach to work in the tryptophan and methionine malabsorption syndromes, where the defects in the low-capacity systems could be bypassed by loading with the amino acid in question to the point where the (less efficient) shared system would come into play. Here, however, we face the possibility that the damage in these diseases may be done prenatally, especially in view of their apparent rarity, which suggests that some affected individuals may die in utero. This could be evaluated by determining the heterozygote rates in the population and by looking for a deficiency of homozygotes relative to their predicted frequency.

As will be seen from the above, there is plenty of material to keep researchers in this active field busy for some time. There is, moreover, one additional possibility that is at least worth a brief mention: manipulation of transport mechanisms for therapeutic purposes. Considering what I have just said about the incompleteness of present knowledge in this field, such a notion may seem far off indeed. Yet in at least one experiment it has been done effectively enough to suggest that there are real possibilities here. A few years ago, an artificial amino acid was used to treat malignancy in an experimental line of animals; by binding to the cells' transport system, it effectively excluded essential amino acids, leading to starvation of the tumor. Unfortunately, the toxicity of the drug was far too high for clinical use but conceivably some modification of the approach may yet prove practically useful. Further off, one can envision the design of drug molecules to "fit" a given transport system so that they would penetrate only cells with that system. A very long-range goal, perhaps, but one worth working for.

9

The Genetics of Drug Reactions

BERT N. LA DU JR.
New York University

Time was when many doctors tended to assume that if one picked the proper drug and gave it in the recommended dose it would do what it was supposed to do – help the patient without significantly harming him. We now know, of course, that this concept is grossly simplistic; proper drug dosage depends not merely on the drug but on the patient – age, weight, the particular stage of disease or diseases, even sex.

Yet even giving the fullest weight to all these factors, we still find cases in which the "proper" dose of a given drug will be toxic to some people and/or ineffective in others. And it is becoming apparent that a major factor in these "residual" drug anomalies is heredity – the unique genetic make-up of each individual in question. The study of these hereditary differences in drug response is known as pharmacogenetics. The growth of knowledge in the field has obvious practical implications for the devising of drug regimens that are fully rational, in the sense of being tailored to the individual patient. In addition, to the degree that discussion of these problems helps focus attention on the extent and range of variability in drug responses generally, it can, I think, prove beneficial in our application of all drug therapies.

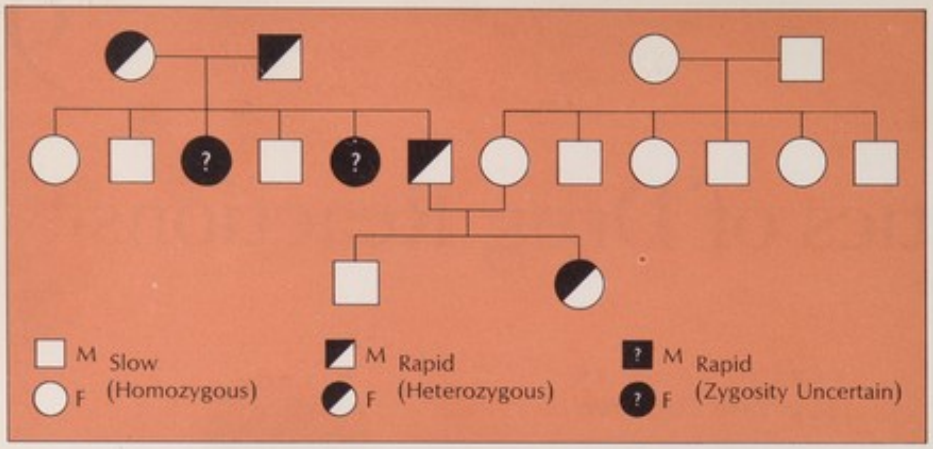
Pharmacogenetics is a relatively new "specialty"; the very name was coined only about 10 years ago. Yet the fact that it has come to the fore so recently is certainly surprising; with hindsight, indeed, one might expect the subject to have drawn attention many years ago. The existence of different drug responses in different species is hardly a recondite concept; in fact, it constitutes a notorious pitfall in the path of the pharmacologic researcher. Obviously, these differences must stem from genetic differences among species. The concept of hereditary physiologic differences among members of our own species is, if anything, even more venerable, dating back at least to around 1905 and the concept of "inborn errors of metabolism." What could be more expectable, then, than that certain inherited metabolic defects in man would influence the individual's responses to drugs?

In fact, anomalous drug responses have long been part of the recognized symptomatology of some inherited diseases. In congenital porphyria, for example, barbiturates can trigger an attack, sometimes with fatal results. Individuals with Crigler-Najjar syndrome are abnormally sensitive to salicylates, those with von Gierke's disease show a decreased blood sugar response to epinephrine and glucagon, sufferers from Down's syndrome are hypersensitive to atropine. Instances like these are not particularly startling; since the patient's metabolism is pathologically deranged to begin with, a pathologic or at least abnormal reaction to a drug is a reasonable expectation.

Rather more remarkable are a number of metabolic abnormalities (one cannot quite call them errors) that are not expressed under "normal" conditions and that seem to be essentially or wholly benign. However, in the presence of certain drugs they express their potentially dangerous character. In this category is the abnormal hemoglobin known as Zurich. In heterozygotes, the condition is apparently harmless unless the individual receives sulfonamides or other oxidizing drugs, which produce methemoglobinemia and a fulminating hemolytic anemia. Similar adverse effects have been observed with some other abnormal hemoglobins.

Another example is inherited deficiency of the enzyme glucose-6-phosphate dehydrogenase, a condition thought to affect some tens of millions of people. Here again the pathology is essentially drug dependent, and hemolysis results from a number of compounds – the antimalarial primaquine, para-amino salicylic acid, phenacetin, and the sulfonamides – as well as from certain constituents of the broad, or fava, bean. Many variants of this enzyme defect have been identified; not all result in drug sensitivity.

Even more subtle, though not necessarily less serious, are metabolic abnormalities in which the only detectable "pathology" is an accentuation or a diminution of the usual response to the drug. One of these idiosyncrasies,



Inheritance of isoniazid excretion traits shows that rapid (R) *gene is dominant over slow* (r) *gene. Slow excretors must be* rr *homozygotes; rapid may be either* RR *homozygotes or* Rr *heterozygotes. The latter two groups can be distinguished only to the extent that the traits of their offspring permit inferences to be drawn about their zygosity.*

indeed, is so common as to make the term "abnormality" of dubious relevance. This involves isoniazid and some related compounds.

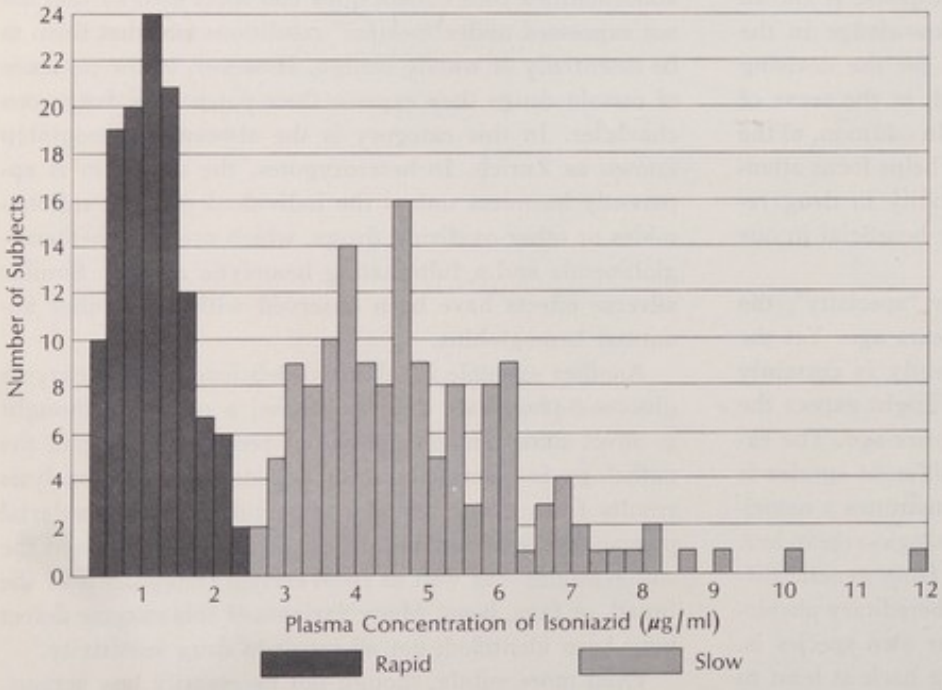
Not long after isoniazid (INH) was introduced as an antitubercular agent in the early 1950's, clinicians noted that patients showed remarkably wide variation in their metabolism of the compound, whether this was measured by the decline in serum levels of INH or the rate of its excretion in the urine as acetyl-INH. Studies of identical and fraternal twins showed that the metabolic differences were largely hereditary; in fact, calculations indicated that genetic factors accounted for about 97% of the observed differences.

Studies of INH metabolism rates in larger groups then revealed a clear bimodal distribution, with the population divided into "rapid inactivators" and "slow inactivators." When family pedigrees were analyzed in the light of these findings, it appeared that INH inactivation is controlled by two autosomal alleles at a single gene locus, identified as *R* (rapid) and *r* (slow). Both homozygotes and heterozygotes for the "rapid" allele (*RR* and *Rr*) are rapid inactivators; only the "slow" homozygotes (*rr*) are slow inactivators. In other words, slow inactivation of INH is a recessive trait.

What is extraordinary in this instance is that something like half the population falls into the slow group, a frequency far above that for most recessive traits. Why this should be so is not known. In the few cases in which recessive genes are known to be widespread in a given population, the gene can be expected to confer some advantage on its possessor. Thus sickle-cell trait, which reaches a frequency of around 40% in some African groups, gives a degree of protection to the heterozygote against malignant falciparum malaria, which is endemic in the homelands of these populations. We know of no advantage conferred by the slow INH trait. On the other hand, there is no obvious natural disadvantage in slow inactivation. Indeed, given the high proportion of slow inactivators, it seems unlikely that any such disadvantage could exist. The only disadvantages we know of are drug dependent, in particular an enhanced tendency toward cumulative toxicity – a peripheral neuritis – from INH, which can be controlled by administering the vitamin pyridoxine.

Ideally, one would prefer to prevent the side effects by reducing the dose, since the slow inactivator can derive the same therapeutic benefit from a smaller quantity of the drug. In practice, however, determining whether the patient is in fact a slow inactivator requires rather specialized tests, which cannot be performed in most places. Significantly, slow inactivators are also more likely to show toxic side effects from two drugs chemically related to INH – the antidepressant phenelzine and the antihypertensive hydralazine. In the latter case they sometimes develop a lupus erythematosus–like syndrome.

Slow inactivation, like many other metabolic anomalies, has been traced to a deficiency of a particular enzyme, in this case *N*-acetyltransferase of the liver, which metabolizes INH and related compounds into the acetylated form in which they are excreted. Present thinking is that slow inactivators simply manufacture less of the enzyme. However, some recent evi-



Measurements of plasma isoniazid levels in 267 individuals, six hours after oral dose of drug, show marked differences in excretion rates. Clear bimodal distribution shows division of population into rapid (left) and slow (right) excretors.

dence suggests that the enzymes associated with the *R* and *r* alleles may differ slightly in their properties, including, conceivably, their efficacy as acetylating agents.

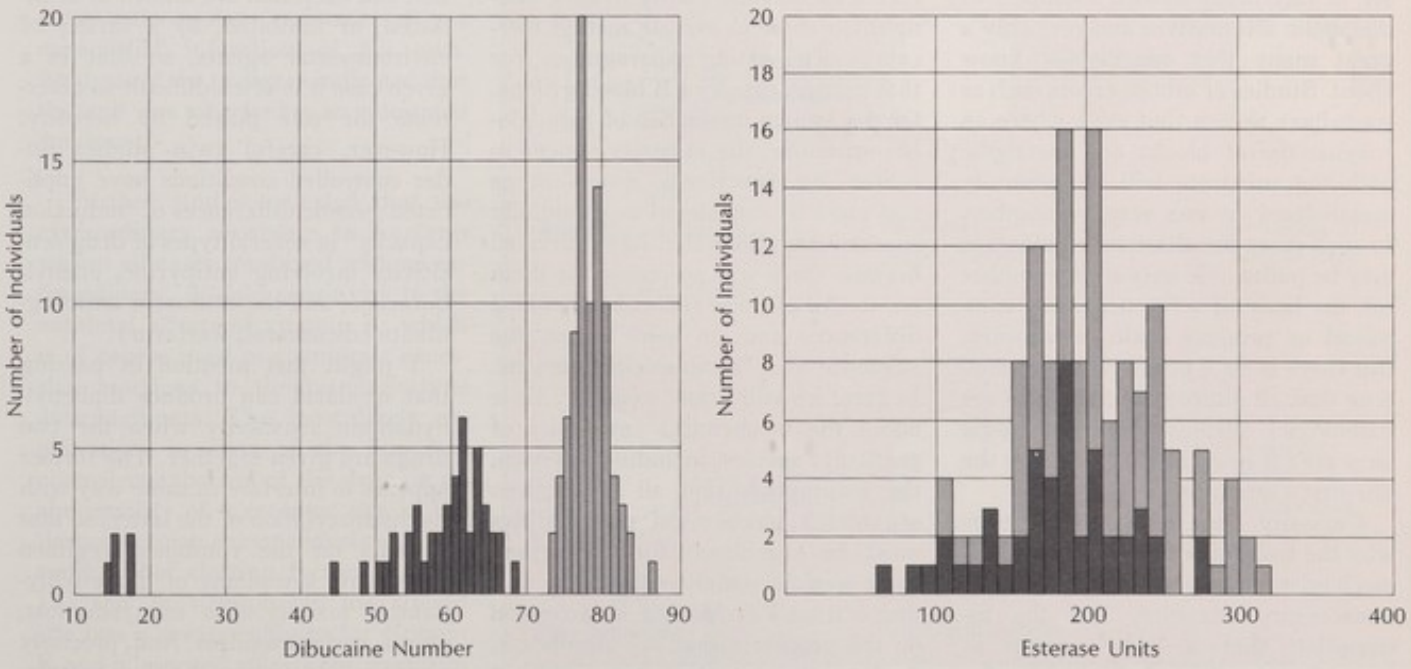
A rather more complex enzymatic situation is involved in another subtle, and sometimes very serious, inherited drug reaction: hypersensitivity to succinylcholine. In the body this muscular relaxant compound is rapidly hydrolyzed to succinylmonocholine, an inactive metabolite, so that only a fraction of the dose actually reaches the anatomic site of action – the myoneural junctions in skeletal muscles. When used as an adjuvant in surgical anesthesia, it of course affects the respiratory muscles as well, producing a brief apnea normally lasting only a few minutes. Occasionally, however, individuals show prolonged respiratory paralysis and may require artificial respiration for an hour or more.

Biochemical and neurologic studies of such individuals showed that the problem did not involve any abnormality of the myoneural receptor sites but, rather, abnormally high serum levels of the drug. Plainly, it was being hydrolyzed much more slowly than normal. This pointed to some problem with the hydrolyzing enzyme plasma pseudocholinesterase (PPC-ase), but further research failed to reveal any obvious deficiency in the enzyme. Population studies showed a normal distribution of PPC-ase levels, and even individuals known to be succinylcholine-sensitive possessed enzyme levels that, though generally on the low side, were by no means low enough to account for their anomalous reactions to the drug.

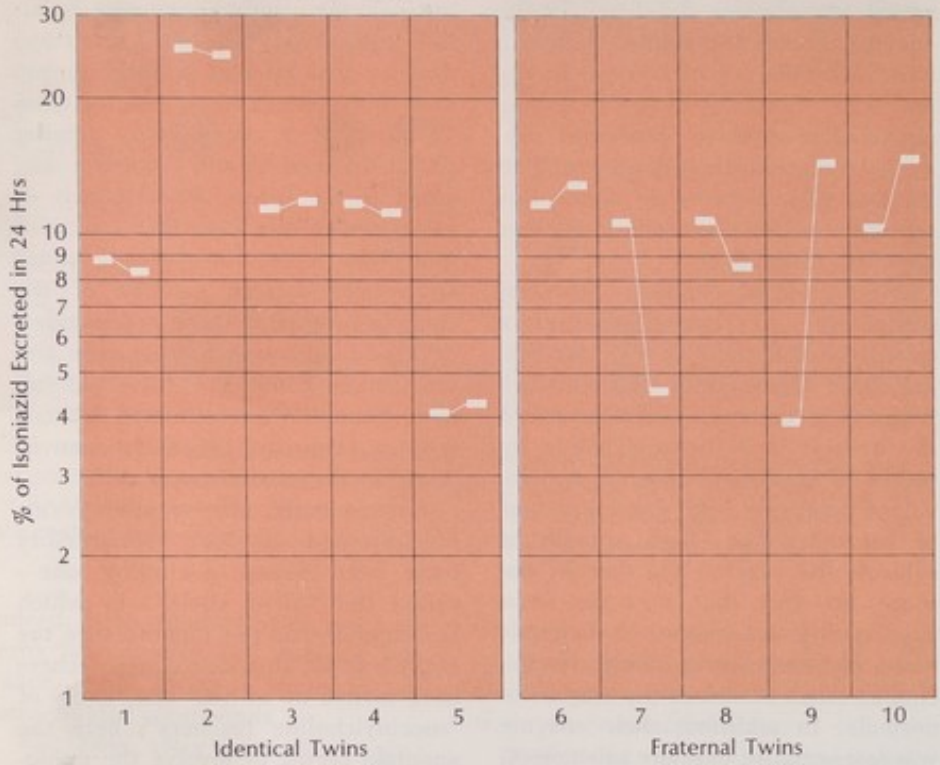
More refined studies, however, turned up not a quantitative but a qualitative difference in PPC-ase that had been obscured by the standard tests for its activity. These measured the hydrolysis of benzoylcholine by means of spectrophotometry, and involved relatively high concentrations of benzoylcholine – high enough to saturate the enzyme and thereby obscure the fact that PPC-ase from succinylcholine-sensitive individuals had a distinctly lower affinity for the choline esters than does the enzyme in normals. In addition, their enzyme was *less* sensitive to many substances, such as dibucaine and fluoride, that normally inhibit its action. By "playing off" the inhibitor against the benzoylcholine substrate the researchers were able to demonstrate the lowered affinity property of the atypical enzyme and also to classify individuals according to their "dibucaine number," i.e., the percentage of inhibition of "their" enzyme. Grouped in this manner, the original normal distribution immediately separated into three distinct populations: a large group with dibucaine numbers ranging from 73 upward, a considerably smaller group between 45 and 70, and a very small group below 20. Analysis of pedigrees on this basis then indicated that those in the intermediate group were heterozygous carriers of the trait, in contrast to those in the under-20 group, who were homozygotes and true succinylcholine-sensitive individuals. The latter are estimated to occur with a frequency of about one in 3,000 in the general population.

Subsequently, several other forms of inherited PPC-ase abnormality have been found, including one – called the "silent allele" – in which homozygotes do not manufacture the enzyme at all. In addition, researchers have turned up at least one family of "succinylcholine resisters"; here the anomaly seems to involve the manufacture of unusually high quantities of the plasma enzyme.

I should stress that none of the individuals with anomalies of either *N*-acetyltransferase or PPC-ase are in any way abnormal except for their drug reactions – not even the "silent allele" homozygotes with essentially no PPC-ase at all. One must conclude that either the enzymes in question



Dibucaine inhibition test of the enzyme pseudocholinesterase yields a trimodal distribution (graph at left). Small group of mutant homozygotes is hypersensitive to succinylcholine; heterozygotes and normal homozygotes react normally to the drug. In contrast, conventional test of enzyme activity (graph at right) produces a unimodal distribution, with no clear distinction among any of the groups; mutant homozygotes and heterozygotes (deep color) are seen to be scattered through most of the range.



Genetic source of differences in excretion rate of isoniazid is shown by comparisons of identical and fraternal twins. The former in all cases have very similar rates; the latter may but almost invariably do not. From such studies it has been estimated that genetic factors account for about 97% of observed individual differences in isoniazid metabolism.

serve no normal function in the body or that such normal functions as they serve can be carried out adequately via other metabolic pathways involving other enzymes. There are in fact many known examples of metabolic alternatives and probably a great many that we do not know about. Studies of inborn errors such as PKU have shown that even where an enzyme defect blocks one metabolic path the substrate will be generally metabolized in one way or another. In such cases the alternative pathways may be pathologic because they either rob the body of some necessary compound or produce toxic metabolites. But there is no a priori reason to suppose that all alternative pathways are pathologic; obviously the pathologic ones are far more likely to come to the scientist's attention.

Curiosity may lead us to wonder why the body appears to have evolved mechanisms for providing seemingly unnecessary enzymes. But the assumption that a "why" exists in every case involves, I think, an overteleologic view of human evolution. There are, after all, any number of inherited physiologic variations in the human animal, some of which possess, so far as we can tell, no functional or evolutionary significance whatever. These include such obvious traits as straight vs wavy hair or long vs round heads as well as more subtle differences, such as the ability to taste phenylthiourea or to excrete methyl mercaptan after eating asparagus or, for that matter, the A vs B blood groups. Of the known anomalies of hemoglobin structure, the majority appear to confer no functional disadvantage and must be considered as essentially random mutations that have survived because there was no reason for them not to. All of these reflect biochemical differences and, in some cases, the existence of a "nonessential" enzyme. In fact, we still know relatively little about the biochemical variability of particular species, including our own; the assumption that all of a given organism's biochemical potentialities must be significant for the organism's welfare mainly reflects the fact that science has focused its attention on the processes that *are* significant. And to the extent that we subject human beings to more and more "unnatural" biochemical challenges, i.e., by giving them drugs, we are likely to turn up more and more "nonfunctional" biochemical processes.

But the puzzle of "unnecessary" enzymes does not exhaust the complexities of inherited anomalous drug reactions. In the case of the anticonvulsant drug diphenylhydantoin, for example, some individuals show inherited sensitivity resulting from a lag in conversion of the drug to a more soluble compound that is one stage in its disposition. This lag is thought to reflect some abnormality in the liver enzyme microsomal hydroxylase, but according to present evidence the enzyme itself seems normal. There are indications, rather, that the problem may be an inherited inability to "induce" the enzyme. Diphenylhydantoin, like many other drugs (notably, e.g., the barbiturates), has the ability to step up production of the enzyme(s) that metabolizes it; in "sensitive" individuals this induction may not occur, so that the enzyme, though otherwise normal, remains at its basal level. Exactly why this should occur (if it does) is uncertain, since we know very little about the mechanisms that control enzyme induction, apart from the fact that some of them appear to be inheritable.

The problem is further complicated by the fact that the liver microsomal enzymes involved in drug detoxification and excretion are known to be induced, or inhibited, by a variety of environmental agents, so that in a given case it is often difficult to determine the role played by heredity. However, careful twin studies under controlled conditions have implicated genetic differences of "induction capacity" in several types of drug sensitivity involving antipyrine, phenylbutazone, and the coumarin anticoagulants (dicumarol, warfarin).

I might just mention in passing that isoniazid can produce diphenylhydantoin sensitivity when the two drugs are given together. The former appears to interfere in some way with the hydroxylation of the latter, so that patients on the combined regimen may show symptoms of diphenylhydantoin toxicity such as nystagmus, ataxia, and drowsiness. And, precisely as one would expect, the slow acetylators of isoniazid are the most likely to show diphenylhydantoin intolerance.

All these abnormalities, whether of enzyme structure or enzyme induc-

tion, essentially concern the rate at which the drug is removed from the body, and thus the tissue concentration of the drug that will be produced by a given dose. In other words, the individuals are abnormal only as regards their reaction to a given *dose* of the drug in question; considered in terms of the serum or tissue level of the drug, their reactions are within the normal range.

A quite different type of inherited drug anomaly is inherited resistance to coumarin-type anticoagulants (not to be confused with the inherited sensitivity cited above). The condition was first observed in an elderly man who was receiving warfarin prophylactically following a myocardial infarction. The usual dose was ineffective, and the desired increase in prothrombin response was only achieved when the man was given a daily dose of 145 mg – some 20 times the usual quantity. Yet careful studies showed that the patient did not metabolize warfarin at an abnormally rapid rate; serum levels of the drug and its distribution in the tissues were normal with normal doses and elevated in proportion to the elevated dose.

Tests of relatives turned up five other individuals with similar warfarin resistance, including the patient's identical twin, who had shown the same anticoagulant resistance after a myocardial infarction of his own. Analysis of the pedigree indicated that the trait was inherited as an autosomal dominant.

Further studies revealed that the extraordinary resistance to warfarin was in all cases combined with an extraordinary, 20-fold sensitivity to the antidotal effects of vitamin K, which is of course used to counteract excessive reactions to the coumarin-type anticoagulants. The most likely explanation, in view of the apparent normal metabolism of the drug, is an abnormality of a receptor site in the tissues or in an enzyme involved in the synthesis of clotting factors II, VII, IX, and X. Presumably the altered site has a greater affinity for vitamin K and a lessened affinity for warfarin, markedly enhancing the former's ability to compete with the latter in binding to the site.

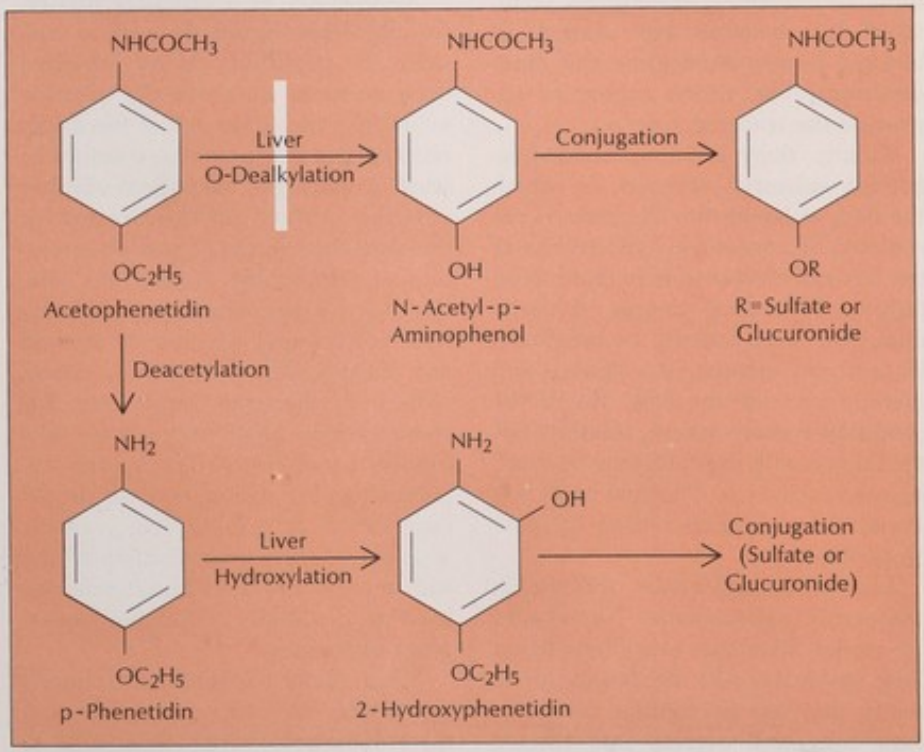
As with other inherited disorders, the mechanisms of inherited drug anomalies are not always well understood, so that a general classification is difficult. Broadly speaking, we find that the disorders, and indeed individual differences in drug reactions generally, do *not* involve the rate at which the drug is absorbed or the way it is distributed in the tissues. Probably this is because these processes are determined mainly by the physicochemical properties of the drug itself (its water and lipid solubility, ionizing properties, etc.). Individually anomalous drug reactions, on the other hand, may be roughly classified according to the basic mechanism involved.

In a *receptor site* disorder, a normal dose of the drug produces a normal blood level, but the patient, because of his "insensitive" receptor sites, requires an abnormally high dose, and blood level, to achieve the desired response. Here the only example presently known is coumarin anticoagulant resistance. (In theory, one can conceive of a converse situation, in which the receptor would possess an abnormally *high* affinity for the drug, but I know of no actual example.)

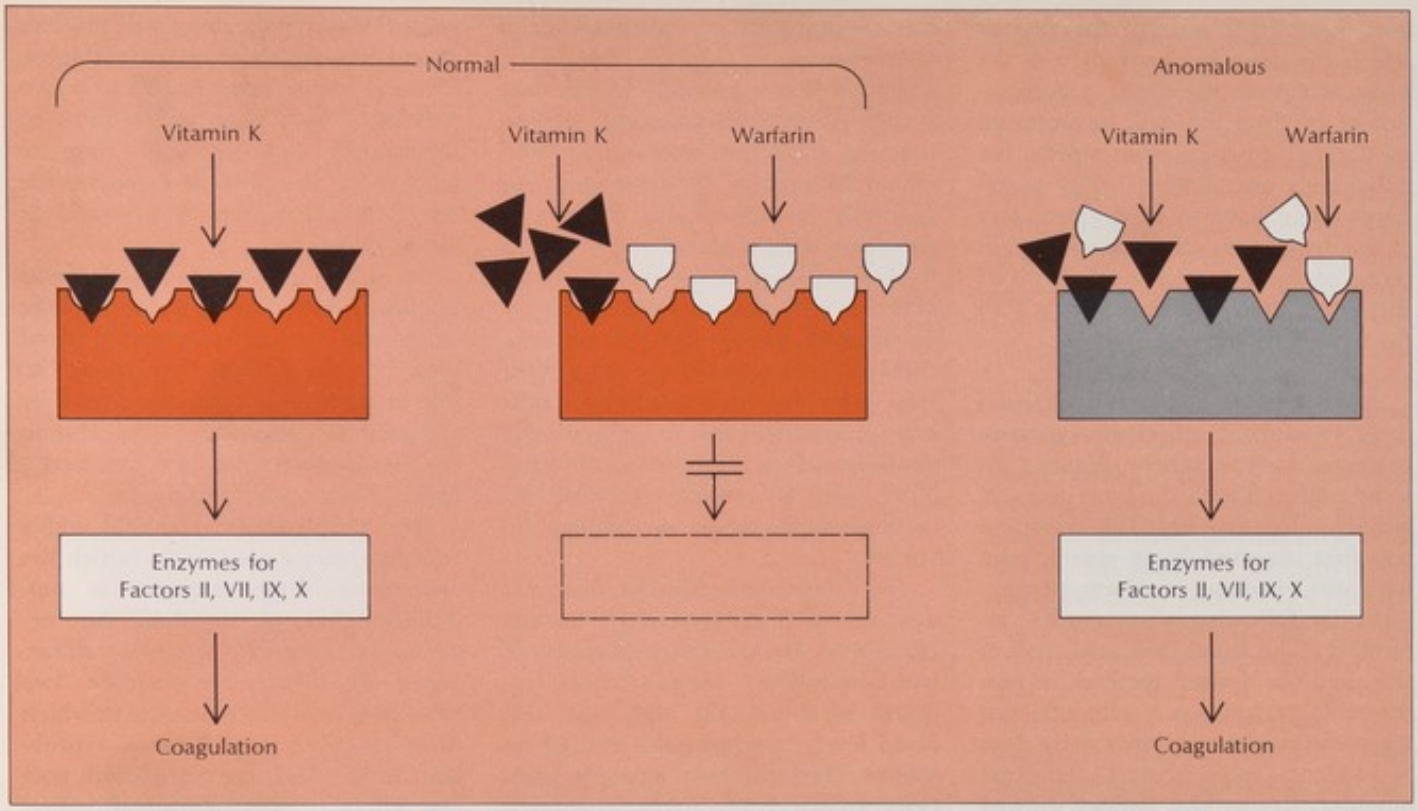
In *drug metabolism* disorders, a normal dose of the drug will produce an abnormally high or low blood level. The patient reacts normally to a normal blood level, but will require an abnormally low or high dose to achieve it. The obvious examples are the slow inactivators of isoniazid or succinylcholine.

In what might be called *tissue metabolism* disorders, a normal dose of the drug produces a normal blood level but the patient "overreacts" to this blood level because his tissues or red cells are peculiarly vulnerable to the drug. Examples are the hemoglobin Zurich and favism.

For completeness, I should add a few *anatomic disorders*, in which the anomalous drug response is not, strictly speaking, metabolic but depends on inherited anatomic differences. There are, for example, two types of inheritable glaucoma in which drug reactions may become a problem. In the first, associated with narrow angle of the anterior chamber, patients may develop increased intraocular pressure in response to atropine and other mydriatics, which tend to block rather than facilitate the flow of the aqueous humor. In the second, which appears to result from a struc-



Normal metabolic path of acetophenetidin proceeds via dealkylation to conjugated form in which metabolized drug is excreted. Inherited enzyme defect blocks normal path (white bar), shunting drug into abnormal path. This produces 2-hydroxyphenetidin as intermediate, which is toxic, inducing hemolysis and methemoglobinemia.



Inherited resistance to anticoagulants of the coumarin type is thought to involve abnormal receptor sites in tissues or in an enzyme. In normal individuals, warfarin binds preferentially to the sites, thereby displacing vitamin K and blocking the formation of coagulation factors. In certain abnormal individuals, who are rare, the sites are more receptive to vitamin K. The result is that the patient is both insensitive to the anticoagulant and hypersensitive to the antidotal effects of the vitamin.

tural difference in the trabecular meshwork system, drainage is facilitated by atropine but slowed or blocked by dexamethasone and other corticosteroids, which appear to aggravate the trabecular defect.

Finally, there is the problem of inherited subaortic stenosis, in which the flow of blood into the aorta is obstructed by concentric hypertrophy of the left ventricular muscle. Here mild failure, with some degree of aortic dilatation, may actually be beneficial, since it may alleviate the stenosis and thereby increase the flow. But if the condition is more severe, treatment of the failure with digitalis may be fatal, since a restoration of normal heart size would decrease rather than increase blood flow.

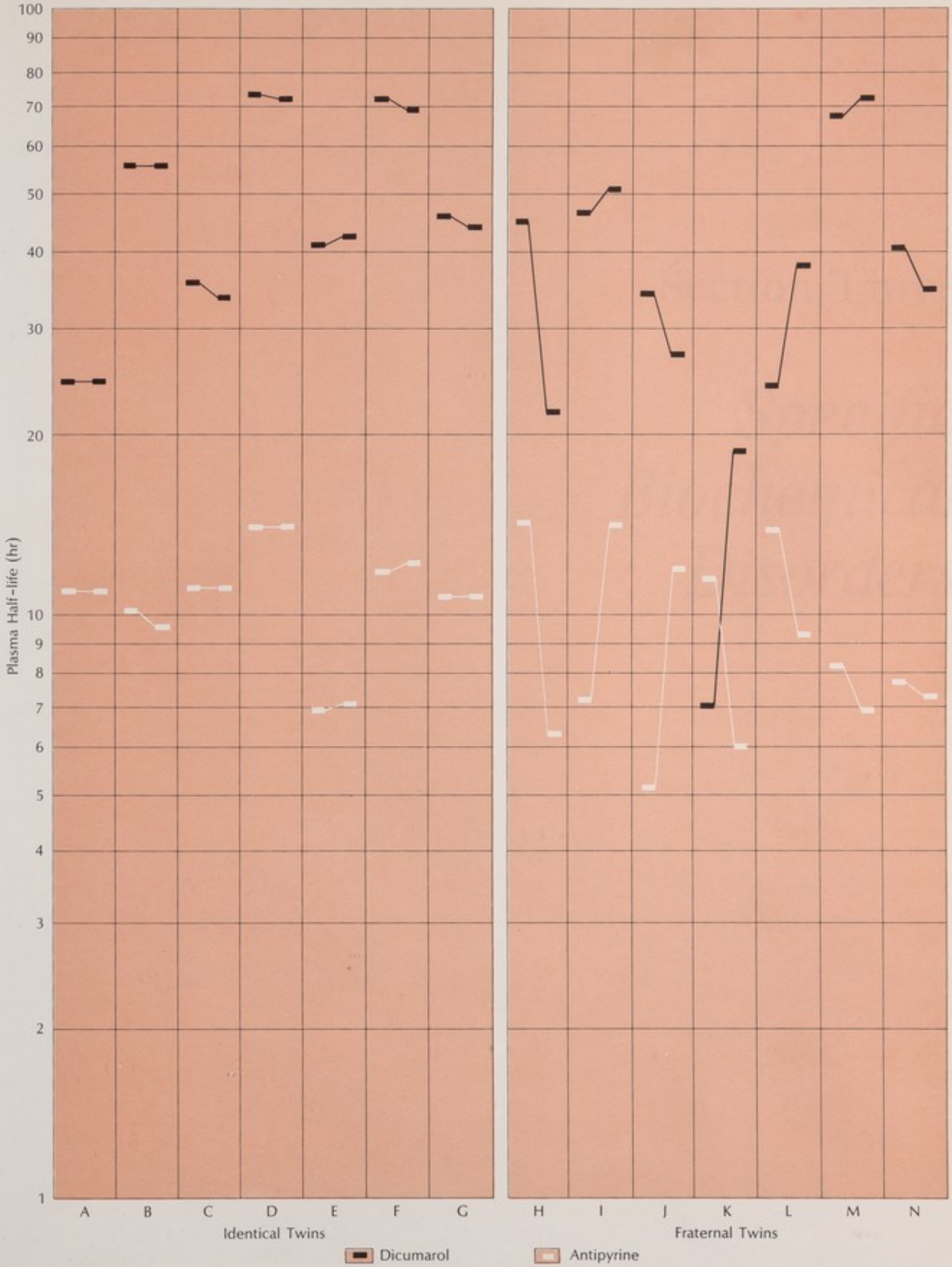
Like all classifications, this one is necessarily approximate. Nor should the reader infer that every hereditary drug anomaly will be found to fit neatly into one or another category; some may well overlap into two or three. This is especially likely to be true of drug anomalies that, like some genetic diseases, are polygenic and multifactorial – and there is already evidence that such anomalies exist.

Meanwhile, as a practical matter, the physician would do well to consider the possibility of an inherited drug anomaly whenever the question arises of prescribing one of the drugs cited herein. In particular, it would be worth inquiring of the patient whether any close relative has been treated for the same disease and, if possible, ascertaining (preferably through the relative's physician) what drugs the relative received and whether he showed any inexplicable reactions to them. Where a physician treats (or has treated) more than one member of a family, he of course has special opportunities for observation; he should especially bear in mind that an otherwise inexplicable drug reaction of any sort in one of several siblings raises at least the possibility of a similar occurrence in another.

With further studies and clinical experience, we can expect more and more inheritable drug anomalies to turn up. I suspect, too, that increasingly we will find them falling into clusters, so that an established sensitivity to one drug will imply a presumptive sensitivity to several (this is already the case with, for example, primaquine sensitivity in glucose-6-phosphate dehydrogenase deficiency).

The need for further research as well as for an enhanced sensitivity in the profession generally to the hereditary factor in pharmacotherapeutics can be seen from a brief cautionary tale. Not so many years ago, a very extensive and expensive study was undertaken to ascertain why an occasional patient (perhaps one in 10,000) showed severe toxic reactions to the anesthetic halothane. It apparently occurred to no one connected with the study that heredity might be of significance. The results of the study were rather inconclusive – and so far as evaluating possible hereditary factors are concerned, will remain so, even with present hindsight, since no data bearing on heredity were ever obtained. If we are ever to determine whether halothane sensitivity is in fact inheritable, we will have to start all over from the beginning and collect the proper data.

One hopes that such an omission would not occur today.



Twin studies have demonstrated genetic control over the metabolism of dicumarol, antipyrine, and phenylbutazone (not shown); in all cases, metabolism was found to be far more similar in identical than in fraternal twins (data from Vesell and Page).

Section Three

Specific Biochemical Disorders

10

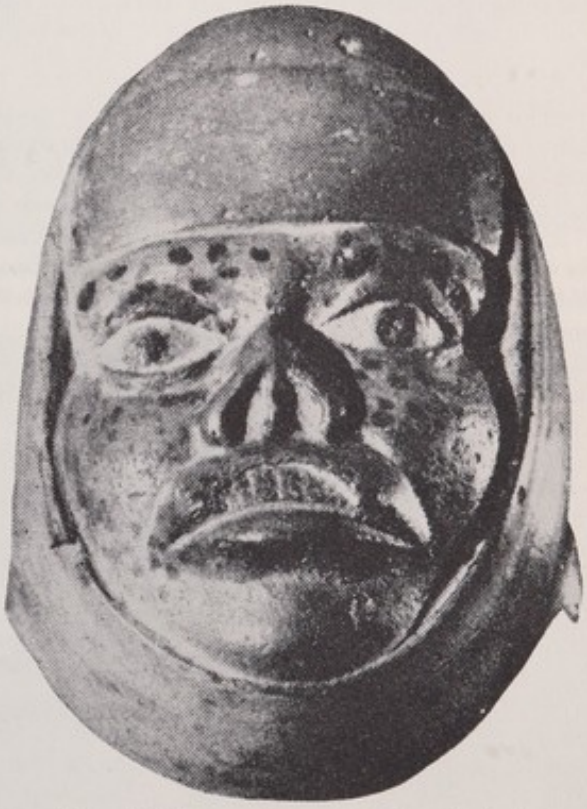
Lesch-Nyhan Syndrome and the X-Linked Uric Acidurias

J. EDWIN SEEGMILLER
University of California, San Diego

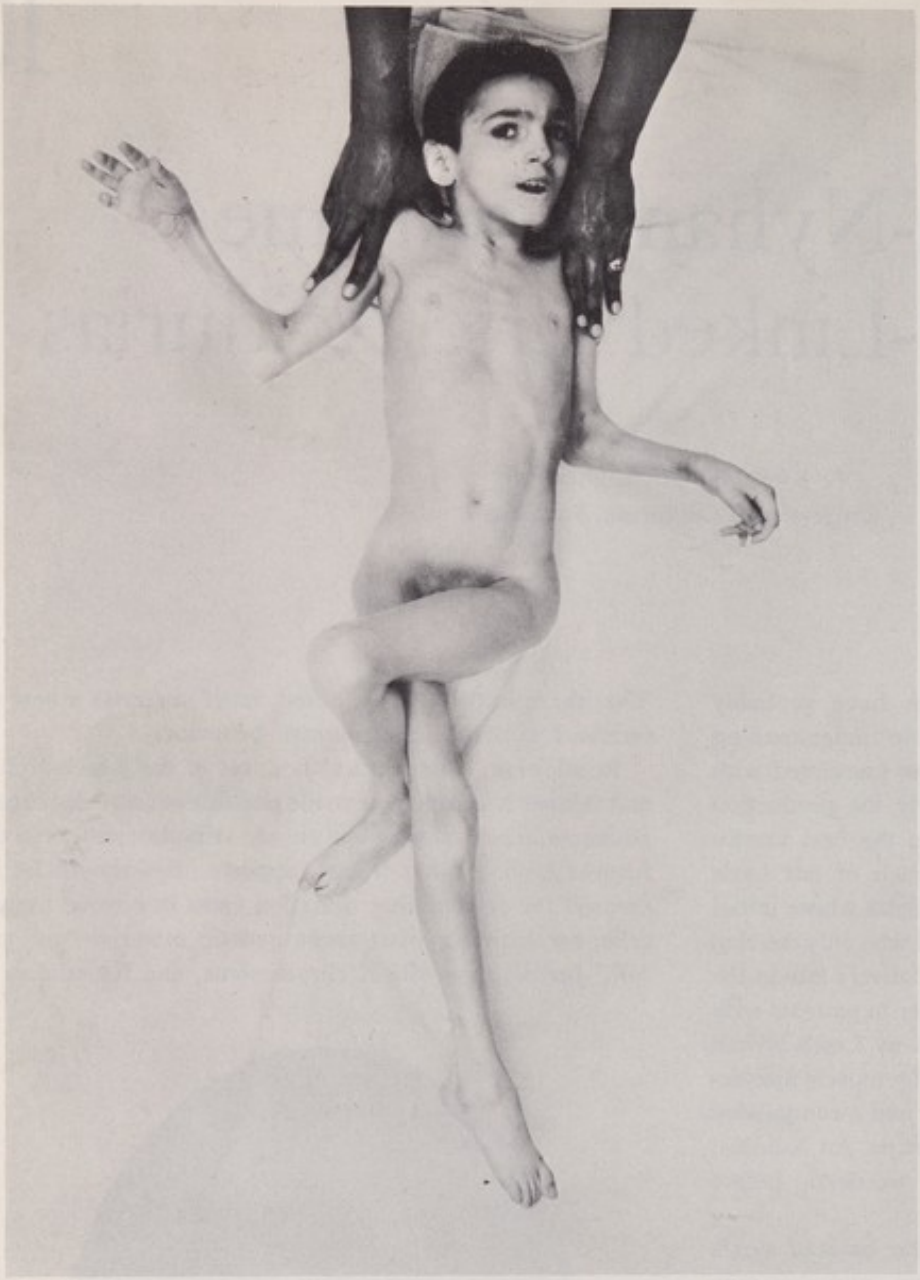
Increasing evidence suggests that we have probably reached the first landmark on the trail to understanding the group of inherited molecular diseases associated with defective purine metabolism, particularly the production of excess uric acid. Gout is, of course, the best known disease in this broad category. But much of our basic understanding so far has come from patients whose initial symptoms are principally neurologic and who only develop gouty arthritis or renal complications relatively late in the course of their disease. Specifically, I refer to patients with the rare form of cerebral palsy known as Lesch-Nyhan syndrome. The disease is characterized by muscle spasticity, choreoathetosis, mental retardation, and a compulsion to bite away the lips, tongue, and fingertips. An X-linked recessive condition, it usually leads to death before puberty.

As with the discovery of the molecular basis of sickle cell disease, the findings in Lesch-Nyhan syndrome may have broad implications for other disorders involving uric aciduria, for other polygenic diseases, and even for improved understanding of compulsive, aggressive behavior. Within three years from the description of the syndrome, the responsible enzyme defect had been identified and has since been found in other gouty patients having fewer or even no neurologic symptoms. This, then, is the first precise identification of a genetic flaw of purine metabolism involved in gout. The defect, which involves a deficiency of hypoxanthine-guanine phosphoribosyltransferase (HPRT), will be discussed in detail later. At this point I should like simply to emphasize that until the Lesch-Nyhan enzyme defect was uncovered, it was not known that HPRT played an important role in brain cell metabolism. In fact, Lesch-Nyhan is the first condition identified in which a specific biochemical abnormality can be associated with a specific pattern of aberrant behavior. How many other forms of compulsive behavior and even outright aggression have genetic and biochemical fundations? That the question can be asked, itself suggests a new direction for investigating human behavior.

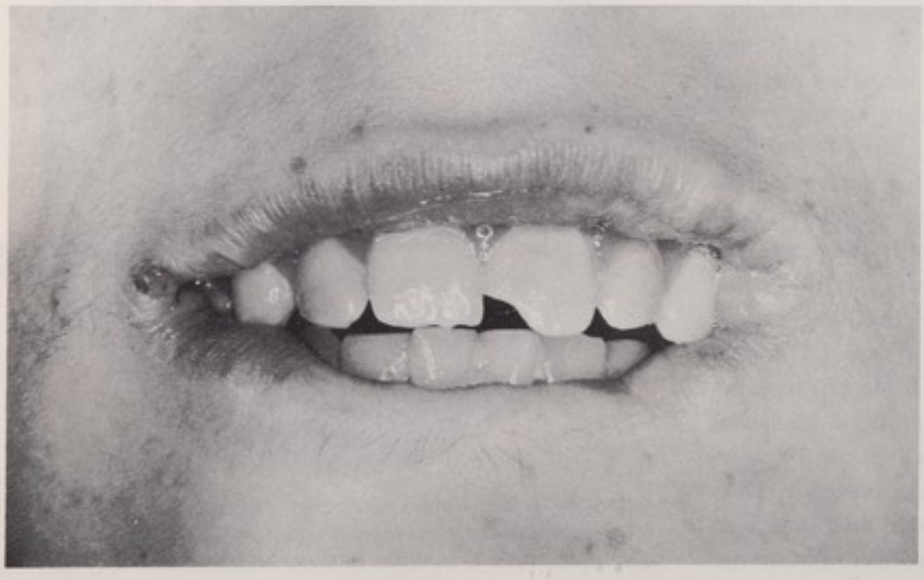
In addition, the technical features of the Lesch-Nyhan and related research have made possible several significant advances in cell biology and should stimulate others in the future. Among other things, opportunities should be increased for determining mutation rates in normal human cells, for learning what turns genetic material "on" and "off," for mapping the X chromosome, and for studies of



Although Lesch-Nyhan syndrome was identified only in 1964, its manifestations are probably centuries old. Pre-Incan ceramic dug up in Peru is thought to show characteristic self-mutilation.



The abnormal posturing and spasticity associated with Lesch-Nyhan syndrome is apparent in child shown above. His mutilation of the lips can be seen in photograph below.



genetic transformation and cell hybridization.

History

Although the term "Lesch-Nyhan syndrome" did not enter the medical literature until 1964, the unusual combination of symptoms is probably centuries if not millenia old. Dr. Nyhan has described "a pre-Inca ceramic dug up near Lima, Peru, on which is illustrated a boy who died in childhood. He had well-demarcated mutilation of lips and nose. This could be an example of the syndrome."

The contemporary story of the X-linked uric acidurias, however, began less than a decade ago. A four-year-old boy with cerebral palsy was admitted to the Johns Hopkins Hospital in 1963; his immediate problem was hematuria, which followed two days of fever and vomiting. An alert house officer observed crystals in the boy's urine, and it was first thought the child had cystinuria. Analysis showed, however, that the crystals were composed of uric acid.

Dr. Michael Lesch, who was then a medical student, and Dr. William L. Nyhan, then on the faculty at Johns Hopkins and now Chairman of the Department of Pediatrics at the University of California, San Diego, studied this boy carefully. They found similar clinical and biochemical abnormalities in his older brother. Their paper on the two children was published in April 1964 and defined the syndrome that now goes by their names. Since then more than 150 individuals with the clinical features of Lesch-Nyhan syndrome and hyperuricemia have been identified.

Both boys demonstrated the neurologic features already described. Drs. Lesch and Nyhan also found that these children excreted from three to six times as much uric acid in their daily urine as did control subjects of the same size and age. Actually, the amount per kg of body weight was two or three times the uric acid excreted by certain adults with severe clinical gout.

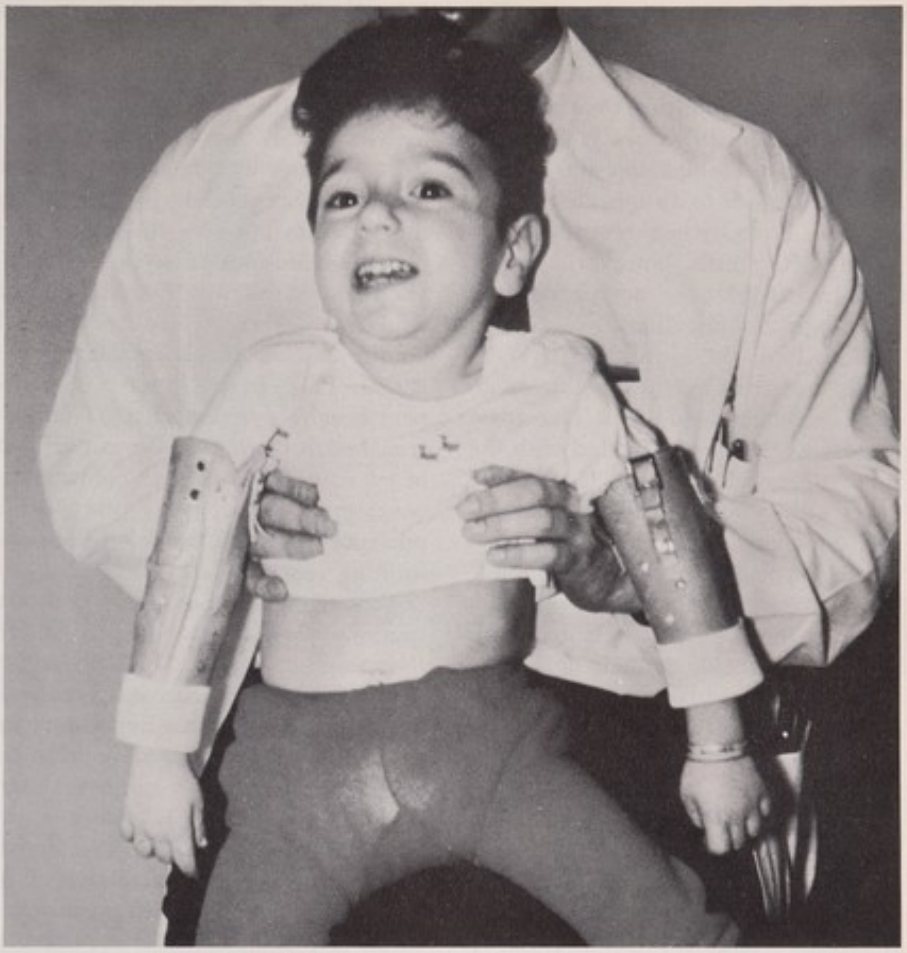
A child with this syndrome almost always appears normal at birth and for the first few months thereafter, although overproduction of uric acid is already occurring at birth. The first signs of difficulty may be observed

between the ages of six and nine months. A child's mother may notice flecks of brownish-orange crystals in the diaper, because uric acid has precipitated out of the urine. A sign of the neurologic deficit is that the child cannot hold up his head as well as a normal child. Athetosis or choreoathetosis may develop subsequently; virtually all have dysarthria.

Another early symptom of the disease is frequent vomiting. One mother who carried this genetic defect recalled that her two male children who died in infancy vomited repeatedly. She thought they were suffering from some sort of food intolerance and tried all sorts of foods without success. Many children actually have dysphagia. But, in addition, vomiting may also be related to compulsive, aggressive behavior, for the children often throw up when they become upset, frequently drenching those caring for them.

The compulsive self-mutilation may begin very early or be considerably delayed. In one of our patients, it did not begin until age 14. Two older patients examined by Dr. Nyhan had the enzyme deficiency and were extremely hyperuricemic but neither had mutilated himself. One, aged 14, had only joint aches, while the other, 27, had gouty arthritis and tophi. They also manifested less choreoathetosis than other patients with the syndrome. Thus a considerable range of expression is possible, perhaps reflecting the level of residual enzyme activity.

Although these children are mentally retarded by standard tests, one gets the impression they have considerable intelligence. Their motor handicap, particularly as far as speech is concerned, is so great that communication is difficult. We do know, however, that they are fully aware of the painful consequences of their compulsion for mutilation. They want to be protected against self-destruction and they welcome restraints such as protective mittens or having their hands tied down. We have found that splints on the elbows are useful, for they protect the hands but still allow the patient to use them. Once a patient realizes he cannot put his hands in his mouth he relaxes and becomes peaceable. If someone is taking the restraint away, on the other hand, a child may shriek in fear because he knows what he is going to do to himself.



Arm restraints are often required to protect affected child against his compulsion to self-injury, as seen above. Although standard tests usually indicate mental retardation, such children often appear lively and bright (photo courtesy of Dr. Joseph Dancis).

It appears that psychologic stress is involved in the child's behavior, too. One of these children was in our hospital at the National Institutes of Health for more than a year, and we were finally able to understand his difficult, dysarthric speech. He said that the compulsion to chew his fingers and lips came on whenever he became anxious or upset. During one period when this boy felt very secure he went for six months without splints on his elbows. When he felt he needed them, he would announce the fact. I came in one morning and found he was wearing them again. I asked him what was the matter and he said, "My mother's sick and I'm worried about her." One is sometimes tempted to think of this abnormal behavior of the Lesch-Nyhan syndrome patient as fingernail biting with the volume turned up.

These patients are also very aggressive. When they are younger they bite or pinch others, and learn to swear early. Later on they delight in kicking and hitting people, including their doctors. It is amazing that while their choreoathetosis makes so much of their motion aimless, they can still home in on the physicians' eyeglasses and send them flying across the room.

Despite this perverse streak, children with the syndrome are generally well liked by hospital personnel. Compared with the usual patients in schools or institutions for the retarded, they are bright. They apologize quickly for what they have done, and they laugh whenever something funny occurs.

Although one might expect these patients to develop symptoms of gout at an early age, it usually occurs fairly late in the disease. Some develop gouty arthritis; as a group they are also susceptible to kidney damage. We seek to prevent kidney damage and the onset of gout by administering allopurinol to block uric acid for-

mation. Since the oxypurine precursors of uric acid, hypoxanthine and xanthine, and the acid itself all have independent solubilities, the chance of developing kidney stones is reduced. In some cases, allopurinol has caused a uric acid stone to disappear from the kidney. Even when started very shortly after birth, however, allopurinol treatment does not prevent the neurologic complications.

How long these children live depends to a great extent upon the medical care available to them. The oldest patient we have encountered with the fully expressed syndrome is about 25 years old. Since both of his parents are physicians, he has had an unusual degree of medical attention. The majority eventually develop severe kidney damage, but although uremia is often contributory to their death, the immediate cause is usually a respiratory infection, sometimes complicated by aspiration pneumonia from regurgitation.

In summary, male children may be suspected of having the Lesch-Nyhan syndrome if they vomit frequently during the neonatal period, have brownish-orange crystals in their urine or diapers, manifest hematuria or evidence of kidney stones, demonstrate choreoathetosis with or without mental retardation, and are compulsively aggressive or self-destructive, particularly in biting their lips, tongue, or fingers.

Biochemical Studies

I first began investigating the biochemical aspects of Lesch-Nyhan syndrome while I was in the Section on Human Biochemical Genetics at the National Institute of Arthritis and Metabolic Diseases and have continued these studies at the University of California, San Diego. It seemed to me that if we could understand the primary genetic abnormality in this form of cerebral palsy, we might be able to find related defects in the regulatory system for purine metabolism and thereby account for the excessive rate of purine metabolism found in many patients with gout. In the subsequent investigations, important contributions were made by my coworkers, Drs. William N. Kelley, Frederick M. Rosenbloom, Frank Henderson, Wilfred Y. Fujimoto, Martin L. Greene, Theodore Friedmann, and J. A. Boyle.

Previous studies of both bacterial and mammalian cells had outlined the pathways for purine synthesis, together with the feedback control mechanism regulating synthesis. As the diagram opposite illustrates, the purine nucleotides guanylic and adenylic acid and the corresponding di- and triphosphates exert feedback control on the first rate-limiting enzyme reaction. This enzyme, PRPP (5-phosphoribosyl-1-pyrophosphate) amidotransferase, catalyzes the conversion of PRPP and glutamine to phosphoribosylamine. In a sense the purine nucleotides keep a brake on the rate-limiting reaction.

Our original hypothesis was that a mutation had altered the site where the inhibitors attached to the enzyme and made it less sensitive to feedback inhibition. To test this hypothesis we studied the effect of azathioprine on our patients' purine synthesis. Dr. Leif Sorenson and his associates at the University of Chicago had shown that azathioprine, a derivative of 6-mercaptopurine, inhibited purine synthesis in gouty patients who produced excessive uric acid but not in two brothers who had some of the clinical features of Lesch-Nyhan syndrome. We confirmed Dr. Sorenson's findings in two patients with the full-fledged syndrome.

For in vitro studies of purine metabolism we then established cultures of fibroblasts taken from the skin of our patients. De novo purine synthesis was then assessed by blocking metabolism two steps beyond the rate-limiting reaction. This was done with the glutamine analogue azaserine, which results in accumulation of the purine intermediate formylglycinamide ribonucleotide (FGAR). We found that Lesch-Nyhan cells incorporated formate-^{14}C into FGAR at a rate five to six times higher than normal cells, representing an excessive rate of purine synthesis that was quite comparable to the degree of overproduction of uric acid found in patients' 24-hour urine collection.

Up to this point, findings were consistent with the hypothesis that the purine nucleotide inhibitors of purine synthesis were unable to attach effectively to the rate-limiting enzyme PRPP amidotransferase. But then we used another derivative of 6-mercaptopurine, 6-methylmercaptopurine ribonucleoside, and found it did in fact inhibit purine synthesis in both normal and patient cells. This was the first indication that the children had a defect other than the one we had postulated.

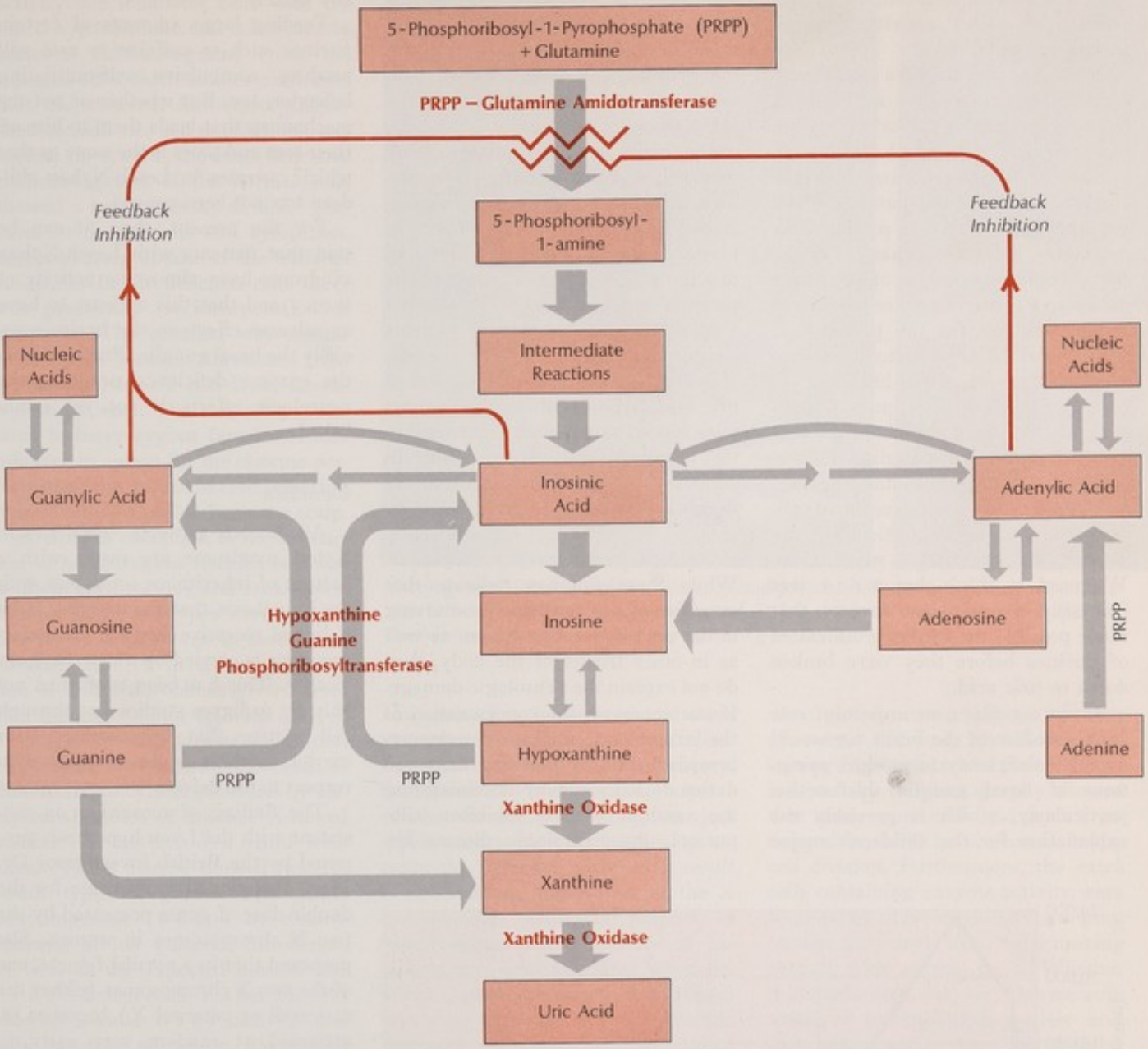
Study of the pathways involved in the metabolism of these two different purines, 6-mercaptopurine and 6-methylmercaptopurine ribonucleoside, subsequently pinpointed the biochemical lesion responsible for Lesch-Nyhan syndrome. The first of these compounds needs hypoxanthine-guanine phosphoribosyltransferase, which converts hypoxanthine or guanine to ribonucleotide form. The other, however, requires a kinase that normally converts adenosine to a nucleotide. This suggested that 6-mercaptopurine's failure to inhibit purine synthesis could have been the result of an HPRT deficiency in the patients' cells.

We accordingly assayed our cultured cells for HPRT activity and found that patients' cells were markedly deficient in the enzyme, which was easily detected in normal fibroblasts. To verify these findings, erythrocytes were also studied. HPRT could not be detected in patients' red cell lysates when either hypoxanthine or guanine served as a substrate but was fully detectable in normal red cells.

The typical patient with the full-blown syndrome has less than 0.005% of normal HPRT activity in his red blood cells and about 2% to 3% in skin fibroblasts. Fibroblasts in affected patients typically show more HPRT activity than their erythrocytes. We believe this is because the mutant enzyme is somewhat unstable and so is gradually inactivated at body temperature and therefore loses activity in the longer-lived red cells. (Fibroblasts replicate every 24 hours; erythrocytes have an average life-span of 120 days.)

If one adds hypoxanthine or guanine to cultures of normal fibroblasts, purine synthesis is inhibited. But if either hypoxanthine or guanine is added to the medium of cells taken from Lesch-Nyhan patients, purine synthesis is stimulated. At present we do not know why the latter should be true, but the cell cultures respond very

Feedback Control of Purine Biosynthesis



As indicated in the diagram above, the purine nucleotides and their polyphosphates normally serve to inhibit 5-phosphoribosyl-1-pyrophosphate (PRPP) glutamine amidotransferase, mediator of the first and presumed rate-limiting reaction controlling purine biosynthesis. The enzyme deficient in children with Lesch-Nyhan syndrome is hypoxanthine-guanine phosphoribosyltransferase (HPRT), which shares the same substrate, PRPP, with the rate-limiting enzyme. Increased PRPP concentrations are found in the cells of children with the disease and appear to contribute to the excessive purine synthesis.

much as patients do in this respect.

The precise way by which the enzyme defect causes the excessive purine synthesis has been studied. We now see that the shortage of HPRT leads to an excess of PRPP, presumably because it cannot be used in the HPRT-catalyzed reaction. PRPP is normally one of the substrates for the rate-limiting enzyme for purine metabolism, PRPP-glutamine amidotransferase. When present in excess it probably causes the rate of metabolism to accelerate. We have found a demonstrable excess of PRPP in Lesch-Nyhan fibroblasts (four times normal) and red cells (10 to 40 times normal).

It is also significant that virtually all children with Lesch-Nyhan syndrome have increased activity of the enzyme adenine phosphoribosyltransferase (APRT) in their red cells. This curious increase superficially resembles a type of controller gene mutation described in bacteria by Monod, in which a decrease in one enzyme is accompanied by a coordinate derepression of another enzyme, but a far simpler explanation was found. Our evidence suggests that the high intracellular concentration of PRPP in these children's cells stabilizes APRT against the gradual spontaneous inactivation of the enzyme that is a natural accompaniment of aging of nor-

mal red cells. As a result, the mutant red cells show a substantially higher amount of APRT activity than do normal red cells.

Now how does the HPRT deficiency seem to be related to the self-destructive and aggressive behavior so commonly observed in the Lesch-Nyhan syndrome? The mechanism responsible for the neurologic damage is not yet known. It does not produce any detectable structural change in the microscopic appearance of the brain at autopsy. But we have examined autopsy tissues for the presence of HPRT and found it virtually absent from all regions of the brains of affected children, while in normal brain it is actually most abundant, particularly in the basal ganglia. This is somewhat of a surprise, since the reaction the enzyme catalyzes had not been thought of as particularly important for the body's metabolism. We used to think that HPRT was primarily a scavenger enzyme that made possible the thrifty reutilization of purines before they were broken down to uric acid.

If HPRT plays an important role in metabolism of the brain, we would expect a deficiency to produce symptoms of basal ganglia dysfunction particularly, which is precisely the explanation for the children's major symptoms of spasticity and choreoathetosis. The evidence suggests that the local biochemical consequences of the enzyme defect are responsible for the neurologic damage, rather than some substance that gets into the brain secondarily.

Originally it was tempting to think that uric acid was causing the neurologic damage since it is largely excluded from the cerebrospinal fluid in normal subjects. However, we found that the uric acid concentration in our patients' cerebrospinal fluid was not very different from that in patients having spinal taps for other reasons. However, the oxypurine precursors of uric acid, hypoxanthine and xanthine, were indeed substantially increased in the patients' cerebrospinal fluid. In addition, the CSF concentration of these oxypurines was higher than their serum concentration, while in nonaffected subjects the reverse was true. While these findings indicate that excessive purine synthesis is occurring in the central nervous system as well as in other tissues of the body, they do not explain the neurologic damage. If one increases the concentration of the oxypurines in plasma and cerebrospinal fluid by blocking their oxidation to uric acid by administering the xanthine oxidase inhibitor allopurinol, the neurologic disease becomes no worse, nor does it become any better.

Feeding large amounts of certain purines such as caffeine to rats will produce compulsive self-mutilating behavior, too. But whether or not the mechanism that leads them to bite off their toes and paws is the same as that which operates in Lesch-Nyhan children has not been proved.

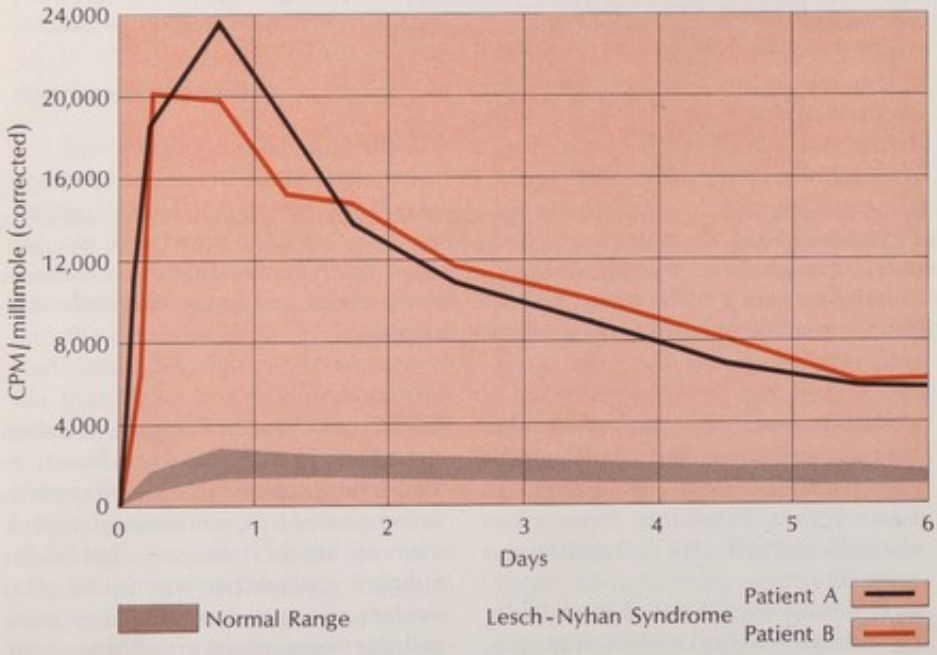
For the present, then, it can be said that patients with Lesch-Nyhan syndrome have almost no activity of HPRT and that this appears to have an adverse effect on the brain, especially the basal ganglia. Precisely how the enzyme deficiency produces the neurologic effects is not yet established.

Genetics

All known patients with Lesch-Nyhan syndrome are males with a pattern of inheritance consistent with the hypothesis that the disorder is an X-linked recessive one and the defective gene is carried by a heterozygous mother. This has been confirmed not only by pedigree studies but through cell cultures that demonstrated such carrier mothers to be mosaics with respect to the defect.

The finding of mosaicism is consistent with the Lyon hypothesis proposed by the British investigator Dr. Mary F. Lyon to compensate for the double dose of genes presented by the two X chromosomes in women. She proposed that in a normal female, one of the two X chromosomes (either the maternal or paternal X) becomes inactivated at random very early in embryologic development. Whatever "choice" is made, the cell descendents of these embryonic cells will all carry the same active X chromosome as their "progenitors." A carrier mother could thus be expected to have two cell populations so far as the X chromosome is concerned. In one of these cell populations the X chromosome carrying the defective gene for HPRT would be active and would manifest deficient HPRT activity. Cells in which the defective X chromosome was inactivated would have normal HPRT activity. Fortunately, it is possible to assay single cells for activity of the enzyme.

HPRT is the first enzyme involved in the incorporation of hypoxanthine



Incorporation of labeled glycine (a precursor of purine synthesis) into urinary uric acid was found to be much higher in two Lesch-Nyhan patients than in four normal subjects, demonstrating that accelerated purine synthesis is responsible for patients' uric aciduria (graph adapted from W. N. Kelley et al, Ann Intern Med *73:583, 1969).*

into nucleic acids. To assay for its activity one incubates cells with tritiated hypoxanthine and subsequently prepares radioautograms from the cultured cells. Such studies confirmed that the maternal fibroblasts did indeed contain two cell populations, one normal for HPRT, the other deficient. Radioautograms of the normal cells showed a heavy overlay of silver particles in the nucleus and cytoplasm, demonstrating uptake of the tritiated substrate. In the deficient cells there was practically no incorporation of substrate. The contrast was similar to that seen between fibroblasts from normal individuals and from Lesch-Nyhan patients. The two cell populations in heterozygous females is also reflected in assays for the enzyme performed on the groups of cells adhering to hair follicles pulled from the scalp.

A curious exception to the mosaicism in heterozygous females is found in their erythrocytes and leukocytes. Both of these cell types show normal activity of HPRT even in mothers of affected children in whom examination of either fibroblasts grown from skin biopsies or hair follicles assayed directly show both normal and mutant cells. Some insight into the cause of this discrepancy was provided by Dr. Nyhan's studies of another X-linked marker. A mother of a child with Lesch-Nyhan disease showed two different types of glucose-6-phosphate dehydrogenase (G-6-PD) by electrophoresis of her cultured fibroblasts, as well as the expected normal and HPRT-deficient cells. Yet her red cells showed normal HPRT activity and only one type of G-6-PD. Since both of these genes are X-linked, this result suggests that the HPRT-deficient precursor cells for the hematopoietic system were at a disadvantage compared to normal cells and did not survive the early fetal development. As expected, this selective pressure does not seem to operate in families with less severe degrees of enzyme deficiency where heterozygous females can show a diminished enzyme activity in erythrocytes.

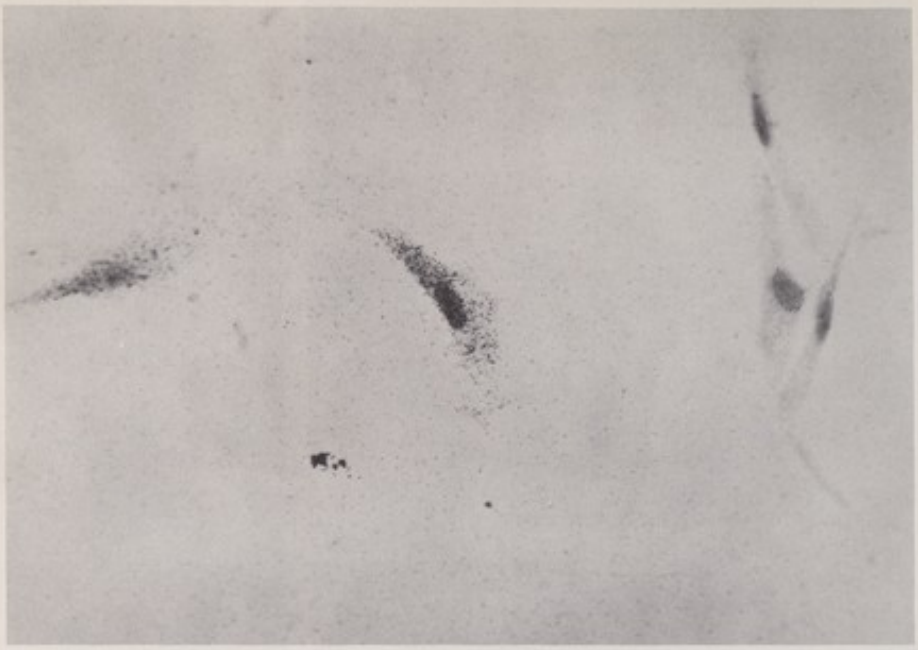
In theory, if the defective gene represents a mutation occurring de novo in the oocyte, a mother might not be a mosaic, but we have not yet found such a case among the mothers of the 12 patients we have tested. Dr. Barbara Migeon of Johns Hopkins University has, however, found some examples in families she has studied. It might even be possible, in theory, for a female to manifest some aspects of the disease, depending on just which embryonic cells happen to express the defective gene as a result of the random inactivation of the X chromosomes. Indeed the majority of heterozygotes show evidence of excessive uric acid synthesis when detailed studies are made, even though they may not show a hyperuricemia. Although one mother we have studied is mentally retarded, we cannot be certain that it is a consequence of her heterozygous state for the HPRT-deficient gene. Obviously a detailed neurologic evaluation and intelligence testing of heterozygotes is needed to answer this question.

The expression of the abnormal gene in the heterozygote is further attenuated by a newly recognized type of interaction between the normal and HPRT-deficient cells designated "metabolic cooperation," which leads to a partial correction of the enzyme defect. Studies in vitro in our laboratory, performed by Dr. John Subak-Sharpe of Glasgow and Dr. Theodore Friedmann, in which a small number of normal fibroblasts were plated among a large number of HPRT-deficient fibroblasts containing ingested carbon particles and followed by radioautography showed that many more cells had HPRT activity than could be accounted for by cell division. Furthermore, the extra cells containing enzyme activity were those HPRT-deficient cells (containing carbon particles) that were making contact with normal cells. Whether it is preformed enzyme, enzyme subunits, or informational nucleic acid that passes across these intracellular bridges remains to be demonstrated.

Theoretically, an affected female could be produced but it would be a rare occurrence. Although male patients usually die before attaining reproductive age, some males without the full clinical manifestations of the syndrome have been found. If a male with this genotype married a heterozygous female, half their female offspring could be homozygous for the defect.

The defect in Lesch-Nyhan syndrome does not appear to result from a gene deletion. Patients showing no detectable HPRT enzyme in their erythrocytes nevertheless show 2% to 3% normal activity in their cultured fibroblasts. Furthermore, recent stud-

Feeding large doses of caffeine, a methylated purine, produces compulsive self-mutilation in rats; mutilation of paws is typical. It has not been proved, however, that the mechanism is the same as the one operating in Lesch-Nyhan children.



Evidence to support the Lyon hypothesis – that in females one of the X chromosomes is inactivated at random, resulting in mosaicism with respect to X-linked genes – has come from radioautographic studies of skin fibroblast cultures established from mothers heterozygous for HPRT deficiency. In radioautogram above, black granules overlying some but not all of the cells demonstrate activity of HPRT, which is required for incorporation of tritiated hypoxanthine into nucleic acid. This photograph and the one on page 111 are reproduced from W. Y. Fujimoto et al, Lancet *2:512, 1968.*

ies by Drs. Charles S. Rubin, Joseph Dancis, Lily C. Yip, Robert C. Nowinski, and M. Earl Balis at the Sloan-Kettering Division of Cornell University Graduate School of Medical Science and the New York University School of Medicine are compatible with the prevailing view that a mutation in gene structure is involved. These investigators purified HPRT, prepared antisera against it in rabbits and rats, and then exposed erythrocyte lysates from five patients with Lesch-Nyhan syndrome to the antisera. The lysates blocked inactivation of normal enzyme by the antibody. The cells apparently contained essentially normal amounts of a protein immunologically identifiable as HPRT, but lacked catalytic activity. The researchers reasoned that such a change is probably not due to a mutation in a regulator gene or to a gene deletion, but results from some change in a structural gene.

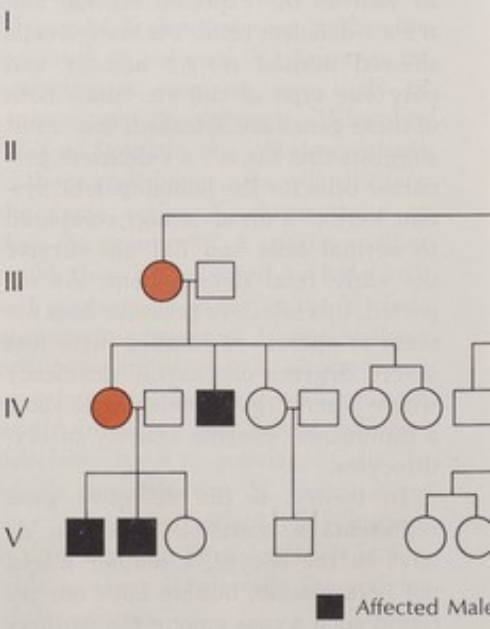
In discussing the genetics of this syndrome, one naturally also thinks of gout. For centuries the tendency of gout to run in families has been recognized. Furthermore, it is predominantly a disease of males, presumably because the serum urate of women averages around 1 mg/100 ml lower than that of men, so that fewer women develop the hyperuricemia required for deposition of the crystals of monosodium urate in and about the joints and kidney responsible for the clinical symptoms of gout. Although only about 5% of the cases occur in women, they are usually not seen until after the menopause when their serum uric acid rises and approaches that normally found in males. Sir Archibald Garrod, who first recognized the importance of inborn errors of metabolism, placed gout among these disorders as far back as 1931.

Unlike many other inborn errors of metabolism, no single enzyme defect has been found so far to be responsible for gout. In the 1940's, there were reports that hyperuricemia was caused by a single dominant gene with incomplete penetrance. But subsequent studies, utilizing more precise methods for determining uric acid concentrations, demonstrated there was no single mechanism that separated hyperuricemics from normal subjects. In fact, it has been shown that while some gouty patients have excessive uric acid production, others appear to have normal production, so far as we can tell.

Today the view is accepted that many different genetic and environmental factors probably contribute to the hyperuricemia that eventually produces gout. It has also been learned that a considerable number of other relatively common hereditary human diseases represent the interaction of polygenic inheritance and environmental factors.

As mentioned earlier, we originally hoped that Lesch-Nyhan syndrome would give us some insight into a defect in purine metabolism that might explain the problem encountered in some cases of gout. After identifying the HPRT defect in Lesch-Nyhan patients, we tested many of our gouty patients for the same defect. We have found a small number with the deficiency, from whom it has been possible to learn a great deal.

Thus, when deficiency in HPRT activity in red cells is on the order of 0.1% to 17% of normal, a broad range of clinical difficulties can be identified. These patients all synthesize excess uric acid roughly in inverse proportion to the amount of active enzyme available. They appear to have a much more malignant form of gout than do the usual gouty patients. Arthritis develops at an earlier



This large kindred in which Lesch-Nyhan syndrome was identified in three out of

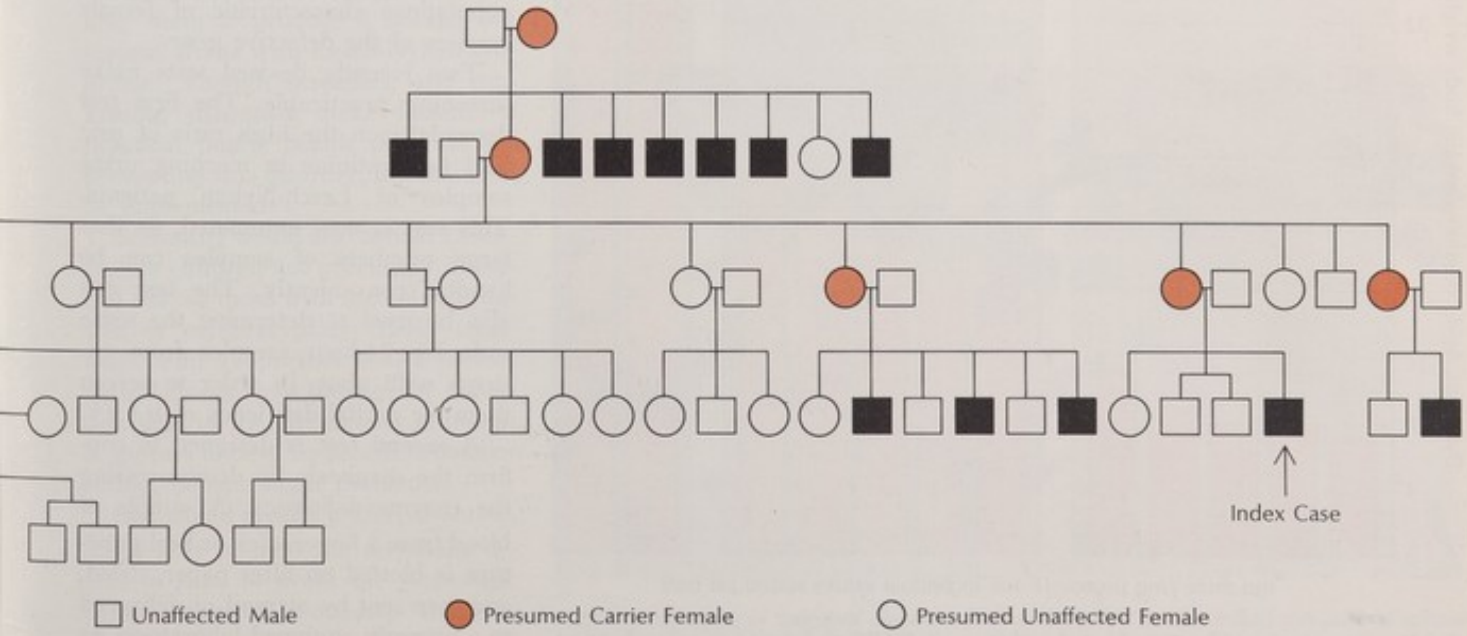
age; they have more renal calculi and more renal damage.

The exact incidence of the HPRT defect among gouty patients has not been determined. Only five such men, three of whom were in the same family, were discovered to have it among 110 gouty patients who entered the Clinical Center at NIH for detailed metabolic studies over a period of 14 years. Dr. Alexander Gutman of Mount Sinai School of Medicine in New York City has found only seven among 425 patients that he has been following, with five of them in a single family. All told, we know of only 28 hyperuricemic patients with the less severe HPRT defect. The clinical expression of the neurologic defect is greatly attenuated by even small quantities of residual enzyme. Thus patients who show more than 1.0% normal activity in their red cells have no neurologic dysfunction, while the eight patients with 0.01% to 0.5% of normal activity showed a wide spectrum of neurologic dysfunction, but the presenting symptoms were so variable that one could not possibly have identified the patients as having the same disorder simply on clinical evidence. But these 28 individuals all have a similar biochemical deficiency, the same one found in Lesch-Nyhan syndrome. A change in terminology may be in order; possibly the whole group of diseases should be called X-linked uric aciduria, with the Lesch-Nyhan patients viewed as having the most severe form.

Evidence has been found for many different types of mutations in the HPRT locus. Four different families illustrate the situation. The three J brothers, all with gouty arthritis and tophi (two also had renal calculi), had erythrocytes demonstrating about 1% of normal HPRT activity when either hypoxanthine or guanine was used as a substrate; their enzyme was less stable than normal when heated. None had any neurologic complications. The two L brothers had erythrocytes with HPRT enzyme activity of about 0.5% of normal with the substrate guanine, and of approximately 10% of normal with hypoxanthine; their enzyme proved to be more stable than normal when heated. One brother had gout, tophi, and renal calculi; the younger brother had a history of uric acid renal calculi; but both had neurologic symptoms classed as a spinocerebellar syndrome. One had mild spastic quadriplegia, the other was mentally retarded. In the third family, Mr. S had gout, tophi, and calculi. His erythrocytes had about 10% activity with either substrate, but the enzyme was more stable than the J family's upon heating, though not as stable as the normal form. In still another family, studied by Drs. J. A. McDonald and W. N. Kelley at Duke University, activity could be restored in vitro by use of a substantially higher concentration of PRPP in the assay system, indicating a decreased affinity of the mutant enzyme for PRPP.

We cannot explain all these differences in molecular terms yet, but we presume the amino acid sequence of the same HPRT enzyme molecule has been coded differently by each of the mutations. This difference in amino acid sequence, most likely at certain sites in the enzyme of affected members of each of the families, probably leads in turn to differences in tertiary structure, enzyme activity, and stability of each family's HPRT protein.

The X-linked uric acidurias can be thought of as representing only the tip of the gout iceberg. Undoubtedly there are many other defects in purine metabolism among the gouty patients. For example, two other gouty patients with overproduction of purines appear to have the sort of biochemical defect we originally postulated. Their fibroblasts did not respond normally in culture when exposed to substances that inhibit purine synthesis through the feedback mechanism. One of several working hypotheses is that their mutation produced a change in the structure of PRPP amidotransferase, the rate-limiting enzyme, to prevent the inhibitors of purine synthesis from



five generations was reported by Dr. Nyhan and coworkers in 1966. Note consistency of inheritance pattern in the third generation with that expected of an X-linked recessive: transmission through the heterozygous mother to (approximately) half her sons.

binding or in other ways to decrease the enzyme activity.

Genetic defects other than those involving purine metabolism also may initiate the hyperuricemia that is the final common pathologic mechanism in gout. Patients with glycogen storage disease type 1, with deficient glucose-6-phosphatase in the liver, who live to adulthood, acquire gout and hyperlipidemia with severe cardiovascular disease as their major clinical problems. Hence this is an example of at least one defect in carbohydrate metabolism that can bring about overproduction of purines. We have also identified one case in which hyperuricemia accompanied a primary defect in amino acid metabolism. Cells of this mentally retarded infant girl with branched-chain ketoacidemia (maple syrup urine disease) had from 15% to 25% of normal decarboxylase activity for the branched-chain keto acids derived from leucine, isoleucine, and valine. She had not only elevated keto acids but also hyperuricemia. Yet when the keto aciduria was controlled by a semisynthetic diet containing minimal requirements of these branched-chain amino acids, her serum uric acid also declined to the normal range and rose when each of the branched-chain amino acids was added to the diet.

And, of course, other as yet unidentified factors appear to play some role in the expression of gouty arthritis as well as in other clinical consequences of elevated uric acid. Dr. Arthur Hall investigated the occurrence of gout among the more than 5,000 people studied in the Framingham Heart Disease Epidemiology Study. He found that 83% of those whose serum uric acid exceeded 9 mg/100 ml at the start of the study had gout 12 years later. But if the level was between 8 and 9 mg, only 25% got the disease; if it was between 7 and 8 mg, only 15% did. So while the severity and duration of hyperuricemia influence the situation, not all individuals who are defined as hyperuricemic develop gout, and we do not know what other factors are involved in this attenuation.

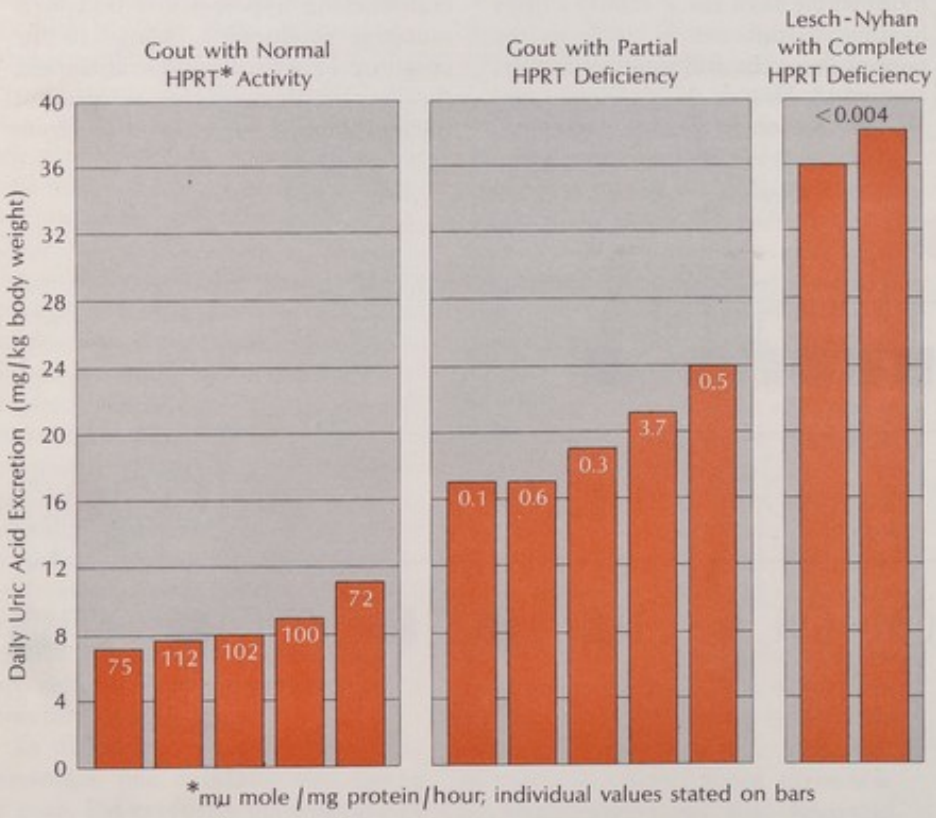
With amniocentesis as the starting point, antenatal diagnosis has paved the way to prevention of Lesch-Nyhan syndrome. Some amniotic fluid cells are examined for heterochromatin bodies to establish the child's sex. Other cells from the fluid are cultured and incubated with tritiated hypoxanthine. Radioautography will reveal the HPRT activity of the cells and identify which HPRT genotype is present in the fetus: normal male or female, female carrier, or affected male.

Ideally, amniocentesis and culturing of cells should only be attempted when the expectant mother is a known carrier of the HPRT defect. Her enzyme picture can be established by examining her fibroblasts in cell culture or by analyzing her hair follicles for the enzyme. Prenatal diagnosis under these circumstances makes possible a preventive program involving selective abortion of proved affected fetuses. The amniocentesis is performed by an obstetrician experienced in this procedure at 16 to 18 weeks of gestation, and the culturing and assay of the amniotic cells require an additional three to four weeks.

So far as screening is concerned, there is no justification for a massive program to identify potential carriers of defective HPRT genes in the population at large, but identification of families at risk would be very worthwhile. This could start with screening of populations from mental hospitals, cerebral palsy clinics, and schools for the handicapped. Once affected patients have been located, their mothers and female blood relatives should have skin biopsies and examination of their cultured fibroblasts for the two cell populations characteristic of female carriers of the defective gene.

Two recently devised tests make screening practicable. The first test depends upon the high ratio of uric acid to creatinine in morning urine samples of Lesch-Nyhan patients. This test is now automated, so that large numbers of samples can be handled conveniently. The test can also be used to determine the same ratio for 24-hour samples from patients with gout, in order to screen them for partial deficiency of HPRT.

A second test is designed to confirm the diagnosis by demonstrating the enzyme deficiency. A sample of blood from a finger stick or heel puncture is blotted on filter paper, dried, and then sent by airmail or delivered to a properly equipped laboratory; no refrigeration is required. The laboratory measures the incorporation of



Studies in patients with gout showed a degree of HPRT deficiency in some but not others, and a positive correlation between the extent of HPRT deficiency and the daily urinary uric acid level. Data are from Kelley et al., Ann Intern Med *70:178, 1969.*

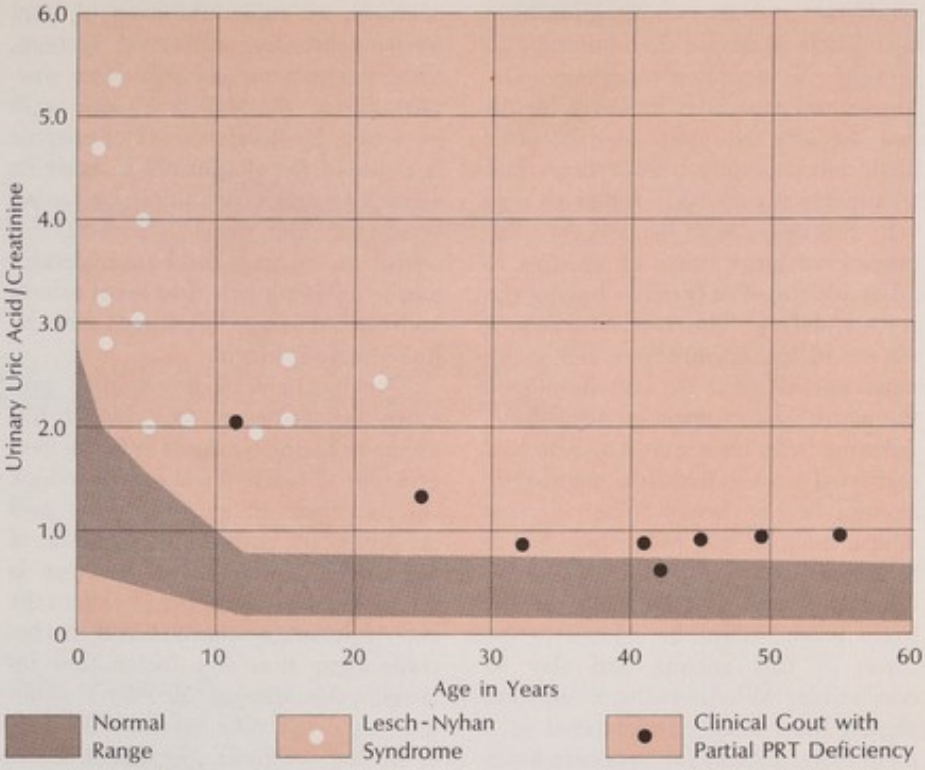
hypoxanthine-8-^{14}C into inosinic acid as an indication of the HPRT activity. Blood from affected children shows virtually no incorporation.

Together these two procedures permit the early identification of affected males. However, the women and girls who may later bear HPRT-defective sons can be identified only by the more cumbersome method of skin biopsy for culture of fibroblasts, and subsequent radioautography, or by biochemical tests of the roots of hairs freshly removed from their scalps. Both of these tests for carriers are technically difficult and their reliability and accuracy for finding a large group of potential carriers remains to be determined. In addition, the random nature of the inactivation of the X chromosome raises the theoretical possibility that no HPRT-deficient cells may be present in the particular site selected for skin biopsy or hair-follicle examination.

These considerations suggest that amniocentesis and prenatal diagnosis should be performed on all female blood relatives of an affected child, at least until sufficient experience has accrued for statistical evaluation of the chances for error in identification of carriers of this disease. These females may already be mothers of retarded or cerebral-palsied children, or may be siblings or even more distant blood relatives of such children. Either way, genetic counseling, amniocentesis, and prenatal diagnosis could help them to produce a normal son instead of one with Lesch-Nyhan syndrome. Though screening and confirming diagnosis make feasible a practical public health program for preventing this disorder, it remains to be implemented.

Screening would also benefit Lesch-Nyhan children not previously identified among those with mental retardation or cerebral palsy. Early treatment with allopurinol would protect their kidneys from damage by excess uric acid in the circulation, although it appears to be of no benefit in ameliorating the neurologic problems.

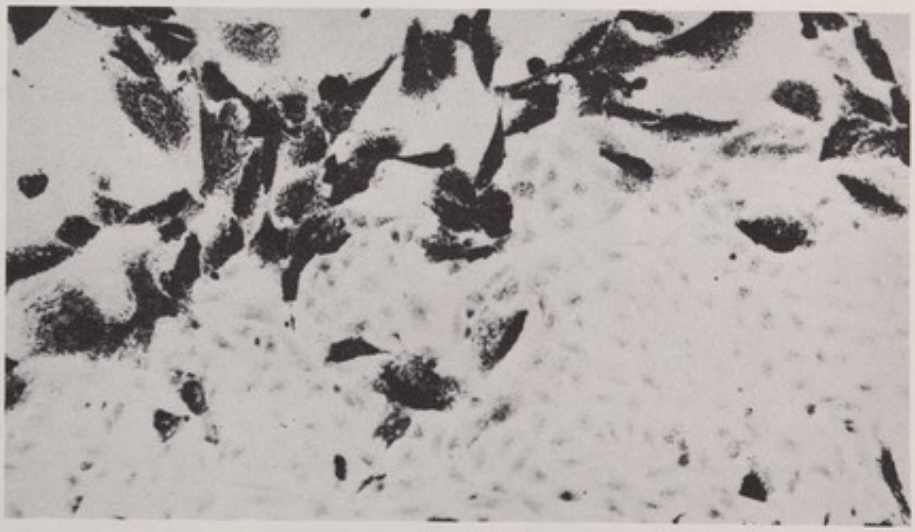
One would greatly prefer, of course, to be able to treat the metabolic defect in time to avoid the neurologic complications. Can one hope that prevention by selective abortion will eventually prove to be only a stopgap arrangement? One need only recall



The high uric acid to creatinine ratio observed in morning urine of Lesch-Nyhan patients and in some gouty patients with partial HPRT deficiency provides the basis for a useful screening test (adapted from J. M. Kaufman et al, J Pediat *73:583, 1968).*

the example of maternal-child Rh incompatibility to appreciate that this may come about. Only a very few years ago prevention of the Rh disorder was a complex procedure, required repeated amniocentesis to check on the course of the fetus, and exchange transfusion at birth (sometimes with transfusions even before birth). But now this disease, erythroblastosis fetalis, can be prevented by a simple injection of Rh antibodies given to the Rh-negative mother at the time of delivery of each of her children (see Clarke, Chapter 26).

Although we are not even at the halfway mark in treatment of the neurologic problem of Lesch-Nyhan



Radioautography of cultured amniotic fluid cells (taken at 17 weeks from patient whose skin fibroblast culture appears on page 108) showed two cell populations, with and without HPRT activity (as indicated by incorporation of ^{3}H-hypoxanthine), thus demonstrating that the fetus was a female and heterozygous for HPRT deficiency.

syndrome, at least two attempts have been made to devise chemotherapy to prevent the neurologic damage. Dr. Marks and associates in Texas identified an affected child shortly after birth and promptly began allopurinol treatment; however, it failed to prevent the neurologic disease. We administered large doses of adenine to a boy identified at birth as having the HPRT defect. The concentration of PRPP in his erythrocytes fell to almost normal, but he still developed the neurologic syndrome on schedule. Adenine was chosen because it had corrected a megaloblastic macrocytic anemia in one Lesch-Nyhan patient whose anemia had not responded to folic acid or vitamin B_{12}. We postulated that perhaps the cause of the brain lesion might be related to the cause of this anemia and also be correctible. Whether the compound should have been administered in a form more suitable for uptake in brain cells, or whether adenine is simply the wrong agent for therapy, is not clear, but various biochemical approaches will continue to be tried.

Research into the X-linked uric acidurias is also clarifying certain puzzling differences in response to the drug, allopurinol, used in the treatment of gout. Allopurinol acts in two ways. Its primary effect is to inhibit xanthine oxidase, so that it lowers the uric acid levels in serum and urine in all patients and increases the concentration of the oxypurine precursors of uric acid, hypoxanthine and xanthine. The second effect of allopurinol is to inhibit purine synthesis de novo. However, in patients with X-linked uric aciduria, no such inhibition of total purine synthesis is observed. Instead, these patients merely substitute oxypurines for the deficit in uric acid excretion. Evidently the HPRT enzyme is required for allopurinol to exert its secondary action of inhibiting purine synthesis. The drug is nevertheless useful in treating the hyperuricemia and in avoiding uric acid renal calculi and renal damage in patients with X-linked uric aciduria.

The biochemical information provided by studies of this disease has given us further insights into the general type of biochemical lesions to look for as causes of excessive uric acid synthesis in gout. The position of HPRT in the metabolic pathway is not one that would appear to affect the rate of purine synthesis, but it has become clear that any factor that increases the amount of PRPP available to a cell may increase the rate of purine synthesis. So this takes us back one step further at the biochemical level.

Cell biology has already made countless contributions to medical science and, conversely, the study of the human inborn errors of metabolism has made contributions to cell biology. This interaction will undoubtedly continue. Thus, apart from demonstrating the value of skin fibroblast cultures in the study of metabolism, the Lesch-Nyhan research has shed light on a new type of cell-to-cell interaction mentioned above, the process known as metabolic cooperation.

Studies using these mutant human cells have also opened up additional possibilities in cell biology for consideration. Can we some day repair enzyme defects by manipulating human cell lines? At least we know this can be accomplished in cell culture. Hybridization of one human cell strain deficient in G-6-PD and another deficient in HPRT produced a cell strain in which both defects had been repaired. This was carried out in a collaborative study with Drs. Wilfred Fujimoto, Marcello Siniscalco, Hilary Koprowski, Harry Eagle, and Harold Klinger.

Another group of workers corrected the HPRT defect in mouse cells by cell fusion with a chicken erythrocyte. This suggests that we are at the threshold of exploring genetic combinations in human cells that could not be achieved by natural biologic matings. Such combinations should add to our knowledge of human genetics.

One hesitates to be too optimistic about the clinical potentials of basic research. Missed opportunities and unforeseen obstacles can indefinitely delay or simply prevent future application of today's findings. But the rapid advance in dealing with Lesch-Nyhan syndrome since its recognition less than a decade ago suggests that a powerful tool may have become available for studying the genetic and biochemical bases of compulsive and aggressive behavior. Such behavior may have served our primitive ancestors well in the long struggle for survival, but today it threatens our species with extinction. We could do far worse than to apply the best talents and insights of medicine to this problem.

11

The Porphyric Diseases of Man

RICHARD D. LEVERE *and* ATTALLAH KAPPAS
State University of New York and Rockefeller University

The striking wine-red urine of some porphyric patients may partly explain why these diseases enjoy a reputation so disproportionate to their clinical frequency. There are, however, reasons more important than the chromatic characteristics of urine for paying special attention to this group of disorders.

The porphyrins are intermediates along a biochemical pathway whose principal product, heme, is vital to the respiratory activity of cells and, as recently shown, also essential for the organism's interactions with its environment. One of heme's functions, for example, is to participate in the mechanisms by which the body purges itself of toxic agents. Thus, what may have begun as an intellectual curiosity about a striking but uncommon group of genetic disorders has evolved into several lines of investigation having broad relevance to human physiology.

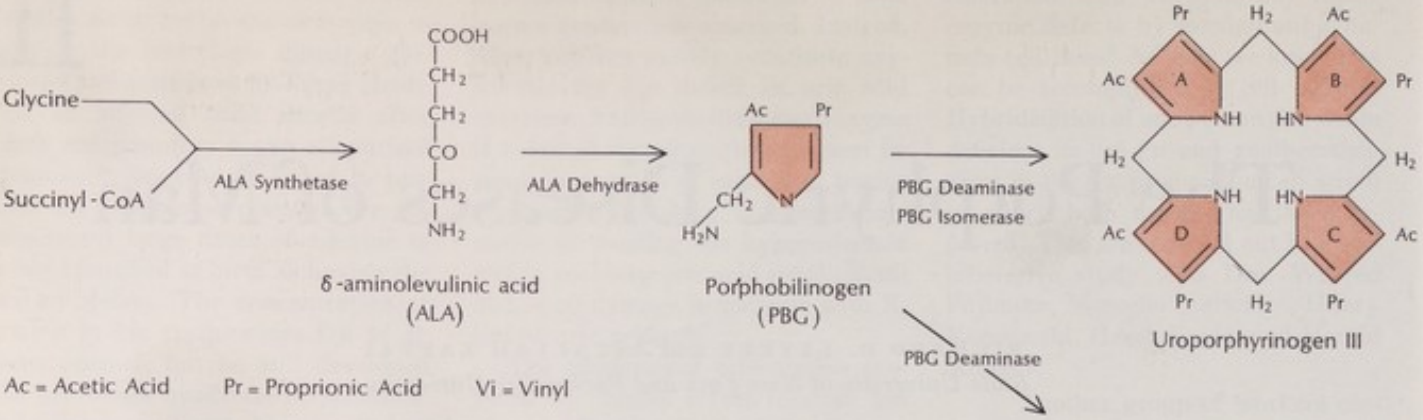
From a biochemical point of view, the pathway of heme synthesis is also an extremely useful system for study since the essential chemical dimensions of the sequence have been well elucidated, and the rate of activity of the pathway can be controlled by a variety of exogenous chemicals in various experimental systems.

The pathway of heme synthesis begins the same way in all aerobes – with two simple and prevalent molecules: glycine from the general amino acid pool and, from the Krebs cycle, succinate that has been activated by binding to coenzyme A. These two molecules are combined by the enzyme δ-aminolevulinic acid (ALA) synthetase to form the aminoketone known as δ-aminolevulinic acid. This step is a critical one in the control of heme synthesis because ALA is the first molecule that is essentially fully committed to the formation of the end product of the pathway, heme. Next, two molecules of ALA are enzymatically combined to form a ring-shaped molecule, the monopyrrole porphobilinogen (PBG). The enzyme involved at this stage, ALA dehydrase, is one known to be inhibited in lead poisoning, with the result that ALA accumulates and is excreted in the urine in excess amounts.

In the next step, four PBG molecules combine to form the tetrapyrrole uroporphyrinogen. Under normal circumstances the PBG is acted upon by two enzymes, one that deaminates it and another that isomerizes it to produce the natural tetrapyrrole isomer, uroporphyrinogen III, which proceeds to heme synthesis. A very small amount of PBG fails to be isomerized normally, resulting in the formation of type I isomers; these substances cannot go on to form heme. Uroporphyrinogen III is then decarboxylated to coproporphyrinogen III, which is further modified by oxidation and decarboxylation to form protoporphyrinogen IX. Autoxidation of this intermediate then yields protoporphyrin IX, the phylogenetic fork in the road where plants and animals take separate paths.

In the animal cell an iron atom is inserted into the protoporphyrin IX molecule to form the metalloporphyrin heme. In plants magnesium is inserted instead, beginning a process of further modification that produces chlorophyll. Plants, however, can synthesize some heme.

From this outline of the heme synthetic process (see illustrations on pages 114 and 115) it will be seen that any blockage of the pathway after the formation of ALA and prior to the insertion of iron into protoporphyrin IX will result in the excess accumulation and excretion of chemical intermediates such as ALA, PBG, uro-, copro-, and protoporphyrinogens. These "porphyrinogen" compounds, or reduced porphyrins containing six extra hydrogen atoms, are the true intermediates in the heme synthetic pathway. They are colorless and nonfluorescent in this reduced state, but when oxidized, as in voided urine, they become "porphyrins," assume a characteristic wine-red color under visible light, and thus provide one of the principal chromatic manifestations of certain porphyrias of man. The oxidized porphyrins, when excited by ultraviolet light in the long wavelength region, (~ 400A), also discharge energy in the form of an intense reddish-orange fluores-



The porphyrin-heme pathway: The formation of both the type III and type I series of isomers is shown. Type III isomers differ from those of type I in that one pyrrole ring (ring D) has been "flipped over" by the PBG-isomerase enzyme in the formation of the tetrapyrrole structure. The type I pathway cannot continue beyond the point indicated by the heavy bar (opposite page) because of the absence of an enzymic analogue of coproporphyrinogen III oxidase; the type III isomer proceeds to heme synthesis.

cence, thus making possible a second important urine test for certain porphyrias.

The latter characteristic of porphyrins, i.e., their ability to intensely absorb UV radiation, also accounts for the characteristic photosensitivity of certain porphyric patients. Porphyrins in the peripheral circulation and extracellular fluids, by absorbing solar UV radiation, become photosensitizing agents, capable of greatly damaging the cellular structures in the skin.

Porphyrins in any of their stereochemical constructions appear to play no currently demonstrable physiologic role other than to serve as chemical precursors for the synthesis of heme. Essentially, therefore, a porphyrin disorder is a heme disorder and one that may find expression in any function performed by heme proteins.

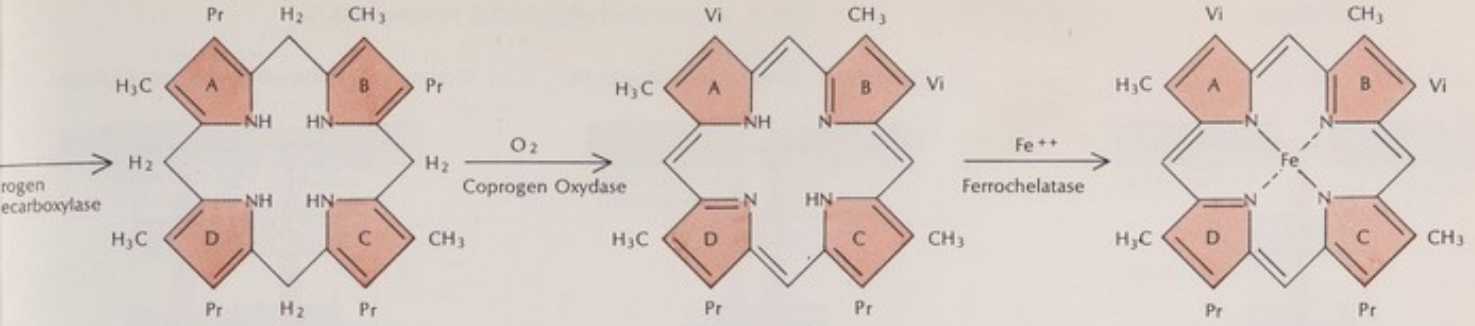
One significant heme protein function about which we have only recently begun to acquire knowledge is the detoxification role assigned to the hepatic heme protein, cytochrome P-450. This cytochrome is the terminal oxidase of an enzyme sequence localized in the endoplasmic reticulum and serves there as an activator of oxygen for hydroxylation reactions for various substrates. Cytochrome P-450 binds molecular oxygen, O_2, one atom of which will appear in the form of H_2O while the other atom is transferred to a great variety of exogenous and endogenous chemicals that are brought to the liver cell for metabolic disposal. The transfer of the oxygen atom to these substrates (usually in the form of an $-OH$ group) generally serves to biologically inactivate them or to increase their conjugability and solubility for excretory purposes. It is interesting that this oxidative pathway for drug metabolism in the liver is highly developed in all terrestrial species for whom, of course, the diffusion of endogenous or exogenous toxins into the infinite volume of surrounding ocean no longer serves as an available detoxification mechanism.

The sole function of heme is to act as a prosthetic group for such heme proteins as the hemoglobins, cytochromes, myoglobins, catalases, and peroxidases. Its functional role completed, its tetrapyrrole ring is ruptured to form the linear tetrapyrrole bilirubin that is conjugated and excreted along pathways previously described (see Arias, Chapter 12).

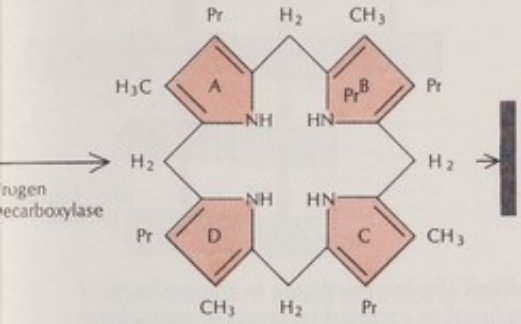
There are two main categories of porphyric disorders of man, based on the tissue in which the biochemical lesion is localized: erythropoietic and hepatic. The hepatic porphyrias may be further classified etiologically into those that are inherited and those that are acquired (see table on page 119). Only the major forms of the porphyrias are discussed in the following sections.

Two principal erythropoietic types of porphyria, both inherited, have been described. The first of these, congenital erythropoietic porphyria, is an extremely rare autosomal recessive disorder. The biochemical disturbance in this disease is reflected in a marked overproduction of uroporphyrin I and coproporphyrin I in circulating erythrocytes and bone marrow cells. While this would seem to reflect a defect of PBG isomerase activity, it is obvious that such a deficiency must be a relative one, since a total absence of PBG isomerase would be incompatible with the production of heme and therefore with life. Moreover, these patients do have normal complements of hemoglobin in their erythrocytes so that at least normal amounts of type III isomers must also be produced. In fact, recent evidence indicates that the PBG isomerase defect is a relative one and that there may be as well an overproduction of ALA and PBG with which the isomerase cannot cope, the surplus being channeled into the nonfunctional type I tetrapyrrole structure.

This disorder thus appears to represent a deficiency of normal enzymatic control mechanisms at a step prior to the formation of uroporphyrin, most probably at the level of ALA synthetase, associated with an additional defect involving the relative inability of PBG isomerase to adapt to this overproduction of its substrate.



Coproporphyrinogen III — Protoporphyrin IX — Heme



Coproporphyrinogen I

The disease expresses itself clinically shortly after birth, principally through the dramatic appearance of red urine. Photosensitivity is another cardinal manifestation, resulting in pronounced vesicle and bulla formation; these heal slowly and are prone to secondary infection and ulceration, with severe, permanent mutilation of light-exposed skin. Deposition of porphyrins in the teeth and nails produces brownish-red discoloration and red fluorescence when these tissues are exposed to ultraviolet light. Splenomegaly is common, as is a hemolytic anemia that responds to splenectomy.

These vivid expressions of the disease have led to at least one interesting bit of nonmedical speculation. It has been proposed that the legendary werewolves of the Middle Ages may have been persons afflicted with congenital erythropoietic porphyria, having fluorescent teeth and nails, mutilated and deformed ears, nose, and eyelids, and nocturnal habits because of their sensitivity to light.

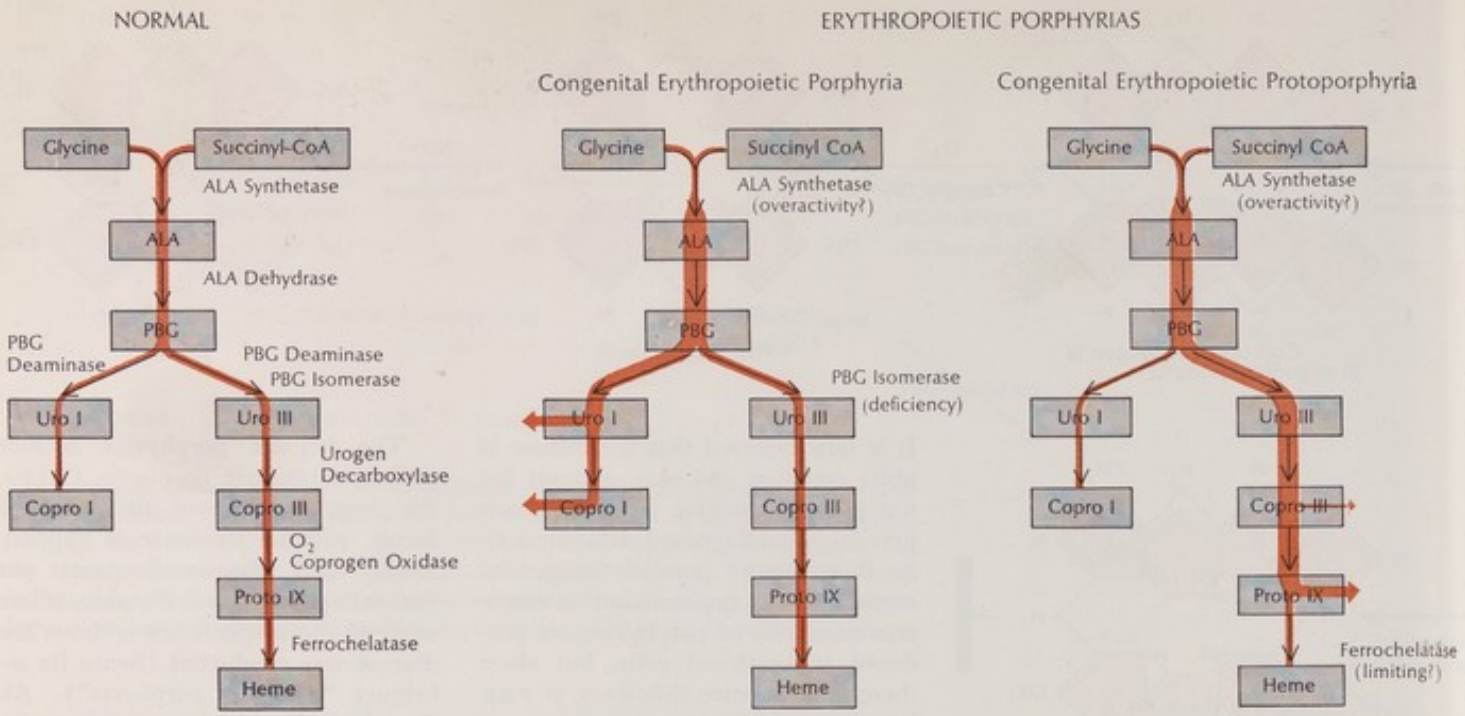
The second major entity within this category is known as congenital erythropoietic protoporphyria. Apparently transmitted as an autosomal dominant trait, the disorder was first described in 1961; a large number of cases have been reported since then. It is now believed that the disease is quite common and may account for many light-sensitive skin reactions previously undiagnosed. Biochemically, this disorder parallels congenital erythropoietic porphyria in that excessive quantities of porphyrins are produced in erythroid cells, but since there is no relative deficiency of PBG isomerase activity the excessive quantities of porphyrins produced are all in the normal type III series. Laboratory findings are those of abnormally high protoporphyrin levels in circulating erythrocytes, bone marrow cells, and plasma. In contrast to congenital erythropoietic porphyria, there is no increase in fecal or urinary uroporphyrins. Also, recent evidence suggests that the current erythropoietic classification of this disorder may possibly be too restrictive, since certain experimental data would be consistent with an hepatic origin for a portion of the excess porphyrins that are synthesized. If this proves to be the case, this will be the only porphyric disease with biochemical lesions in both hepatic and erythropoietic cells.

The main clinical manifestation of erythropoietic protoporphyria is usually mild photosensitivity, resulting in some patients in slight eczematization and inflammation in exposed skin. Onset of symptoms is usually during the first few years of life, but may be delayed until adulthood. In a few patients hypersplenism with associated hemolytic anemia, which responds favorably to splenectomy, has been reported. Although anemia is not common, recent laboratory evidence does suggest that the increased protoporphyrin content of erythrocytes makes them inordinately susceptible to lysis when subjected in vitro to ultraviolet light. This hemolysis can be inhibited by reducing agents.

The hepatic porphyrias include acute intermittent porphyria (AIP), the most common of the inherited forms, with an incidence of approximately 1.5 diagnosed cases per 100,000 population in Sweden, where most of the original research on the disease was conducted (hence its sobriquet "Swedish porphyria"). Although its incidence appears to be highest in that country, especially in some of its northern reaches, it has now been described in all corners and races of the world. Recently it has attained some renown from the speculation that England's King George III may have been one of its victims.

Inherited as an autosomal dominant, acute intermittent porphyria nevertheless afflicts three females to every two males, very rarely manifests itself before puberty, varies in the degree in which the genetic defect is clinically expressed, and appears to be susceptible to endocrinologic exacerbation. The premenstrual period and pregnancy may be especially difficult for certain women with this disorder.

Unfortunately for ease of diagnosis, the most classic acute clinical presentation of AIP is that of colicky abdominal pain, often associated with vomiting, constipation, or, infrequently, diarrhea. Fever, leukocytosis, proteinuria, and tachycardia – in any combination – may accompany the pain. In short, AIP can mimic almost any surgical emergency in the abdomen. Not a few such patients are subjected to needless laparotomy because physicians overlook the possibility of porphyria. The abdominal pain is most probably related to the neurologic involvement that is so characteristic of AIP, the autonomic nervous system being affected in some manner, as dis-



Schematic shows known or suggested biochemical lesions occurring in the different porphyrias described in the text. The approximate degree to which the intermediates in the pathway of heme synthesis are produced or excreted in excess of normal (far

cussed below. Almost all patients show some degree of behavioral disturbance, and, indeed, a frank psychiatric disorder may be the initial presenting picture. The incidence of porphyria in mental institutions appears in fact to be significantly higher than in the general population. In some individuals, the presenting neurologic manifestation may be epileptiform convulsions, paraplegia, quadriplegia, uniplegia, or lesser motor or sensory involvements. One interesting phenomenon, containing a pitfall for clinical management of these individuals, has only recently been recognized. For some unknown reason these patients may secrete inappropriate amounts of antidiuretic hormone and present with spurious dehydration manifested by hyposmolar serum and hyperosmolar urine. It is dangerous to force fluids in such patients and fluid restriction is indicated even though the clinical signs seem to demand the reverse.

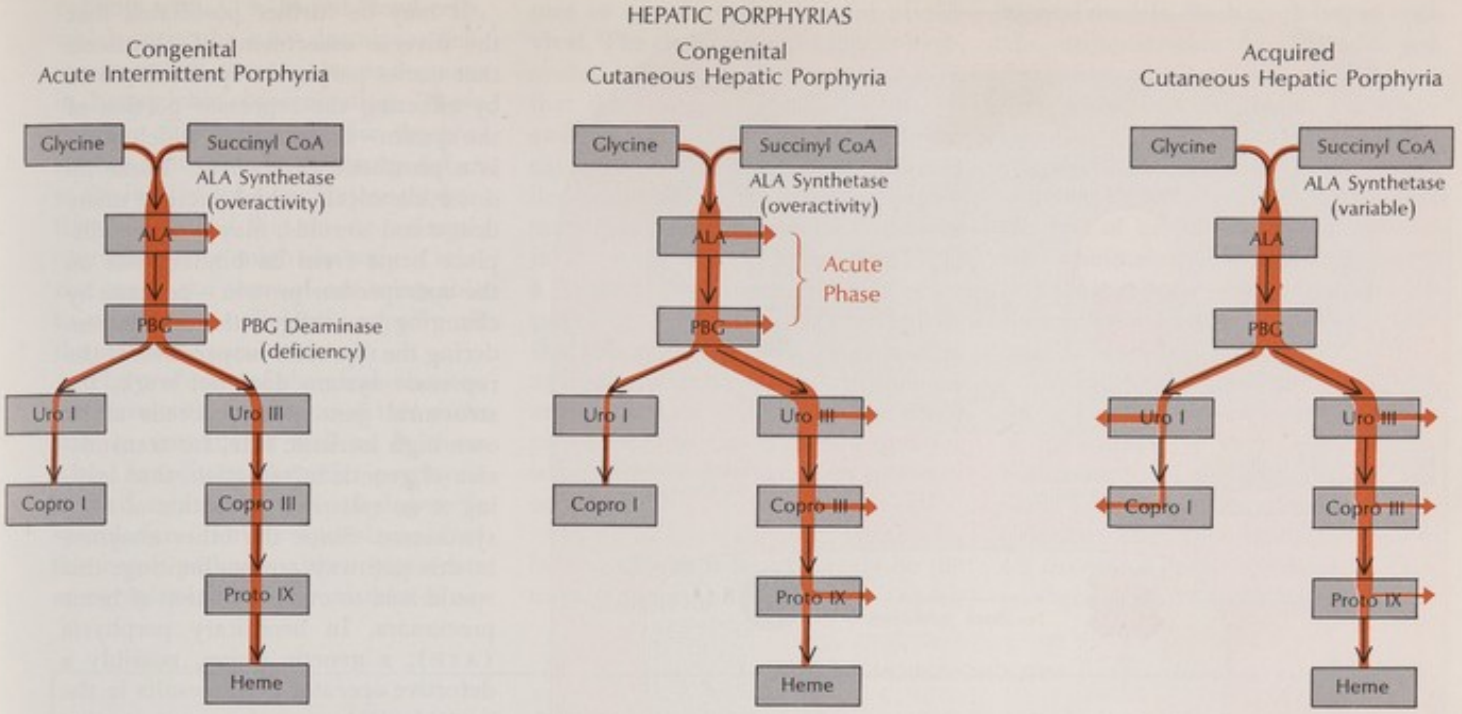
Patients with AIP do not generally have high urine levels of porphyrins but rather excrete excessive quantities of ALA and PBG. They are not photosensitive and their urine when voided is usually colorless – an important distinction from the other porphyrias. However, after voiding, the ALA and PBG contained in urine undergo spontaneous, nonenzymatic conversion to uroporphyrin, which, being oxidized on standing, yields the wine-red color associated with the porphyrias. A key laboratory finding is produced by the Watson-Schwartz test (see photos on page 120). The basis of this test is the insolubility in chloroform (or, more recently and accurately, in *n*-butanol) of the red complex formed by the interaction of PBG in the urine and Ehrlich's reagent.

A treatment hazard, whose implications will be discussed more fully later, is that a number of drugs may exacerbate or precipitate the acute disease in individuals carrying the genetic lesion of AIP. Unfortunately, some of these drugs, i.e., barbiturates, sulfonamides, and analeptic and anticonvulsive compounds, might ordinarily be used by physicians for treatment of the puzzling signs and symptoms with which the disease presents.

The other major congenital variety of hepatic porphyria is cutaneous hepatic porphyria. Although it is known as South African porphyria (porphyria variegata) because of its high incidence among the Boers, it is not limited to this ethnic group. This disease is transmitted as an autosomal dominant disorder, and has its onset later than AIP, usually in the fourth to fifth decade of life. During the acute stage of the disease the urine contains large amounts of ALA, PBG, uroporphyrin, and coproporphyrin; the stool contains large quantities of all the porphyrins as well. In the quiescent phase, the urine findings may be almost normal but high porphyrin levels can be found in the stool. These patients are subject to abdominal pains and neurologic disturbances as in AIP. In addition, they have photosensitivity (which helps distinguish them from patients with AIP). For some reason the skin lesions are more common among men, the neurologic signs more frequent in women.

While only hereditary factors have been implicated thus far in erythropoietic porphyria, hepatic porphyria may also be acquired. The range of expression of the acquired disease is broad, from severe manifestations duplicating those seen in the most extreme cases of congenital cutaneous hepatic porphyria to an entirely asymptomatic porphyrinuria. In many cases studied, acquired hepatic porphyria seems to be associated with the ingestion of exogenous porphyria-inducing agents against a background of preexisting hepatic deficiency of some type, usually cirrhotic.

Alcohol seems to be one inciting agent, although the manner of its involvement in acquired porphyria is un-



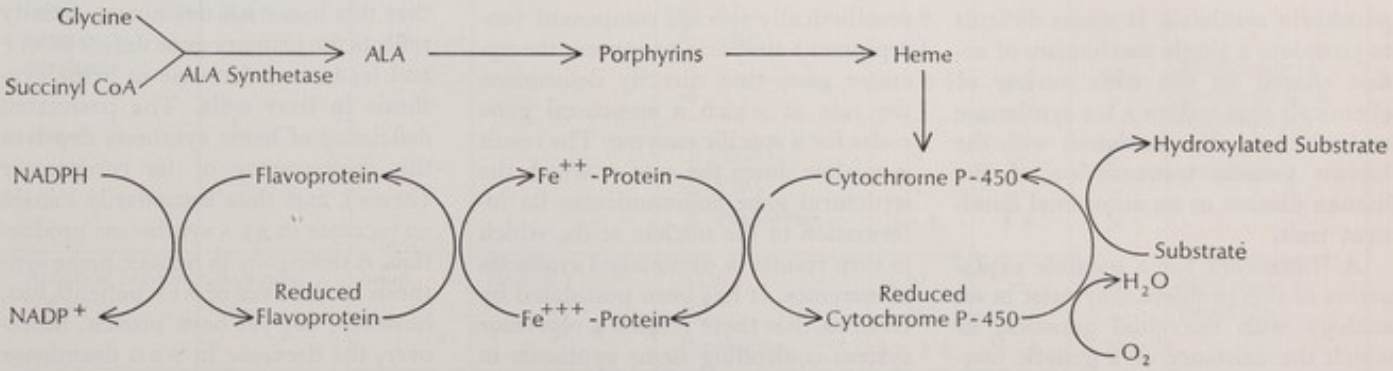
left) is indicated by the thickness of the color lines and the arrows. The large amounts of porphyrins excreted by patients with the cutaneous hepatic porphyrias may reflect enzymatic abnormalities at these levels of the pathway, but this is still speculative.

clear. A high incidence of the disease, without familial pattern, has been observed among the Bantus of South Africa, a tribe that consumes large quantities of an iron-rich adulterated alcoholic beverage. An especially definitive relationship between exogenous chemicals and acquired hepatic porphyria was seen with the chemical agent hexachlorbenzene, a wheat fungicide whose ingestion proved to be responsible for an epidemic of acquired cutaneous porphyria in Turkey. It was subsequently shown that hexachlorbenzene could induce hepatic porphyria in experimental animals.

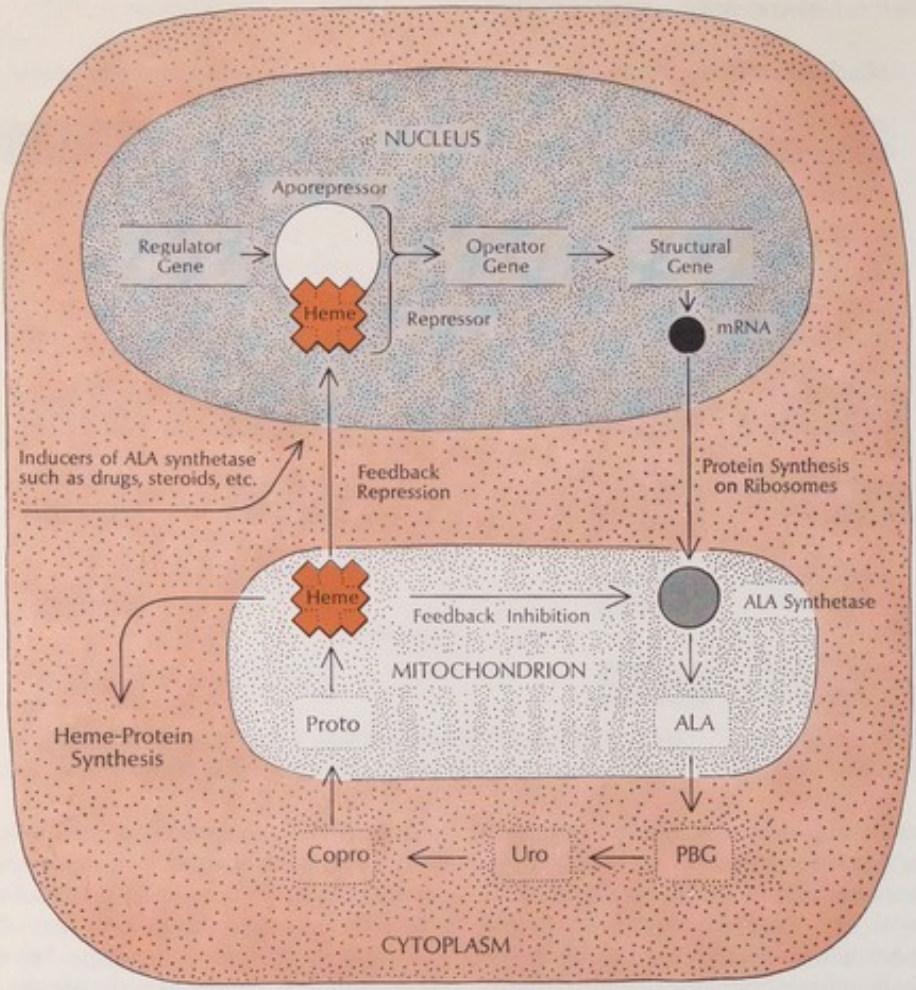
The Turkish experience was the first clear demonstration that a chemical could induce hepatic porphyria in a human population free of genetic susceptibility. Since these observations with hexachlorbenzene were made, abnormalities in hepatic porphyrin metabolism have been described in patients under treatment with various hormones such as stilbestrol and oral contraceptive steroids, and various drugs such as griseofulvin, the sulfonamides, and tolbutamide. Apparently, many of the drugs that can exacerbate the hereditary hepatic porphyric disease may also initiate porphyrinuria in seemingly normal persons—with what clinical sequelae, if any, not known.

Very importantly, however, these drugs have made it possible to establish experimental models of hepatic porphyria in animals of various types and in tissue cultures of liver cells. In the laboratory, a large number of structurally diverse chemicals have been found to induce hepatic porphyrin synthesis. The chemicals appear to act in a common manner in the biosynthetic pathway: via induction of the mitochondrial enzyme ALA synthetase, which controls the rate limiting step in hepatic porphyria synthesis.

A problem in studies of experimental hepatic porphyria is that of reconciling the chemical heterogeneity of the provocative agents with the apparent uniformity of their effect on



The drug-metabolizing enzyme sequence localized in liver cell microsomes is related to the heme pathway through cytochrome P-450, the heme protein that serves as the terminal oxidase in the sequence. The substrates oxidized may be either endogenous or exogenous in origin and are generally (though not always) biologically inactivated by such oxidative transformation.



Schematic depicts the control mechanism through which heme synthesis is regulated in liver cells, as postulated by Granick. In this formulation, production of ALA synthetase is governed by a repressor-operator mechanism. The repressor consists of an aporepressor protein, with heme acting as a corepressor. When heme is displaced from its binding site by a steroid or drug, the repressor is inactivated, the structural gene for ALA synthetase can code for more messenger RNA, and more enzyme is synthesized. The concentration of active inducers in the cell may be dependent on the rate at which they are inactivated by chemical transformation (oxidation, conjugation, etc.). In liver cells, feedback repression of ALA synthetase formation by heme appears to be the principal control; in erythroid cells, feedback inhibition of the enzyme may be the control for heme synthesis. Recent evidence suggests certain inducers effect translational rather than transcriptional mechanisms in regulating ALA synthetase.

porphyrin synthesis. It seems difficult to postulate a single mechanism of action shared by the wide variety of chemicals that induce ALA synthetase and which is also consistent with the known genetic transmission of the human disease as an autosomal dominant trait.

A framework for a possible explanation of this problem may exist in an analogy with microbial genetics, in which the existence of a genetic control mechanism responsive to induction and repression is presupposed. A regulator gene is postulated that makes a protein (aporepressor) with a prosthetically specific component (corepressor) that acts to repress the operator gene that directly determines the rate at which a structural gene codes for a specific enzyme. The result is to slow down the rate at which the structural gene communicates its information to the nucleic acids, which in turn results in diminished synthesis of enzymes. It has been postulated by Granick that there is such a repressor system controlling heme synthesis in the liver and that heme, the end product of the pathway, is a corepressor in this control system, thus regulating the rate of its own formation.

It may be further postulated that the diverse assortment of chemicals that excite porphyrin synthesis do so by affecting the repressor portion of the system – that entity of which heme is a prosthetic component. These inducer chemicals, which include many drugs and steroids, may block or displace heme from its binding site on the aporepressor protein – perhaps by changing its configuration – thus rendering the repressor inoperative. If the repressor system does not work, the structural gene then proceeds at its own high intrinsic rate, for transmission of genetic information, thus leading to an enhanced formation of ALA synthetase. Since the other enzymes in this pathway are nonlimiting, this would lead to overproduction of heme precursors. In hereditary porphyria (AIP), a genetic lesion, possibly a defective operator gene, results in the control mechanism for ALA synthetase being less repressed than normally and responding inappropriately to various drugs. This lesion is manifested by an increased formation of ALA synthetase, and a high level of this enzyme activity in the liver now appears to be a characteristic abnormality in AIP patients.

Recently, a decreased activity of PBG deaminase has been demonstrated in the erythrocytes and liver cells of patients with AIP. In addition, a decrease in the activity of this enzyme has been noted in the red blood cells of some asymptomatic family members. In our own studies of AIP family lineages, decreased PBG deaminase in red blood cells has been found in a group of prepubertal children in whom the clinical syndrome is, as we have noted above, never manifest. It has been suggested by some investigators that this low PBG deaminase activity reflects the primary gene defect in AIP and leads to a decrease in heme synthesis in liver cells. The postulated deficiency of heme synthesis deprives the aporepressor of its corepressor (heme) and thus secondarily causes an increase in ALA synthetase production. A deficiency in hepatic heme synthesis in the liver of AIP patients has, however, not yet been proven. Moreover, the decrease in PBG deaminase activity averages only about 50% of normal and there is, as yet, no demonstrated decrease in heme-proteins such as cytochrome P-450 in the livers of

patients with AIP. In red blood cells, which carry the PBG deaminase deficiency as well, there is also no decrease in hemoglobin concentration. These findings thus still allow for the possibility that there is a primary defect at the level of ALA synthetase production in AIP and that the PBG deaminase deficiency might be only a relative one in the sense of not being able to metabolize sufficiently rapidly all the substrate presented to it.

The profusion of chemicals that can interfere with regulation of this pathway in normals as well as in those with hereditary hepatic porphyria invites teleologic and phylogenetic speculation. This responsiveness of the heme pathway either represents an egregious metabolic mistake of nature or an adaptation helpful to survival. The chances would appear high, evolution being as pragmatic as it is, that there is good reason for the responsiveness of this pathway to chemical influences. This reason probably lies in the role of the liver as the primary detoxification organ and of heme in the form of cytochrome P-450 as a central detoxification agent (see page 117). It is reasonable to suppose that this system evolved for protection against a wide variety of endogenous and exogenous chemicals that could prove toxic to the organism and for which there was no excretory pathway once a species left the seas.

What we know about the porphyrin-heme pathway is based largely on the unique opportunities to study experimental models. Such work began with the demonstration by Schmid and Schwartz, working in C. J. Watson's laboratory, that Sedormid, a barbiturate-like compound, could induce a porphyria-like disease. A next key step was Granick's demonstration that the rate of hepatic heme synthesis in whole animals and in liver cells grown in tissue culture was determined by the rate of formation of ALA synthetase.

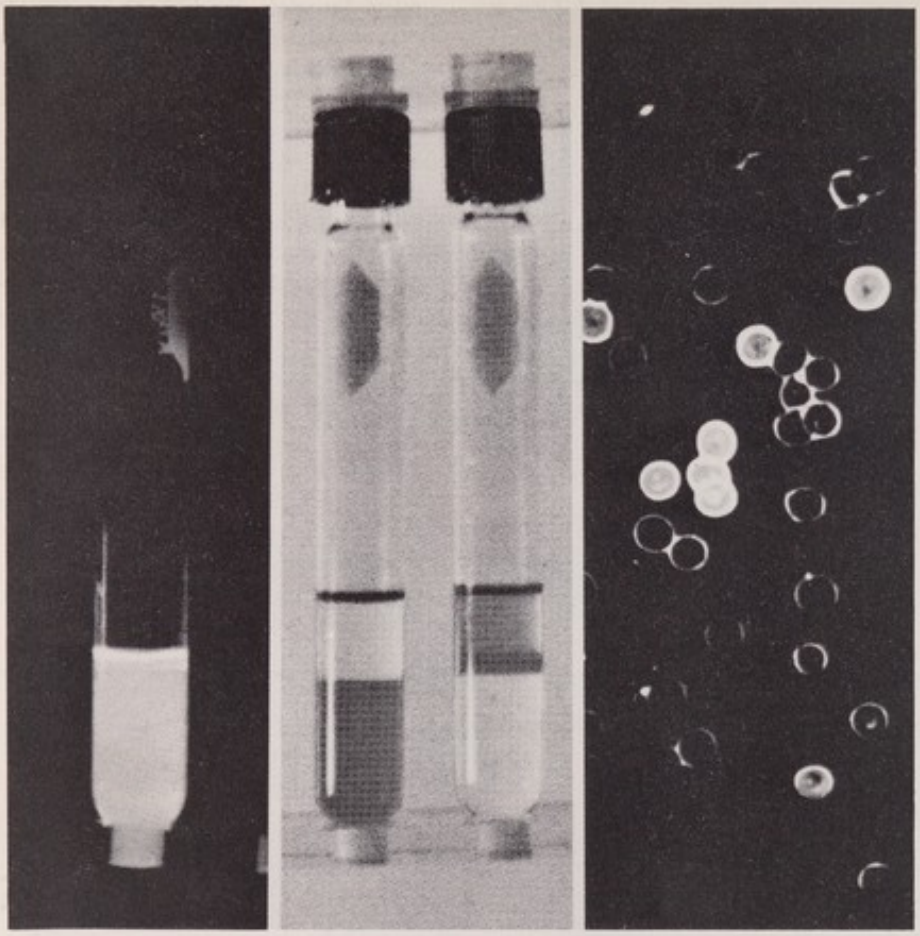
Granick extensively explored the subject of controls on heme synthesis in his studies of liver cell cultures. Following his observation that the livers of animals treated with porphyria-inducing chemicals contained up to 40 times the normal amount of ALA synthetase, he demonstrated that add-

The Porphyrias: Chemical Findings and Clinical Aspects

Disorder	*Inheritance*	*Age of Clinical Onset*	*Primary Organ Involvement*	*Chemical Findings*	*Urine*	*Feces*	*Red Blood Cells*	*Primary Symptoms*
Congenital erythropoietic porphyria	Autosomal recessive	Birth to 5 years	Erythropoietic	Uroporphyrin I	↑↑	↑	↑↑	Severe photosensitivity
				Coproporphyrin I	↑	↑↑	↑	
				Protoporphyrin	N	N	N	
Congenital erythropoietic protoporphyria*	Autosomal dominant	Usually childhood	Erythropoietic	Uroporphyrin III	N	N	N	Mild photosensitivity
				Coproporphyrin III	↑	↑	N–↑	
				Protoporphyrin	N	↑↑	↑↑	
Acute intermittent porphyria	Autosomal dominant	Adult	Hepatic	Acute Phase: ALA	↑↑	–	–	Mild to severe neurologic-visceral symptoms
				Acute Phase: PBG	↑↑	–	–	
				Acute Phase: Porphyrins	N–↑	N–↑	N	
				Remission: ALA	↑	–	–	
				Remission: PBG	↑	–	–	
				Remission: Porphyrins	N	N–↑	N	
Congenital cutaneous hepatic porphyria	Autosomal dominant	Adult	Hepatic	Acute Phase: ALA	↑↑	–	–	Mild to severe photosensitivity and neurologic-visceral symptoms
				Acute Phase: PBG	↑↑	–	–	
				Acute Phase: Porphyrins	↑–↑↑	↑↑	N	
				Remission: ALA	N	–	–	
				Remission: PBG	N	–	–	
				Remission: Porphyrins	N–↑	↑↑	N	
Acquired hepatic porphyria**	Acquired	Adult	Hepatic	ALA	N	–	–	Mild photosensitivity
				PBG	N	–	–	
				Porphyrins	↑–↑↑	N–↑	N	

* *Recent evidence suggests that the liver may also be involved in this disorder.* *N Normal* ↑ *Increase* ↑↑ *Marked Increase*

** *Genetic factors may also operate in some individuals with this syndrome.*



Urine containing excess amounts of porphyrins fluoresces under ultraviolet light (left). In center photo, test tube on left, in the bottom phase, shows a positive Watson-Schwartz test on urine from a patient with acute intermittent porphyria. Butanol (top phase) has not extracted the chromogen produced in the test. In nonporphyric individuals giving false positive results (test tube on right) butanol does extract the chromogen from the lower aqueous phase. Photomicrograph at right (courtesy Dr. D. Cripps) demonstrates porphyrin fluorescence in circulating erythrocytes under UV light.

ing such chemicals to liver cells growing in tissue cultures induced the same effect seen in vivo and was due to enhanced de novo formation of the enzyme. Subsequently, studies in the laboratories of Tschudy in this country and Nakao in Japan identified increased levels of ALA synthetase in the livers of patients with acute intermittent porphyria.

Certain clinical clues in AIP – i.e., its postpubertal onset, its exacerbation in certain women during menstruation and pregnancy, its female preponderance, etc. – implicated a possible role for the endocrine system in the pathophysiology of this genetic disease. Subsequently it was shown that many steroid metabolites, derived from natural sex hormones, can elicit an experimental hepatic porphyria similar in all respects to that induced by the most potent porphyrin-inducing chemicals and drugs. It is of great interest, in this regard, that patients with clinically manifest AIP have, in recent studies from our laboratories, been shown to generate disproportionately large fractions of such porphyria-inducing 5β steroids when studied with isotopically labeled precursors such as testosterone and dehydroisoandrosterone. This hormone metabolism defect results from a profound deficiency of steroid 5α-reductase activity in the livers of AIP patients. As a result of the 5α-reductase deficiency, natural hormones are shunted to the 5β pathway for their metabolism; and it is via this pathway that potent 5β steroid inducers of the porphyrin-heme pathway are derived. Such 5β steroids may be the porphyrin-inducing humoral agents we earlier identified in the sera of AIP patients during acute exacerbations of the disease.

The abdominal pain and the neurologic abnormalities characteristic of AIP have no clear explanation as yet, and there appears to be no direct correlation between these symptoms and the known biochemical abnormalities of this disease. Recent experiments carried out with Watson and Feldman indicate that ALA, PBG, and a derivative, porphobilin, can significantly impair neuromuscular transmission by impeding acetylcholine release. Such impairment, if it occurs in AIP patients, could relate the autonomic dysfunction characteristic of the disease to the known abnormalities of the heme pathway.

It would be good to report that our increasing knowledge of porphyrin biochemistry has yielded positive answers to the treatment of human porphyria, but that is not yet the case. Indeed, in seeking a specific chemical therapy, one is not sure at which point along the heme pathway it would be most useful to attack the disorder. Should one seek to enhance or suppress the activity of the pathway, and how? Any pharmacologic tampering with the rate of heme synthesis runs the risk of producing Pyrrhic victories. Such manipulation could result in an interference with the body's detoxification capacity; thus, for example, one may diminish to low levels the production of porphyrins and precursors but might possibly thereby restrict the patient's defenses against internal and external toxins. On the other hand, if any chemical intermediate in the pathway is itself toxic, as suggested above, suppressing the activity of this biosynthetic sequence could be highly beneficial. Symptomatic treatment also contains pitfalls since, as has been noted above, many of the clinically indicated drugs are porphyrin-producing in their own right and can have deleterious effects. The physician with a porphyric patient must thus resist the temptation to prescribe medication for every symptom.

The activity of the porphyrin pathway may, however, be slowed down by one relatively innocuous form of treatment. Glucose has been shown to repress the induction of a variety of enzymes, for reasons yet obscure, and

in experimental animals large amounts of glucose suppress ALA synthetase induction and block the response to drugs that ordinarily produce hepatic porphyria. Some patients have been helped symptomatically by glucose when administered in very large amounts, and such treatment may markedly diminish the excessive output of porphyrins and precursors in such patients. Conversely, glucose deprivation may have the opposite effect in patients with porphyria.

With respect to the management of photosensitivity, which accompanies both erythropoietic and hepatic forms of porphyria, the best approach is to guard the patients against exposure to sunlight – by avoidance with sunscreens, ointments, etc. Cholestyramine, which binds porphyrins in the intestinal tract, and extensive phlebotomy have also proved quite useful in the management of cutaneous hepatic porphyria.

The lack of effective treatment for porphyria, while disappointing, is not a public health calamity since these diseases are so rare. But what we have learned about studying the biochemistry of these uncommon disorders of porphyrin-heme metabolism has a much more general clinical relevance. Most physicians will see few porphyrics in their practices, but when they prescribe a drug for almost any condition they may influence heme synthesis in the liver and, secondarily, the rate at which drug metabolism takes place in the liver. We all need to keep informed about the pharmacologic meaning of recent discoveries with respect to hepatic heme synthesis and how it affects and is affected by the administration of various drugs or other environmental agents such as DDT. Consider, for example, the implications of the ability of phenobarbital to stimulate heme formation in the liver, with a consequent increase in cytochrome P-450 and amplified metabolic capacity for drugs. This means that the moment a physician gives another drug in conjunction with phenobarbital (or even an additional barbiturate), he has lost his original basis for calculating dosages and pharmacologic effects.

This phenomenon has considerable relevance to the clinical use of combined drug therapy, since the possibility that one drug which is administered to a patient will significantly alter the activity of another must be kept continually in mind. Hereditary hepatic porphyria may simply be a rare, genetically determined exaggeration of a biochemical process with which we are all dealing regularly in our practices.

12

Genetic Disorders in Bilirubin Metabolism

IRWIN M. ARIAS
Albert Einstein College of Medicine

Although bilirubin was demonstrated in both blood and bile more than 50 years ago, the manner of its transfer from one to another has remained unknown. A window on the process was opened in 1952 when three separate groups of investigators showed that bilirubin in bile is conjugated with glucuronic acid, differing in this respect from normal serum bilirubin, which is unconjugated and bound to albumin. The intrahepatic conjugation, catalyzed by the enzyme glucuronyl transferase, confers high water solubility and thus enables the liver to excrete bilirubin. Once this process was elucidated, it became possible to separate the blood-bile transfer into at least three steps: liver uptake, conjugation, and excretion.

Since 1965 or so we have been learning to study these broad physiologic stages in terms of the finer details of action and reaction at the molecular level. Many of the questions we now ask are biologic questions that go beyond bilirubin as such. It turns out that one or more steps in the mechanism for transferring bilirubin from blood to bile apply not just to bilirubin but to a great variety of organic anions of which bilirubin is only one example. Others are steroid hormones, porphyrins, dyes, metabolites, drugs of many kinds, and radiographic contrast materials.

Nature helped. Sir Archibald Garrod once declared that mutants were nature's gifts to the scientist. It is doubtful that the stages of bilirubin metabolism would have been recognized were it not for the existence of mutants with inherited jaundice, their processing of bilirubin sidetracked at some step along the way and the sequence of metabolic events thereby made susceptible to observation and analysis. The study of human and animal mutants, whose deficiencies are sometimes damaging, sometimes benign, provided a bridge between laboratory and clinic. Demonstrations of the various levels at which bilirubin metabolism may undergo alteration helped increase our ability to recognize and explain clinical conditions.

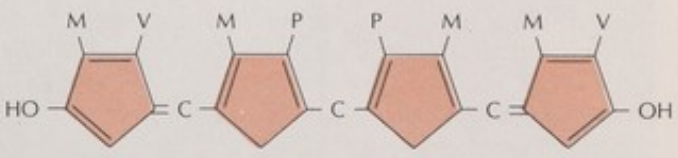
All these themes – the basic biologic implications of bilirubin research, the active flux between clinic and laboratory, the importance of mutants – have fundamentally influenced advances in this field.

For a long time it has been known that when any of a variety of organic anions like bilirubin or sulfobromophthalein (BSP) was injected into an animal, very shortly most of the substance was present in the liver. But the reasons behind the liver's selectivity for such materials eluded us. Some tried to explain it on the basis of hepatic blood flow; however, two other possibilities could be entertained: either the liver cell had some unique property in its plasma membrane that selectively admitted these substances or the liver cell contained some receptor-like material – perhaps a protein that selectively bound bilirubin – or both. Two such protein receptors, named Y and Z proteins, have been recently isolated from liver cytoplasm and show affinity for binding organic anions like bilirubin and BSP.

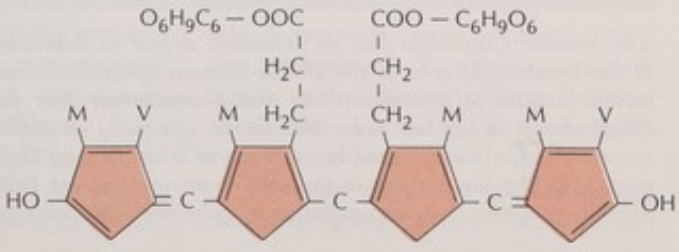
Evidence for believing that Y and Z proteins may be im-

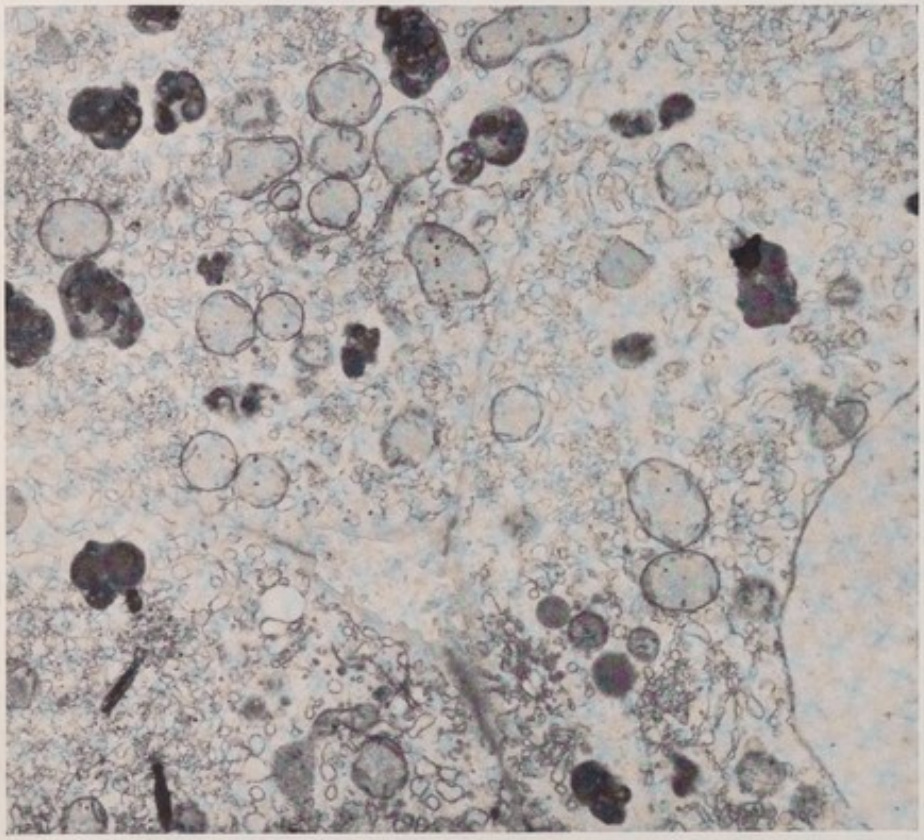
The Bilirubin Molecule

Free Bilirubin



Conjugated Bilirubin





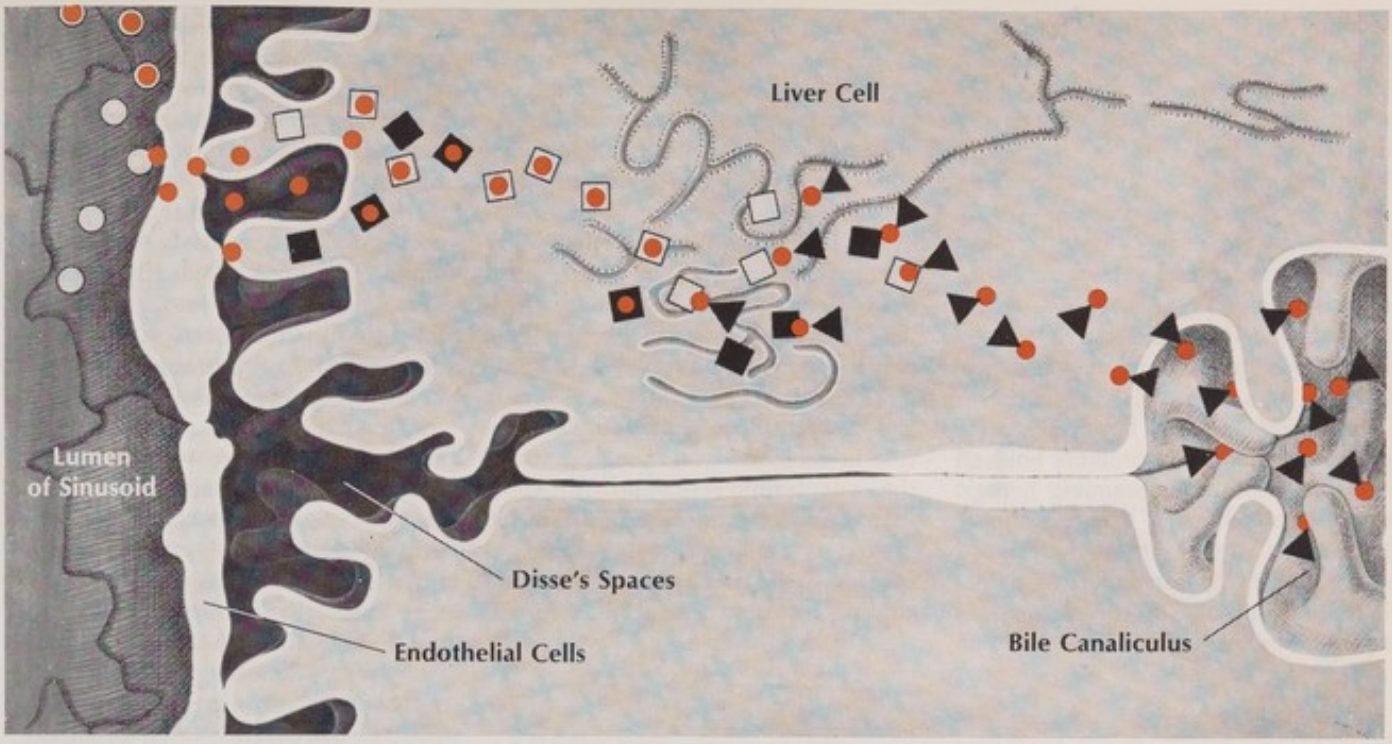
The mutant Corriedale has an inherited defect in bilirubin excretion very similar to the impairment seen in the Dubin-Johnson syndrome in man. The sheep were detected because of photosensitivity that characterizes liver impairment in ruminants. Sheep shown at top has many ulcerations, especially around the eyes. At autopsy the mutant sheep's liver is found to be black, as is the liver of Dubin-Johnson patients (see page 128). The black areas in the electronmicrograph are believed to be melanin-like deposits resulting from accumulation of epinephrine metabolites.

portant in determining the uptake of bilirubin by the liver comes from phylogeny, laboratory experiment, and clinical observation. Thus far, Y and Z have been found only in the liver and only in air-breathing vertebrates. Y and Z may represent an evolutionary adaptation to changed excretory circumstances associated with the transition from sea to land. They are not found in fish or amphibia at that ontogenic stage when anion exchange is usually conducted across the gill. In the laboratory we can study the function of these proteins in relationship to bilirubin transport by administering drugs that more selectively bind to the proteins than either bilirubin or BSP.

The discovery of this experimental model proceeded from the clinical observation that patients in Scandinavia with fish tapeworm who were treated with the male fern extract oleoresin aspidium became jaundiced. They practically doubled their serum bilirubin, all in the unconjugated fraction. The extract's active ingredient, flavaspidic acid, was subsequently isolated and shown to interfere with liver uptake of bilirubin and BSP in man and rabbits. It was demonstrated that flavaspidic acid in vivo or in vitro competes with bilirubin and BSP for binding with, and manifests greater affinity for, the Z protein. We, therefore, propose that the availability of receptor sites for bilirubin transfer from blood into the liver cell is reduced in vivo.

Further evidence came from another clinical observation subjected to experimental study. It was noted a few years ago that a group of businessmen undergoing a detailed annual medical examination showed increased serum concentrations of unconjugated bilirubin that returned to normal within a few days. It was proposed that hyperbilirubinemia (HBR) resulted from some substance to which all the men had been recently exposed. A possibility was iodipamide, a cholecystographic medium administered the day before blood was drawn for bilirubin determination. Laboratory work demonstrated that the organic anion iodipamide, like flavispidic acid, competed with bilirubin and BSP for binding to Z and probably thereby accounted for transient unconjugated hyperbilirubinemia.

Ontogenic studies also suggest that



The stages of uptake, conjugation, and excretion through which bilirubin makes the transition from blood to bile are illustrated in rendering above. Many details of this model remain to be elucidated, but identification of the various levels at which bilirubin metabolism may undergo change has greatly improved ability to recognize and explain clinical conditions.

Y and Z play an important role in hepatic uptake mechanisms. It has been fairly well demonstrated that there is progressive development of Y and Z proteins in newborns; this has been studied extensively in guinea pigs, monkeys, and, to a limited extent, in man. The basic pattern seems the same in all species. Z protein, of which there is a smaller amount than Y in the liver, has the lesser organic anion binding capacity but develops very quickly during fetal life to reach adult levels practically at birth. On the other hand, Y, the major binding protein, is almost absent at birth and develops slowly to reach mature levels in the second or third week of life. It has also been demonstrated in monkeys and humans that the ability to transfer organic anions from serum to liver increases steadily with age from birth on, full development occurring in about two or three weeks. In other words, the maturing ability of the neonate's liver to take up bilirubin generally correlates with the rate of development of Y protein.

The implication of this has led us recently to take a fresh look at the so-called physiologic jaundice so often seen in the newborn infant. Practically every three-to-five-day-old infant develops unconjugated H B R and turns a bit yellow. Until recently it was generally believed that this phenomenon was due to an immaturity of conjugation, largely in terms of the enzyme glucuronyl transferase, even though results of actual measurements of the enzyme in human infants had been inconclusive. However, with the demonstration of Y protein maturation, one must consider the possibility that the condition may be related to uptake and Y protein, rather than to conjugation and glucuronyl transferase. The reassessment is supported by the demonstration that the blood-to-liver transfer rate of other materials is equally affected in the newborn, these materials being organic anions that do not require glucuronide formation, such as B S P, which is conjugated to glutathione, or indocyanine green, which is not conjugated at all. If the problem had to do only with the enzyme, only glucuronide - conjugated substances would presumably be affected.

We have begun to learn more about the Y and Z proteins. We know that Y is a basic protein with a molecular weight of approximately 32,000 and that Z is a different basic protein weighing about 10,000. They also differ in their turnover.

We have learned more, in large measure due to work by Humberto Reyes, an International Research Fellow of the U S P H S. Dr. Reyes observed that it was possible to more than double the amount of Y protein in animals by administering phenobarbital, with a corresponding increase in the liver's uptake capacity. Then, last summer, Dr. A. J. Levi and I, studying a mutant mouse population, found unusually large quantities of Y protein in a dwarf mouse whose defect in the anterior pituitary resembled hypophysectomy. Dr. Reyes took this finding a step further by observing that in hypophysectomized or thyroidectomized rats, the Y protein increased twofold. Moreover, when hypophysectomized or thyroidectomized rats were given phenobarbital, the amount of Y protein doubled again, so that the cytoplasmic fraction of Y, instead of being the normal 2% or 3%, increased to about 10%.

That the effects of phenobarbital were cumulative to those of hypophy-

sectomy or thyroidectomy suggested that the level of Y was regulated at least at two separate points. When thyroxine was administered to the thyroidectomized rat, the normal Y level was restored – thus establishing a role for thyroxine in hepatic uptake.

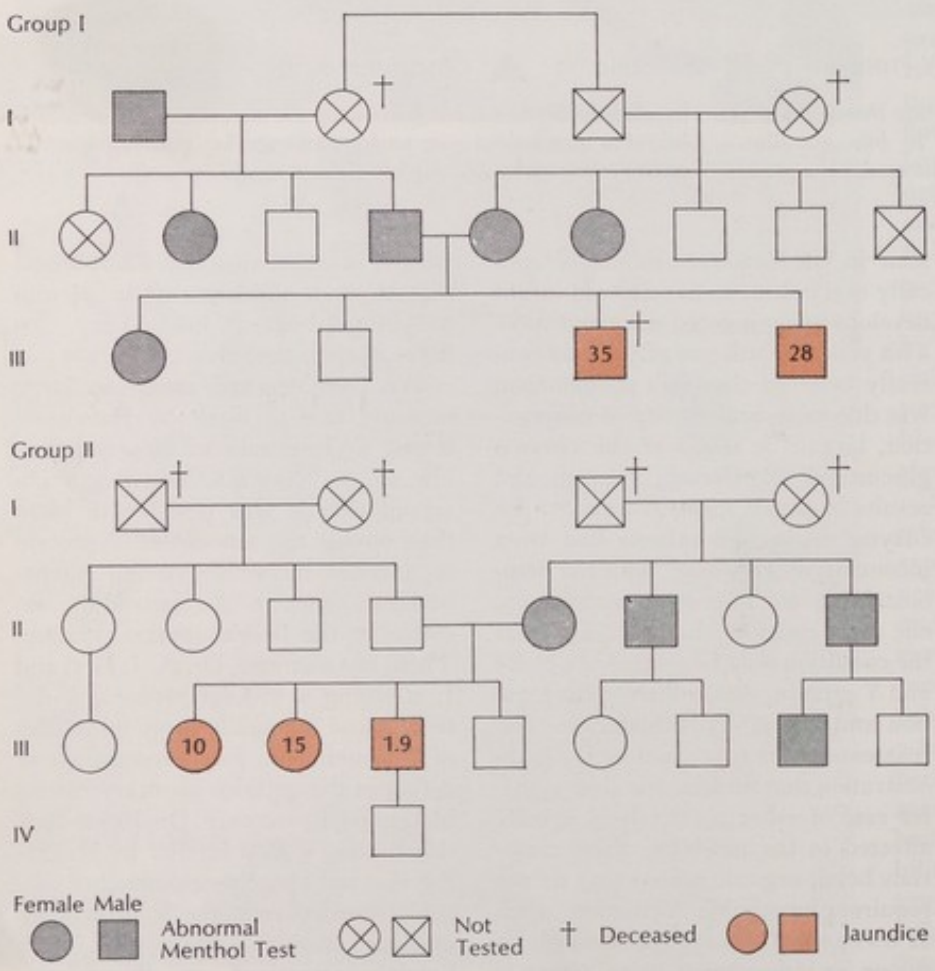
Important pharmacologic implications are contained in this example of a cytoplasmic protein related to uptake that can be induced by drugs and whose steady-state level is modulated by an endogenous hormone.

However, lest it be prematurely supposed that these proteins hold all the answers to understanding hepatic uptake, consider that Y and Z appear normal in the mutant Southdown sheep, which has an inherited defect in hepatic uptake of several organic anions. Obviously, Y and Z are not the only factors determining uptake. While their function may be to bind organic anions in one phase of the uptake process, these organic anions must first cross the plasma membrane, and their passage is most certainly dependent on the status of the membrane. Our guess is that further study of the Southdown sheep will broaden our understanding of uptake by yielding more information on the role played by the membrane.

Phenobarbital also plays a dramatic and as yet incompletely understood role in relation to the membranes of the endoplasmic reticulum. We know that the drug induces a proliferation of these membranes and enhances some of their metabolic activities including bilirubin glucuronide formation. It also increases nonerythropoietic sources of bilirubin synthesis, such as cytochrome P450, a heme protein involved in the microsomal electron transfer system and drug metabolism. This versatility of effect on membranes, added to phenobarbital's effect on Y and Z in the cytoplasm, poses some tangled pharmacologic and physiologic questions. Evidently the drug may work at many levels (or at one level that incites effects in others) to accelerate bilirubin transport at practically every step of the way.

Thus, it seems probable that the chain of events involved in uptake, conjugation, and excretion is more complex than we supposed. What may initially appear straightforward often becomes more complicated after further study. For example, for a while we thought that inherited glucuronyl transferase deficiency was a single entity. The disorder, known since 1952 as the Crigler-Najjar syndrome, now appears to be two entities that are clinically similar but genotypically different. Type I tends to be clinically more severe and is transmitted as an autosomal recessive characteristic. Type II seems less damaging – many patients survive with little or no apparent brain damage – and is probably transmitted as an autosomal dominant trait. Type I patients have no bilirubin glucuronide in their bile, which is colorless, while in type II patients bile is normal in color and contains a small amount of bilirubin glucuronide.

But the major and most dramatic distinction between the two is in their response to phenobarbital. Type I patients show no change in serum bilirubin concentrations while type II patients rapidly become anicteric, their serum bilirubin dropping from levels as high as 20 mg% to 2 mg% in 10 days, with corresponding favorable cosmetic results. With discontinuation of the drug, their jaundice returns. The mechanisms underlying the different response in these two types to phenobarbital are still unknown. The status of Y and Z proteins in these patients has not been studied. That individuals with type I disease are never found in families that produce type II progeny makes it very unlikely that type II is a heterozygote of the same gene defect of which type I is the homozygote. They may represent differences in the composition or structure of glucuronyl transferase or



Genetic heterogeneity of chronic nonhemolytic unconjugated hyperbilirubinemia with glucuronyl transferase deficiency is supported by inheritance studies. Trait for the group I disorder is transmitted as an autosomal recessive; its expression is favored by consanguinity, as in the pedigree shown at top. Group II appears to be transmitted as an autosomal dominant. Deficiency of glucuronide formation is evaluated by menthol tolerance test, in which urinary menthol glucuronide excretion is measured after ingestion of 1 gm of l-menthol (white symbols indicate normal test results). Numbers state serum bilirubin concentrations in cases where these exceeded 1 mg%.

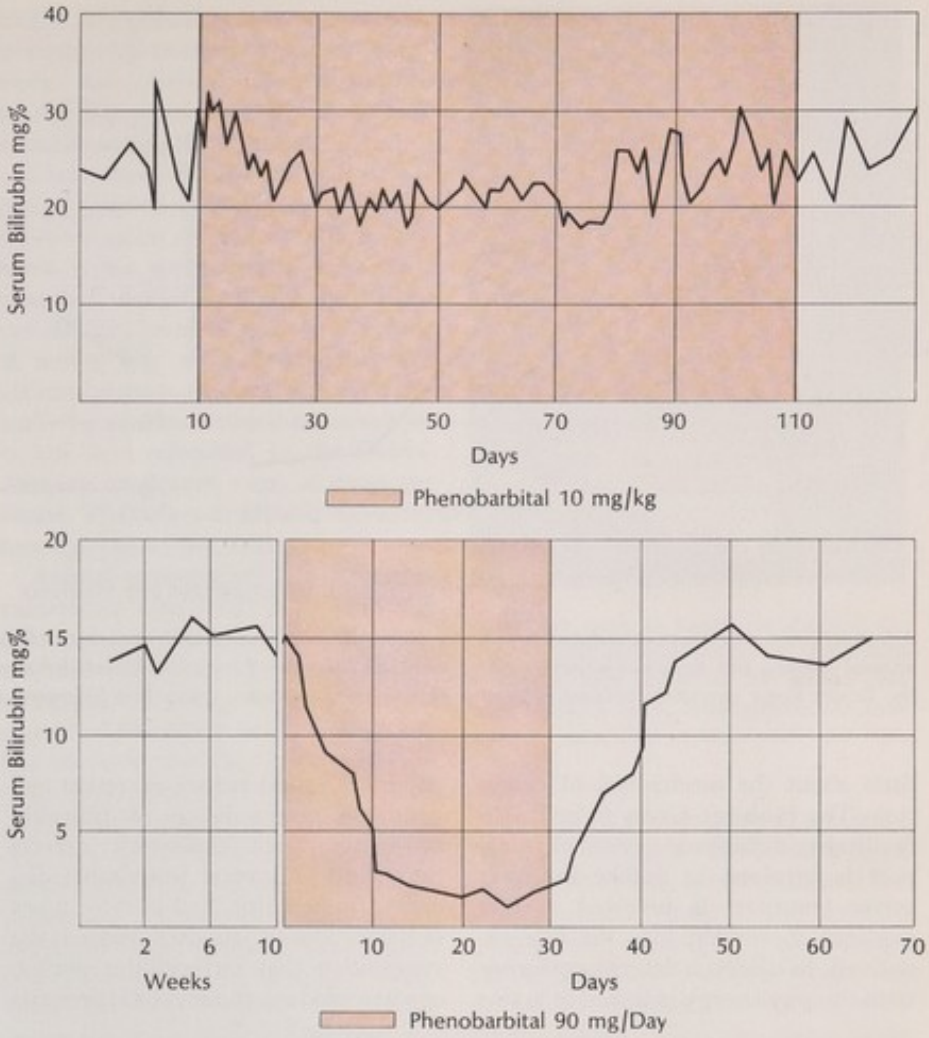
the difference may be at a higher genetic level, involving a regulatory gene.

One cannot assume that a single etiology is responsible for what manifests itself as glucuronyl transferase deficiency. Indeed, it may very well turn out that the controversy concerning the etiology of physiologic jaundice of the newborn – transferase or Y protein deficiency – will be resolved by finding both may be important.

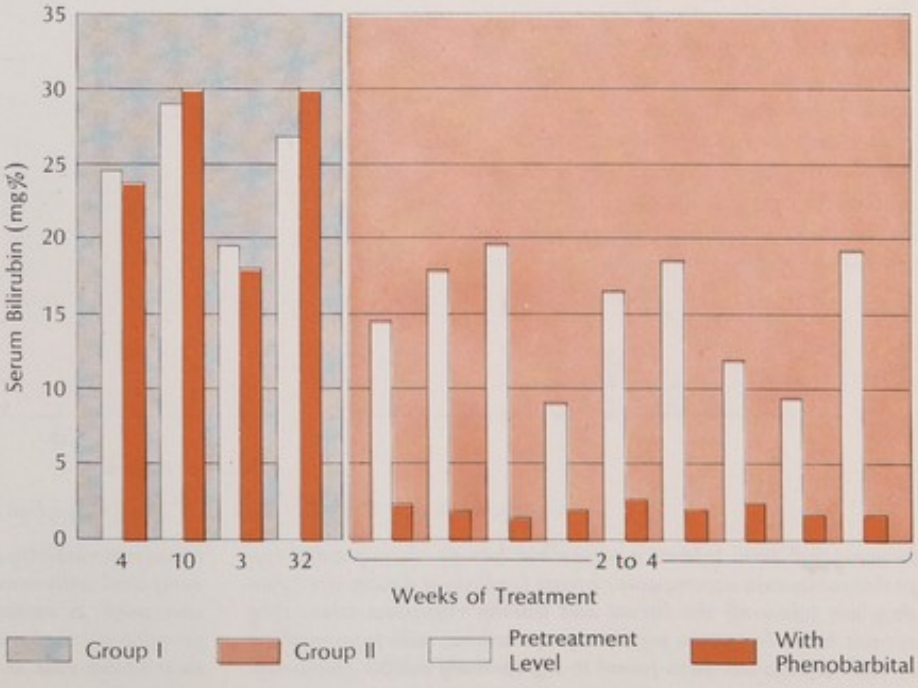
There are other types of glucuronyl transferase deficiency in which inheritable and maternal factors play a part. A small portion of nursing mothers with a probably inherited defect in steroid metabolism excrete in their breast milk an unusual isomer of pregnanediol, pregnane-3-alpha, 20-beta-diol, which has been found to inhibit glucuronyl transferase activity in vitro. Their infants manifest prolonged but benign jaundice from the first to the fourth week of life as long as they are breast-fed. Another group of women, otherwise healthy and normal, regularly give birth to severely jaundiced infants who run a high risk of kernicterus. Examination of serum from these mothers reveals an unidentified factor, present mainly in the last trimester of pregnancy and gradually disappearing after delivery, that inhibits glucuronyl transferase activity in vitro.

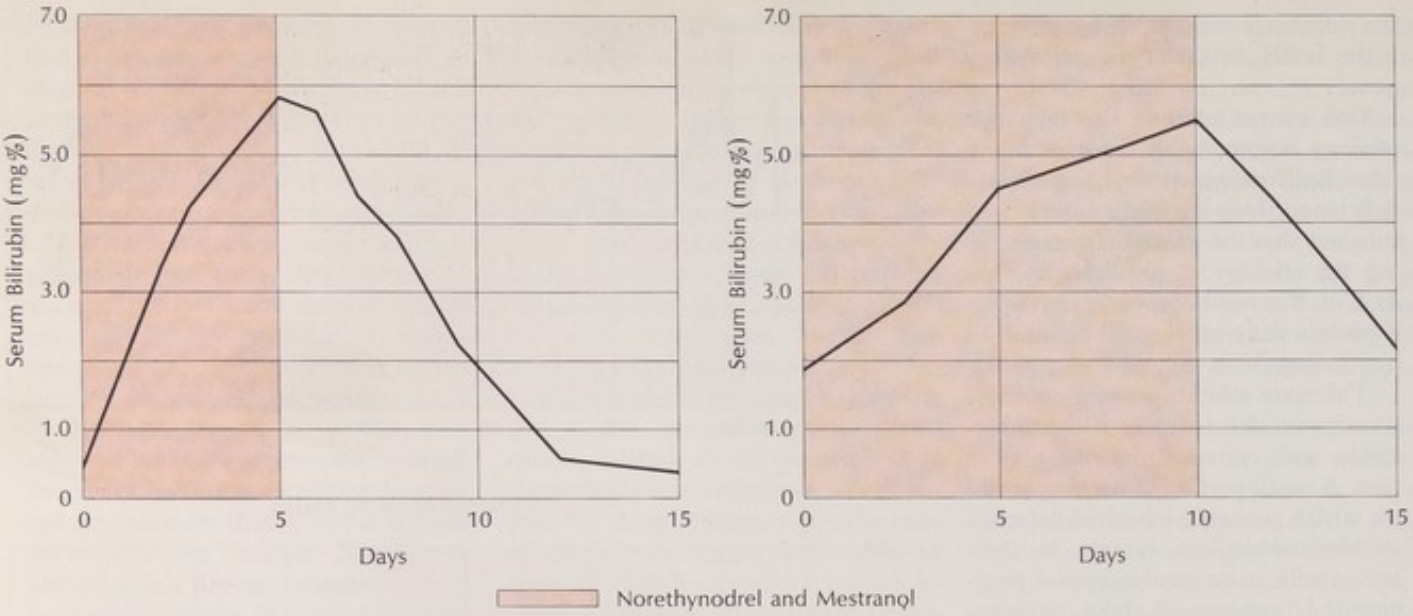
That we know what we do about disorders of bilirubin conjugation – indeed, that we are even aware that conjugation is important in bilirubin metabolism – is in large part due to the existence of a mutant strain of animals, the Gunn rat, that has provided us with a model for experimentation and analysis. These rats have an inherited glucuronyl transferase deficiency similar to the type I syndrome in humans. Their bile is colorless; they cannot conjugate bilirubin, yet they can excrete administered conjugated bilirubin in the bile. Many Gunn rats either succumb to kernicterus as neonates or survive with severe brain damage and are important in the study of the pathogenesis and prevention of brain damage in unconjugated H B R. These mutants made it possible for the first time to separate uptake, conjugation, and excretion.

While we have learned more about the mechanisms of uptake and conjugation, we have learned comparatively



Group I and II patients differ sharply in their response to phenobarbital, a phenomenon tending to confirm their genetic heterogeneity. Above are representative data for a patient in each group. While the more severe hyperbilirubinemia of the group I patient (top) shows no consistent change over a long period, the group II patient's response to phenobarbital is immediate and persists as long as drug is given. Below: similar contrast is seen in graphs summarizing phenobarbital response in patients from each group.

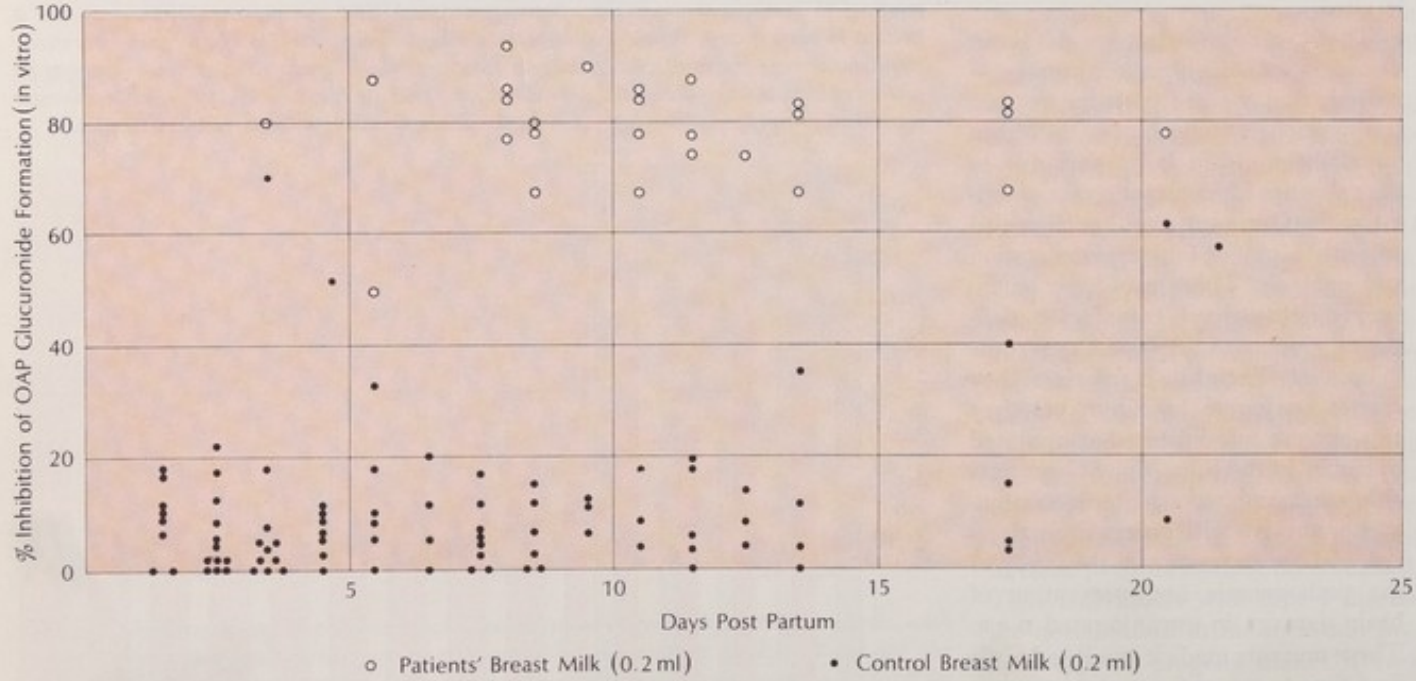




Administration of estrogens reduces hepatic excretory function in normal women but does not result in overt jaundice because of the liver's large excretory reserve capacity. In contrast, women with recurrent cholestasis of pregnancy (graph at left) and those with Dubin-Johnson syndrome (graph at right) manifest overt jaundice following administration of oral contraceptives.

little about the mechanism of excretion. The evidence seems to be that a facilitative diffusional system of transport is involved in uptake but that active transport is involved in bile pigment movement into the bile canaliculi. In theory, a defect interfering with the way energy is linked to active transport would reduce excretion and result in accumulation of bilirubin conjugate. Such cholestatic defects are found in several inheritable disorders, in hepatitis, and in some types of drug-induced jaundice. Indeed, the recognition that intracellular pathology can produce cholestasis represents an advance over theories explaining it all as due to mechanical blockage.

One inheritable human manifestation of impaired excretory function is the Dubin-Johnson syndrome, a benign condition characterized by mild, lifelong conjugated HBR. The livers of Dubin-Johnson patients are found



In some full-term infants a severe but benign unconjugated hyperbilirubinemia accompanies breast feeding; it disappears when they are taken off the breast and usually reappears when they are put back. In guinea pig liver preparations, milk from mothers of such infants has been found to significantly inhibit glucuronyl transferase activity (formation of O-aminophenol glucuronide) as compared with results observed when milk from control mothers was used. A maternal defect in steroid metabolism has been postulated; these mothers excrete pregnane-3(α),20(β)-diol in their breast milk as well as in their urine.

on autopsy to be black, probably the result of accumulation of oxidized epinephrine metabolites and other substances. The impairment of the ability to transfer certain organic anions from liver to bile extends to BSP, porphyrins, and radiographic materials. Yet these persons are otherwise healthy and appear to process bile salts normally. This observation contradicts the general assumption that bilirubin and bile salts are excreted by the same pathways.

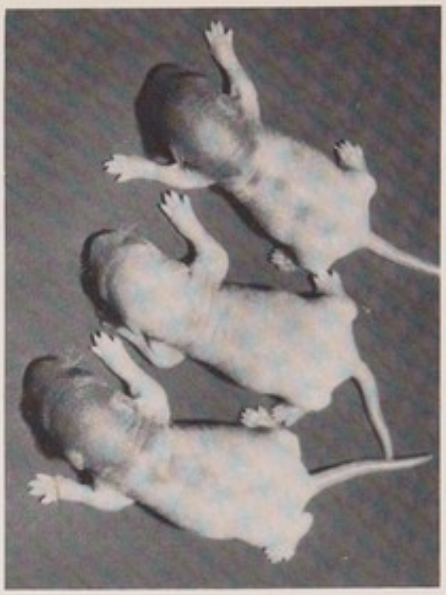
Opportunities for more detailed studies of the phenomenon have come through the discovery of mutant Corriedale sheep in which a seemingly identical disorder occurs. The sheep were detected because of the photosensitivity characteristic of liver defects in ruminants. Their livers, also, were black. Experimental studies revealed that the transport defect is limited to certain organic anions like bilirubin, BSP, and radiographic dyes, while other bile components, particularly bile acids, are unaffected. This fact, in man and sheep, strongly suggests that more than one mechanism operates in hepatic excretion of organic anions.

The jaundice of many Dubin-Johnson patients is kept at borderline or sub-borderline levels because of the liver's large reserve capacity of excretion. In normal women the metabolism of estrogens may further reduce hepatic excretory function, especially in the last trimester of pregnancy, or if oral contraceptives are used. In women with the Dubin-Johnson syndrome, estrogens, pregnancy, or oral contraceptives produce overt jaundice, entirely innocuous except cosmetically, and reversible when the estrogen level subsides. Unfortunately, some obstetricians and surgeons are not alert to this possibility and occasionally make incorrect decisions on the assumption of obstructive jaundice. More commonly, internists fail to make the correct diagnosis and patients are incorrectly treated for years as cases of chronic active hepatitis or cirrhosis.

Cholestasis, or bile secretory failure, is related to impairment of the total excretion of bile, of which bilirubin is but a component. It is characterized morphologically by dilated canaliculi, chemically by an increase in serum bile acid, cholesterol, and alkaline phosphatase activity, and clinically by pruritus related in some way to bile acid retention. In the Dubin-Johnson syndrome, none of these are found. The defect in hepatic excretory function spares bile acids.

Animal mutants with spontaneous cholestasis would facilitate study of bile secretory failure, but they have not yet been found. Cholestasis due to mechanical blockage by gallstones or cancer is relatively well understood, but we have very far to go before we can comprehend the forms of cholestasis that appear to involve bile secretion or whose elucidation lies in pathways of bile acid metabolism and excretion. Two possibly hereditary varieties in man provide some clues. One is the tendency of certain women, often in certain families, to develop cholestasis repeatedly in the last trimester of pregnancy. Another is the rare recurrent familial cholestasis, in which patients suffer repeated attacks of cholestasis without discernible reason. That this affliction may have a genetic basis is suggested by the fact that almost all cases have been found among highly inbred populations.

The limitations of experiments with humans are obvious, and that is one of the reasons our knowledge of the mechanisms responsible for cholestasis is limited. A mutant animal would be helpful. Such an animal model may exist without anyone having thought to associate it with this line of research. After all, the Gunn rat was discovered in the 1930's and initially utilized for a more general study of the heredity process and modes of transmission. Our understanding of metabolism was not ready to utilize this animal until 20 years later, when the glucuronide conjugation of bilirubin was identified and a search begun for models in which deficiency in the conjugation process could be studied. Because of the conscientiousness of a department secretary who hated to dispose of the animals, and because of the scientific devotion and integrity of geneticist William E. Castle, who felt it his responsibility to keep mutational lines alive in the possibility of some day learning more about them, the Gunn rats were still available. Breeding went into high gear and colonies of Gunn rats are now in laboratories all over the world.



Gunn rats serve as model for study of inherited glucuronyl transferase deficiency similar to that found in the human disease. Middle rat is homozygous for trait and is jaundiced. Littermates are not.

13

Alpha-1-Antitrypsin Deficiency

HARVEY L. SHARP
University of Minnesota

Since the turn of the century it has been known that normal human serum inhibits activity of proteolytic enzymes. Some 15 years ago the inhibiting factor was localized chiefly to the alpha-1-globulin fraction and more recently, around 1962, the chief serum component was designated as alpha-1-antitrypsin, a low-molecular-weight glycoprotein with an electrophoretic mobility between albumin and alpha-2-macroglobulin. Approximately 90% of the trypsin inhibitory capacity of human serum could be attributed to activity of this factor. As an antienzyme it soon proved to have a wider range of activity than its name would suggest: Alpha-1-antitrypsin was a strong reactor with elastase, skin collagenase, chymotrypsin, plasmin, and thrombin. It also showed inhibitory activity against fungal and leukocytic proteases but could be inactivated by bacteria, particularly gram-negative organisms.

However, knowing the chemical properties of alpha-1-antitrypsin did not help much in elucidating its biologic functions, which, to this day, remain incompletely defined. The possibility of a protective role – that it neutralizes release of enzymes into body fluids by dying cells – has been suggested by the finding of increased levels of alpha-1-antitrypsin during inflammation and severe infection, when the concentration may rise to three times normal values. Increases with hormonal stimulation, as in pregnancy and with use of contraceptive medication, are easily demonstrated but still to be explained.

But if the presence of this antienzyme is not well understood, its absence is now known to have important and untoward consequences. Thus far two quite disparate disease states, one in adults and the other in children, have been associated with an inborn metabolic defect resulting in a serum deficiency of alpha-1-antitrypsin. The first is a type of pulmonary emphysema first identified around 1963 in Sweden that has since been observed and investigated elsewhere as well. The second, identified much more recently by our group in Minneapolis, is a type of liver cirrhosis that we and others are now intensively studying. Unfortunately, both are usually progressive diseases and, despite attempts at definitive therapy, the outcome is grave.

A deficiency in alpha-1-antitrypsin is readily detectable on serum electrophoresis by the absence of a normally visible alpha-1-globulin band. Swedish workers Carl-Bertil Laurell and Sten Eriksson were routinely surveying serum electrophoretic patterns among hospitalized patients when they discovered a then-new type of dysproteinemia, identifiable as alpha-1-antitrypsin deficiency. Five such cases were found among 1,500 sera they examined over a six-month period. The sera were reexamined by both paper and agarose gel electrophoresis; again the alpha-1-globulin band, usually sharply demarcated, was missing from the electrophoretic pattern. Immunologic studies with antiserum specific for alpha-1-antitrypsin revealed trace amounts present in the serum but electrophoretic mobility was reduced, suggesting a structural abnormality.

When Laurell and Eriksson checked on the clinical condition of the five patients, to their surprise they found that three had severe pulmonary disease; so did the sister of one of the three. The findings clearly suggested "some connection between the degenerative pulmonary disease and alpha-1-antitrypsin deficiency." Family studies were indicated to determine if an inherited defect might underlie this apparent relationship.

Pursuing a detailed genetic and clinical investigation, Eriksson examined serum electrophoretic patterns among relatives of one of the three patients with lung disease and observed three distinct patterns of antitrypsin activity: normal; intermediate, or about 60% of normal; and low, or about 10% of normal. This suggested a recessive type of inheritance for alpha-1-antitrypsin, but a larger family survey was needed. Eriksson next studied relatives of 14 patients with serum alpha-1-antitrypsin deficiency, 213 subjects in all, this time using a technique for determining serum trypsin inhibitory capacity. The method used en-

tails measurement of inhibition when a standard amount of trypsin is added to a given volume of serum; inhibition level is indicated in milligrams of trypsin per milliliter of serum. A variety of substrates can be used for estimating the amount of trypsin not complexed by serum inhibitors. The substrate Eriksson chose (we have done likewise) was benzoyl-arginine-para-nitroaniline; release of para-nitroaniline by action of trypsin on the substrate is measured colorimetrically.

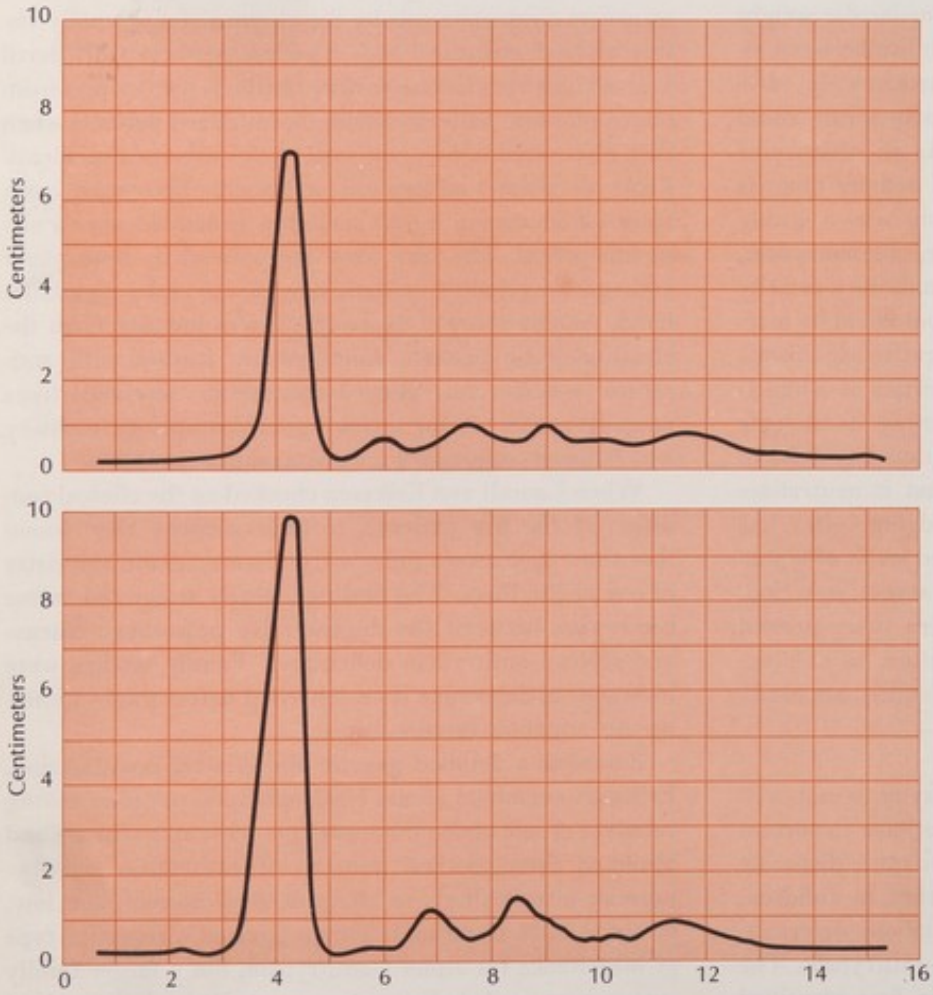
In Eriksson's family studies the assays of antitryptic activity confirmed the trimodal distribution of the serum enzyme: For relatives with a normal antitrypsin level, mean trypsin inhibitory capacity was 1.10 mg/ml; among those with a deficiency, the mean was 0.24 mg/ml; most showed an intermediate inhibitory capacity, the mean here being 0.67 mg/ml. All data seemed to indicate that severe alpha-1-antitrypsin deficiency occurred in individuals homozygous for the genetic defect and that intermediate levels indicated heterozygosity. Autosomal recessive inheritance was suggested by the sex distribution of the deficiency; in several families the defect had evidently been transmitted from father to son, which excluded the possibility of X-linked inheritance.

Eriksson and coworkers tried to further clarify the relationship of alpha-1-antitrypsin deficiency and pulmonary disease by studying a group of 33 hospitalized individuals with the antienzyme defect. Twenty-three showed definite evidence of pulmonary emphysema; however, theirs was clearly an unusual emphysema in several aspects: A higher than usual proportion of the patients were women, and age at onset was remarkably low; 60% of the patients were under 40 and 90% under 50. In nearly half the cases exertional dyspnea had not been preceded by chronic bronchitis, again an atypical finding. Moreover, the emphysema was invariably more pronounced in the lower lobes rather than distributed throughout the lung. X-ray evidence of lower-zone disease was obtained in all 23 cases. From Eriksson's observations, individuals homozygous for the serum protein defect seemed definitely predisposed to this unusual form of pulmonary emphysema in early middle age; heterozygotes appeared to run no increased risk.

To estimate the frequency of the abnormal gene in the Swedish population he surveyed an entire village, analyzing a total of 6,995 serum samples by their electrophoretic patterns and antitryptic activity. On this basis he calculated the frequency of heterozygotes to be 1 in 21.

Further genetic studies by Magne Fagerhol et al., working in Oslo, indicated that an autosomal codominant rather than recessive mode of inheritance may control the genetic defect. Their important contribution was recognizing the presence of several genetic variants of alpha-1-antitrypsin; an improved acid starch gel electrophoretic technique permitted higher resolution in the prealbumin zone and revealed alpha-1-antitrypsin variants not detectable by other methods. Fagerhol and, later, others further identified the variants by use of antigen-antibody crossed electrophoresis and specific rabbit antiserum against human alpha-1-antitrypsin. To date Fagerhol has identified a total of 11 alleles of the principal gene for alpha-1-antitrypsin, presumably determined by a series of codominant genes. No exception to this mode of inheritance has been found.

Collectively the inherited alpha-1-antitrypsin variants constitute the Pi system, so named because the antienzyme is the major protease inhibitor of human serum. Pi allele products can be differentiated by their electrophoretic band spectra, varying protein content of zones in the pattern, and their alpha-1-antitrypsin content. The principal allele for alpha-1-antitrypsin is termed Pi^M; other alleles were detected by the relative electrophoretic mobility of their proteins in reference to Pi^M and are indicated by other letters of the alphabet. Thus Pi^F is a fast-moving variant, Pi^X a slow one; the slowest-moving variant, desig-



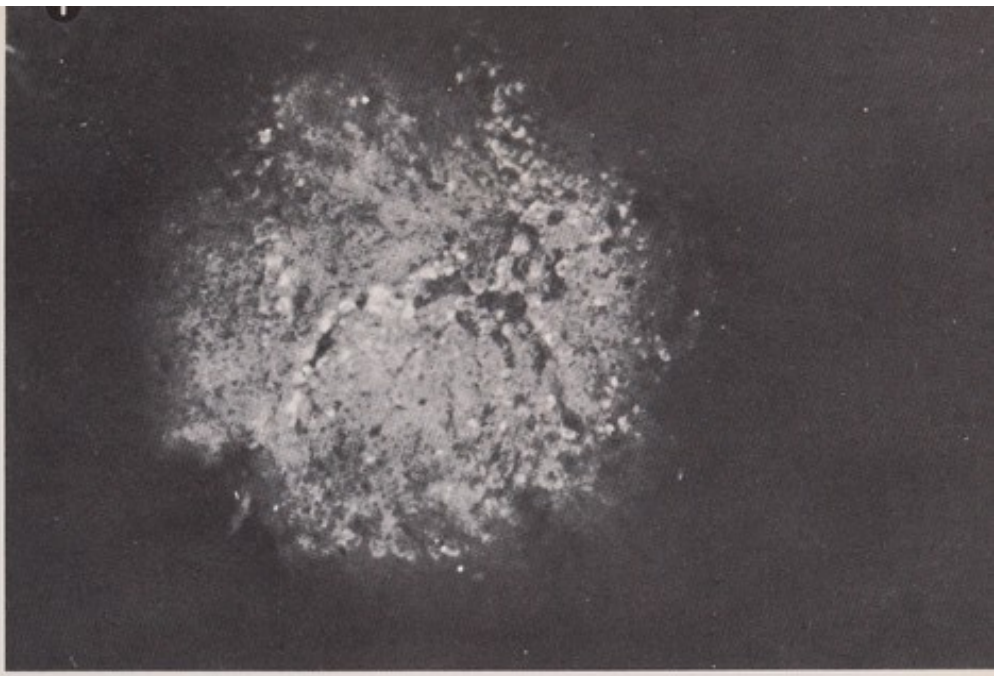
Alpha-1-antitrypsin deficiency is readily detectable by protein electrophoretic patterns. Normal (top) shows alpha globulin peak immediately following high albumin peak. The alpha globulin peak is absent in patient with deficiency of enzyme inhibitor (bottom).

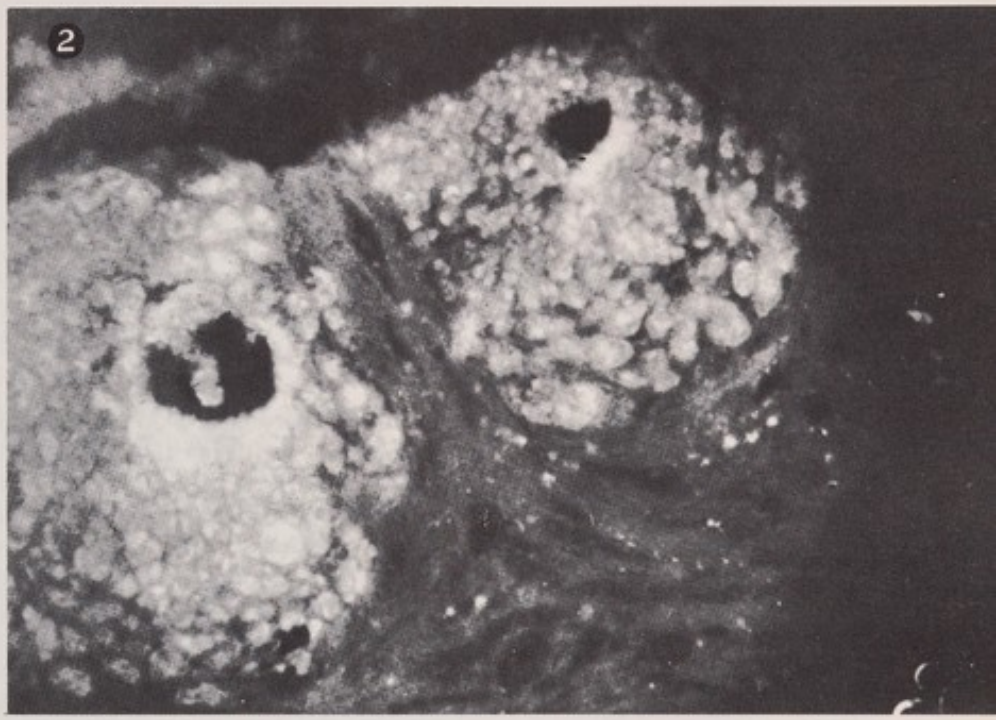
nated Pi^z, has been the one associated with the marked deficiency state.

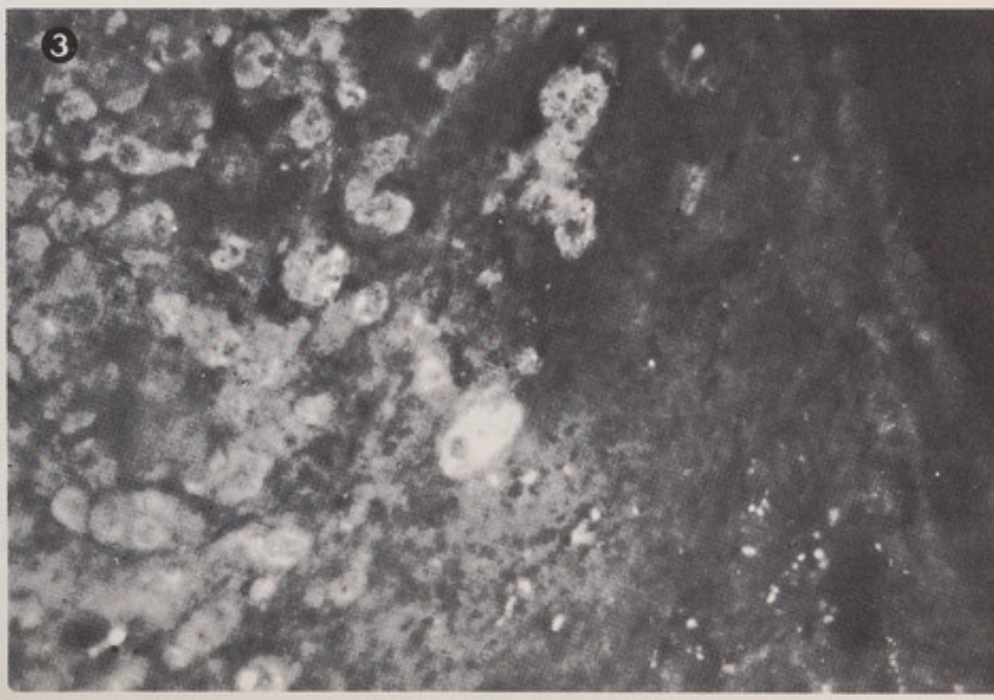
As Fagerhol's work showed, among subjects heterozygous for the deficiency gene (i.e., Pi phenotype M Z), serum levels of alpha-1-antitrypsin are roughly 60% of normal. In subjects homozygous for the deficiency gene (Pi phenotype Z Z), serum levels are between 10% and 20% of normal. These findings are consistent with autosomal codominant inheritance, where each allele is expressed and its product is summated to give the total serum level.

The molecular structure of the several variants of alpha-1-antitrypsin has not been well worked out; the Pi types could result from simple amino acid substitutions caused by point mutations or the substitution of different carbohydrate molecules, but this is not yet known. Actually, it has not been established whether the alpha-1-antitrypsin molecule consists of one or more polypeptide subunits; available genetic data are compatible with the existence of multiple alleles at a single locus, each allele possibly determining a different polypeptide chain.

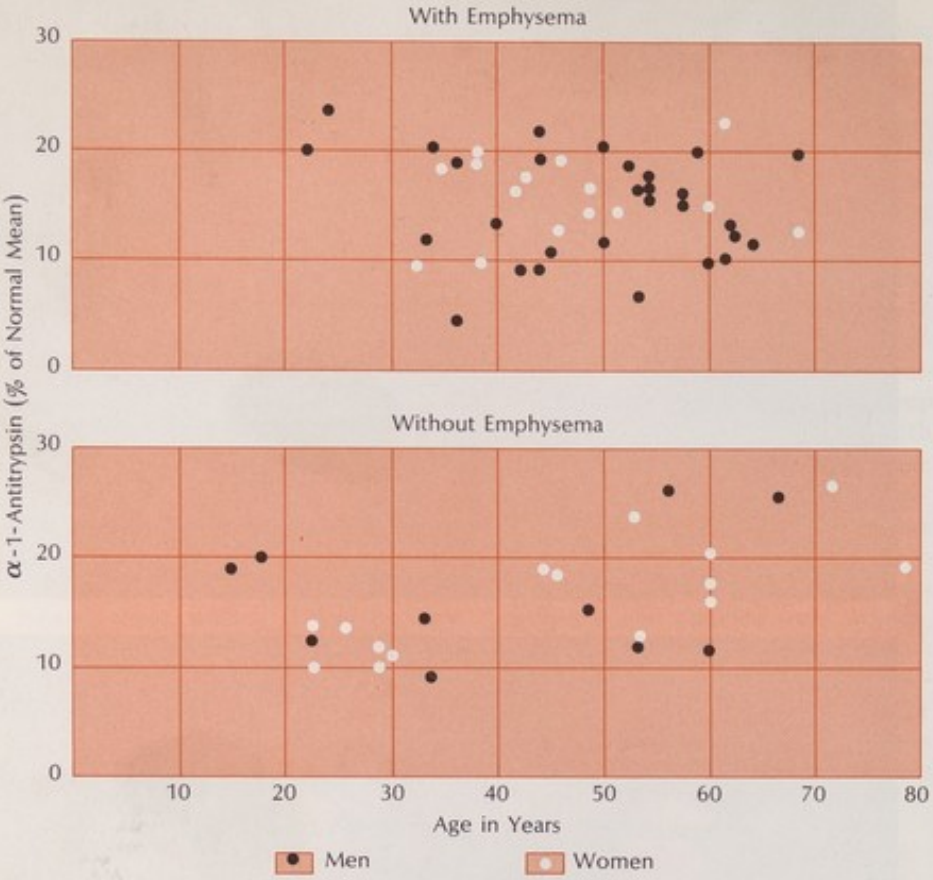
From electrophoretic and immunologic studies on some 13,000 subjects, Fagerhol and coworkers identified 70 persons having Pi phenotype Z Z. Checking into the clinical background, they also, like Eriksson before them, found a high prevalence of chronic pulmonary emphysema. Likewise, they also observed that not all individuals with genetically determined alpha-1-antitrypsin deficiency developed the lung disease. Indeed, among the 70 patients with Pi phenotype Z Z, 44 were found to have emphysema; the other 26 did not. Interestingly, serum levels of alpha-1-antitrypsin were no lower in subjects with emphysema than in those without. This suggested that the expression of the genetic defect in clinical illness







In children with familial cirrhosis associated with alpha-1-antitrypsin deficiency, periportal hepatocytes show marked positive reaction following application of fluorescent antibody against alpha-1-antitrypsin. Fluorescence may be sharply localized as in specimen (1) above; in some patients there is more diffuse involvement affecting virtually all hepatocytes within a lobule. An example of the latter is shown at low power (2) and at higher power (3).



Severe alpha-1-antitrypsin deficiency is not necessarily evidenced as overt illness; thus comparable serum levels are seen in subjects of Pi phenotype ZZ whether they have familial emphysema (upper graph) or are free of the pulmonary disease (lower graph).

may depend on interaction with other genes or perhaps exposure to some environmental agent, or other factors.

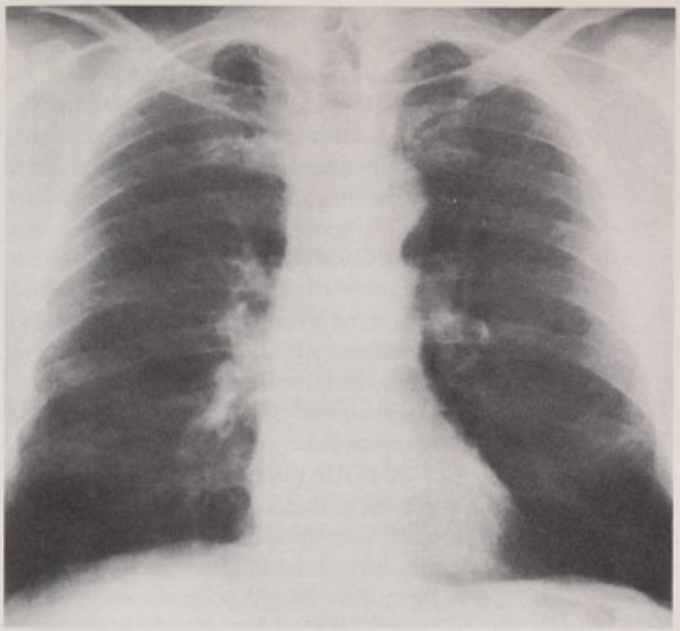
As the Swedish work on the relation of alpha-1-antitrypsin deficiency to pulmonary emphysema became known, investigators elsewhere sought to confirm it. The first example outside of Scandinavia was reported in 1966 by Friedrich Kueppers and coworkers while in New York; their patient was a 53-year-old man who had been ill with pulmonary disease since age 48. He and a brother both proved markedly deficient in serum alpha-1-antitrypsin; his three sons had half the normal amount. As other workers reported additional cases, the destructive pulmonary disease associated with severe alpha-1-antitrypsin deficiency became well characterized. An association between the serum protein defect and early onset panacinar emphysema was firmly established, although the question of a causal relationship remained unresolved.

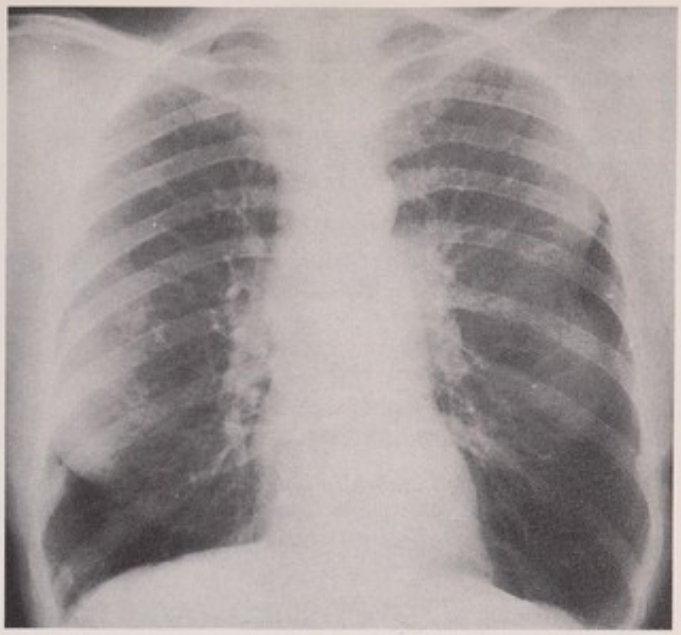
Also unresolved was the possibility of an association between an intermediate serum protein deficiency and the lung disease. Several investigators undertook to define this relationship more clearly. For example, Martin Welch and coworkers in Oklahoma City assessed serum antitryptic activity in 51 presumably healthy adult blood donors and 146 consecutive unselected patients attending a chest clinic. Three in the healthy group were found to have intermediate serum levels of alpha-1-antitrypsin, as did 17 in the group with pulmonary disease, a difference not found statistically significant. Turning then to patients in whom the deficiency had been identified, they studied 18 with intermediate and 13 with severe deficiency with respect to pulmonary disease. All 18 patients with intermediate levels were men; 10 of the 13 with severe deficiency were women. In the latter group the lung disease almost always took the form of primary emphysema with minimal chronic bronchitis; in contrast, a wide assortment of pulmonary diseases affected the patients with intermediate antitrypsin levels.

X-ray and lung scan findings also differed sharply. Characteristically, in the severely deficient group, x-rays showed diffuse loss of vascular markings over both lower lung fields; upper lobe vessels tapered and branched normally and extended to the lung periphery. These findings were confirmed on lung scans, which showed greater perfusion of upper zones. In patients with intermediate antitrypsin levels, radiographic findings ranged from normal to extensive bullous disease. None of the x-rays indicated a pattern of bilateral decrease in lower zone perfusion; none of the lung scans showed bilateral lower-zone disease.

Similar differences obtained in a Boston study by Richard Talamo and coworkers of families of six patients with a combination of pulmonary disease and alpha-1-antitrypsin deficiency. Serum antitrypsin concentration and antitryptic activity were determined in 93 relatives and three additional subjects. Eight of 11 genetically homozygous relatives proved to have obstructive pulmonary disease; findings in a ninth suggested similar lung involvement. Among 34 family members with intermediate antitrypsin levels only one, a 77-year-old man, was found to have pulmonary disease, and he for years had been a heavy smoker. Thus the investigators could conclude that "no remarkable rate of occurrence of pulmonary disease was noted among either genetic heterozygotes or family members with normal alpha-1-antitrypsin function and concentration."

Admittedly, the question of increased risk to heterozygotes remains open. For example, Jack Lieberman and coworkers in California reportedly observed several heterozygous individuals with lung disease among relatives of an affected homozygote. This prompted them to survey patients with pulmonary emphysema at a Veterans Administration hospital in reference to serum alpha-1-antitrypsin deficiency. Of the 66 patients studied, 25.8% proved deficient in the serum protein; seven of these were homozygotes and 10 heterozygotes. The average age of homozygous deficient patients was 44 years; average age of the heterozygotes was 52. When only patients below the age of 50 were considered, 47.8% proved to have some degree of antitrypsin deficiency. In support of predisposition of heterozygotes to a similar type of emphy-

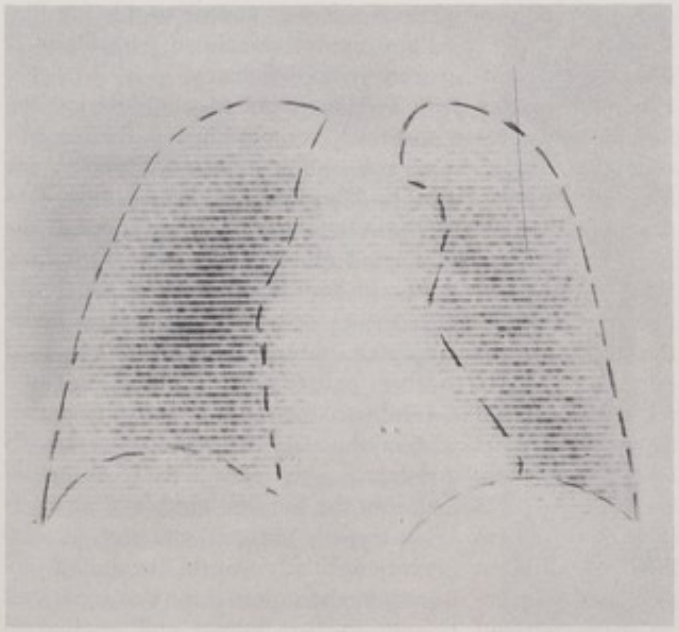


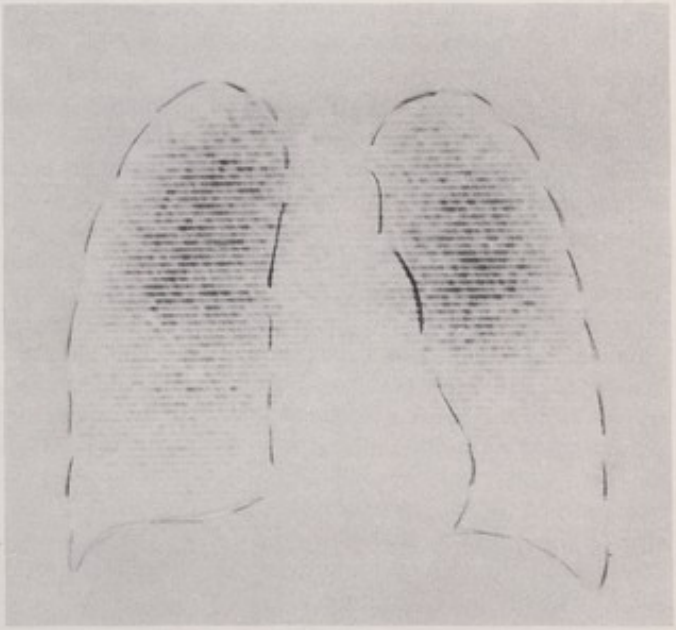


Radiographic findings differ sharply in 55-year-old man with bronchitis and intermediate antienzyme deficiency (left) and in his daughter with primary emphysema and severe deficiency; her x-ray shows loss of vascular markings over both lower fields.

sema, Stevens et al have found pathologic evidence of predominantly lower lobe involvement occurring in heterozygous individuals as well as in the homozygous deficient ones.

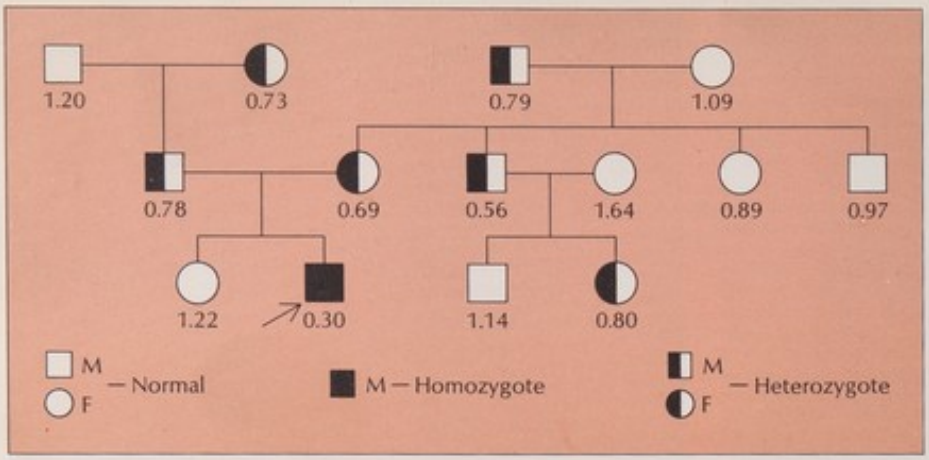
Surely, if individuals heterozygous as well as homozygous for antitrypsin deficiency are predisposed to chronic pulmonary emphysema, it would be important to launch appropriate detection and prevention programs on a public health scale. Disregarding certain national differences, the frequency of heterozygotes in most populations appears to range between 1% and 5%; thus a considerable number of individuals might be at risk of developing pulmonary emphysema at an early age. Since an interaction of genetic and environmental factors is believed to account for increased susceptibility to lung disease in alpha-1-antitrypsin-deficient persons, it might be crucial to identify them early so that at least potential environmental hazards might be controlled. However, although the issue is not yet fully resolved, the evidence obtained by most investigators indicates that the heterozygote is apparently not unduly subject to the familial emphysema; the reported increase in risk in a few stud-





Lung scans confirm preservation of normal lower-zone perfusion in father with moderate alpha-1-antitrypsin deficiency and severe lower-zone vascular decrease in daughter with marked deficit.

Scans of severely deficient patients consistently show bilateral decrease in lower-zone perfusion, not seen in patients with pulmonary disease and intermediate alpha-1-antitrypsin deficit.



Schematic shows family pedigree of homozygous proband (arrow) with cirrhosis and severe antienzyme deficiency. Low trypsin inhibitory capacity (0.30 mg/ml) of his serum contrasts with intermediate and high values in heterozygous and normal relatives.

ies may relate to the selected population examined.

From our work on alpha-1-antitrypsin deficiency in liver disease we have concluded that only in the severe deficiency characteristic of the homozygous state is there predisposition to cirrhosis. No cases have been related to intermediate antitrypsin levels without another genetic defect present. Our first hint of a link between a form of childhood liver disease and a genetic deficiency of alpha-1-antitrypsin came about five years ago at our hospital when Esther Freier was surveying serum electrophoretic patterns in patients with various diseases. Unexpectedly, among the first sera low in alpha-1-antitrypsin was that of a child with hepatic cirrhosis. Both the child and a younger brother who also had liver disease were shown to be alpha-1-antitrypsin deficient.

This discovery prompted a search among other families of children with cirrhosis of undetermined origin. Thus far 18 unrelated families have been identified and studied; data are presented on the first 13. In all, the proband was a child with liver disease associated with serum alpha-1-antitrypsin deficiency initially detected on cellulose acetate protein electrophoresis. In affected children, electrophoretic patterns in the first month of life already showed an alpha-1-antitrypsin deficiency, arguing for an inborn metabolic level. All the children proved genetically homozygous for alpha-1-antitrypsin deficiency; analysis of their sera by Fagerhol's method demonstrated a zz type of serum deficiency. (Other causes of hepatic cirrhosis were of course also looked for but could be excluded by laboratory studies and the clinical course.)

Relatives of the 13 children have been studied by the trypsin inhibitory capacity technique; all but four parents have had intermediate antitrypsin levels indicating heterozygosity (three were mothers on contraceptive medication; as noted earlier, this is associated with an increase in serum antitrypsin levels). Twenty of 31 siblings and 14 of 77 relatives have had low to intermediate antitryptic activity. Most homozygous siblings are also showing evidence of cirrhosis; interestingly, there is a homozygous grandmother with pulmonary emphysema.

Conceivably, in the affected children the serum protein deficiency could be the result of liver disease rather than possibly the cause. However, trypsin inhibitory capacity determinations on children with hepatic cirrhosis attributable to other causes (such as biliary atresia, cystic fibrosis, etc.) have shown normal to above-normal values. (Sera in 9 of 10 alcoholic cirrhotic adults were also in the normal range; the one abnormal result was in a patient whose son is heterozygous for alpha-1-antitrypsin deficiency.)

The clinical manifestations begin in the first year of life. These children all have hepatomegaly or hepatosplenomegaly, although in those without jaundice the organomegaly may be mild and overlooked. Those who have the obstructive jaundice pattern usually return to a normal bilirubin level by the sixth month of life but abnormalities in hepatic enzymes persist, as does the organomegaly. Gradually, other complications of cirrhosis manifest themselves, i.e., esophageal varices, ascites, and infections (most notably, pneumococcal infections). Not all affected children of a family follow the same clinical course, but none has recovered from the initial liver injury. In almost all, the disease progresses relentlessly, with severe portal hypertension developing as a result of their macronodular and micronodular cirrhosis. All have been Australia antigen–negative.

In our initial examination of liver specimens from these children by light microscopy no features peculiar to this form of hepatic cirrhosis could be distinguished. However, an interesting difference emerged when we examined these sections under the electron microscope. An unusual amorphous material was present within the lumen of rough endoplasmic reticulum, which was markedly dilated; no trace of the material could be found in the Golgi apparatus or lysosomes, or free in the cytoplasm. The material was not visualized in all hepatocytes examined; in two of the children, sectioning of multiple blocks of Epon-embedded tissue was necessary for detection. However, nothing resembling this substance could be visualized in any other form of liver disease; it was unique to this familial liver disease associated with alpha-1-antitrypsin deficiency.

To identify and characterize the material, we obtained a fluorescent-tagged rabbit antibody specific for alpha-1-antitrypsin and applied it to frozen liver sections from seven of the affected children. In all, bright fluorescent deposits were seen, especially in the cytoplasm of periportal hepatocytes. Liver specimens obtained from patients with other cirrhotic conditions or chronic active hepatitis showed no significant fluorescence. To determine the specificity of the reaction in the familial cirrhosis, alpha-1-antitrypsin antigen and the fluorescent antibody were incubated together; the supernatant was separated from the precipitate and applied to liver specimens from affected patients. Fluorescence was completely blocked by alpha-1-antitrypsin antigen, con-

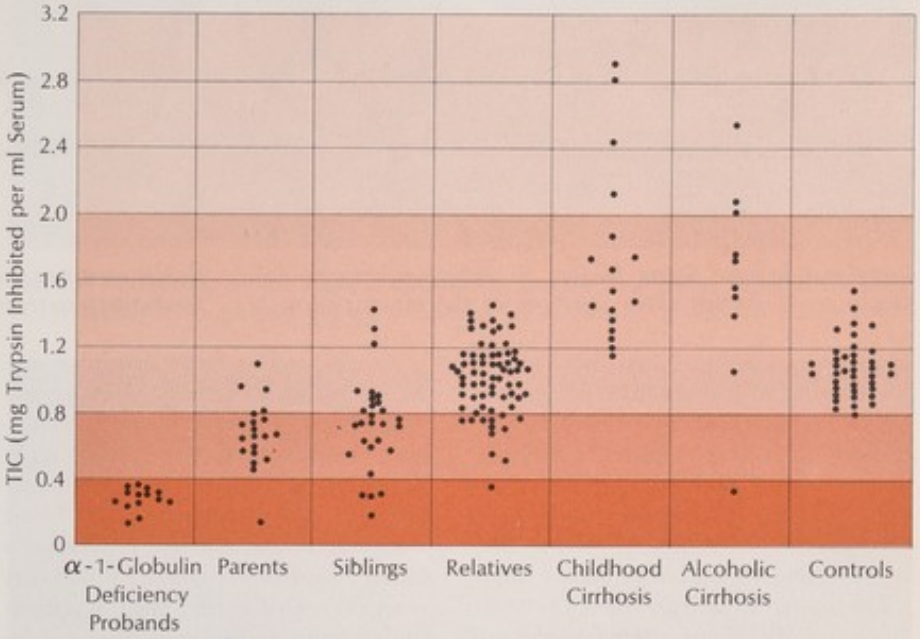
firming the specificity of the fluorescent antibody.

Having localized the positive fluorescence in these cases to the periportal areas of the liver, we returned to the light microscopic sections and examined them more closely. This time a distinct, clumped acidophilic material could be detected in H and E stained sections within the parenchymal cell cytoplasm. The material could be more readily visualized on PAS staining. When tissue sections were pretreated with either diastase or saliva to eliminate PAS staining of glycogen, PAS staining of the abnormal material persisted. Subsequently we have studied the livers of both homozygous and heterozygous individuals without cirrhosis and demonstrated the same material. Lieberman et al also confirmed these findings in Z phenotype emphysema patients.

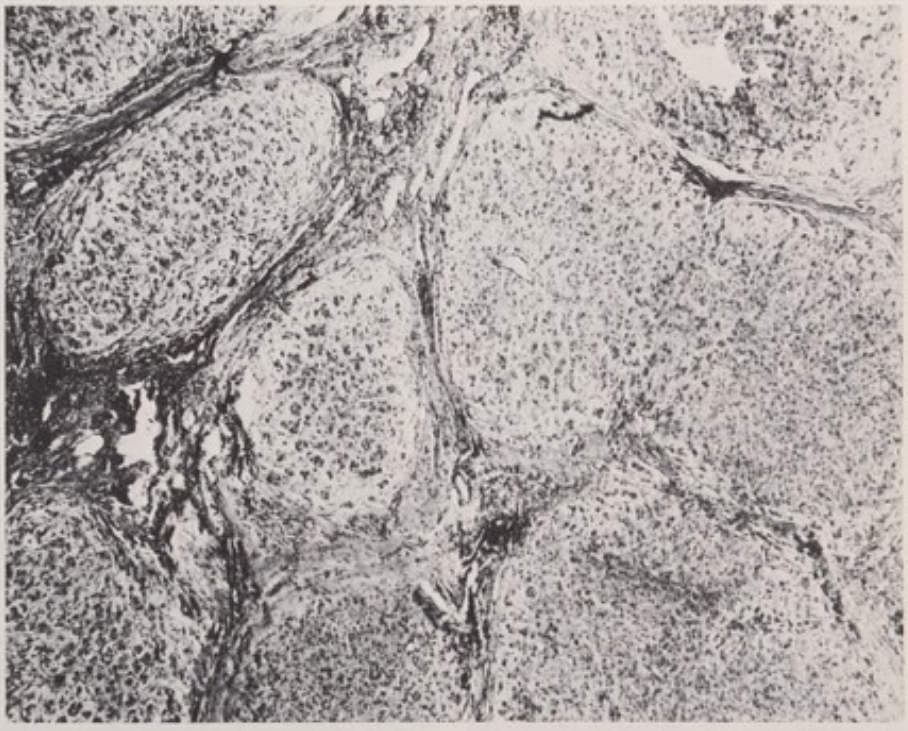
On both histochemical and immunologic grounds the material appeared to be some form of alpha-1-antitrypsin; more than that we could not say Moreover, the findings implied that large amounts must be present in the livers of these children with familial cirrhosis. Normally, the liver is the source of alpha-1-antitrypsin in the blood; as far as is known, it is the only organ that produces the antienzyme. When Dr. Henry Gans removed an animal liver, the serum trypsin inhibitory capacity halved in less than 24 hours.

Of course, a key question is why does alpha-1-antitrypsin accumulate in the liver in these cases? And to what extent does the persistence of the antienzyme in the liver bear a causal relationship to the hepatic damage? The situation clearly differs from that of other liver damage resulting from intrahepatic accumulation of abnormal material. In all other storage diseases, an enzymatic abnormality results in the stored material residing either in the cytoplasm or in lysosomes. In this instance the material is confined to the rough endoplasmic reticulum; the lysosomes and cytoplasm are unaffected. Thus, we lean toward the likelihood of either a structural abnormality in the glycoprotein that interferes with transport out of the liver or a defect in the transport system whereby a normal glycoprotein is produced but not carried to the blood. Recent studies have clarified mechanisms in hepatic transport of another serum protein, albumin, from the ribosomes, where it is synthesized at the rough endoplasmic reticulum, then transported via the smooth endoplasmic reticulum, the Golgi apparatus, and out of the liver. The same has not yet been done for alpha-1-antitrypsin but conceivably there is a defect of some kind early in the transport sequence. To be sure, the possibility that the abnormal presence of this material is causing the hepatic damage still cannot be ruled out, but this is doubtful since it is present in otherwise normal livers.

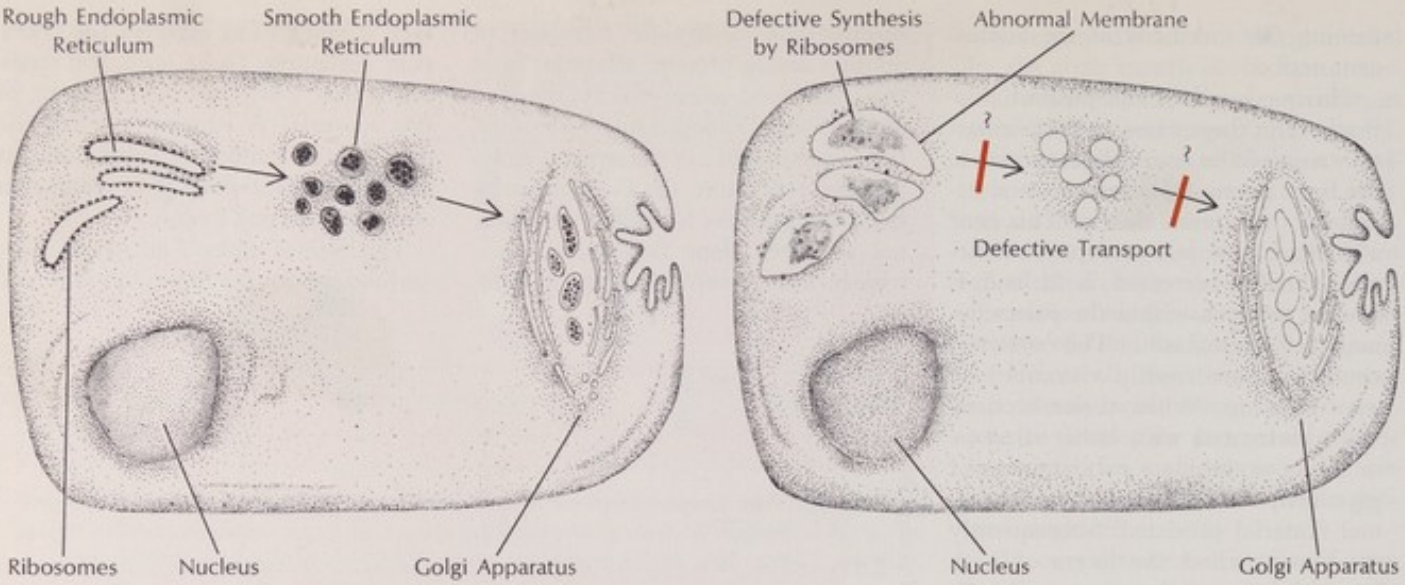
The type of alpha-1-antitrypsin accumulating in the liver has not yet



Assays of serum antitryptic activity show lowest values in children with familial cirrhosis and severe antienzyme deficiency, intermediate values in some parents and other family members, generally high values in patients with other forms of cirrhosis.



Great architectural distortion, with separation of hepatic lobules by fibrous connective tissue, is evident in biopsy specimen from child with familial cirrhosis.



Postulated hepatic defect leading to alpha-1-antitrypsin deficiency could involve either synthesis of the serum protein or its transport into the blood, thus interfering with normal hepatic mechanisms as recently delineated in studies of albumin.

been chemically characterized; we are beginning work on that now. On the basis of the immunogenetic studies it is known that the serum alpha-1-antitrypsin in the genetic deficiency state is usually of the Pi phenotype z z. But is the same alpha-1-antitrypsin variant present in the liver, or is the antitrypsin in the liver altered to the z z type when it reaches the plasma? This we do not know except to say that it is a slow moving glycoprotein. The parents of our cirrhotic children have been studied in terms of genetic typing of the serum proteins and have been found to be Pi phenotype M z and, hence, genetic heterozygotes.

Defining the hepatic material more precisely might bring us closer to a way of effecting its removal from the liver and into the blood, which could be important from a clinical standpoint. Actually, we have tried treating a few of these patients with phenobarbital, knowing that it stimulates proliferation of rough to smooth endoplasmic reticulum, but serum levels of alpha-1-antitrypsin have not been changed as a result.

It is clear that the problem is localized to the liver; the possibility of a circulating inhibitor was ruled out by the finding that on measurement of the trypsin inhibitory capacity of an equal mixture of serum from homozygous and normal individuals, the antitryptic activity is affected only to the expected degree.

We can only speculate as to how the serum deficiency in alpha-1-antitrypsin might contribute to development of cirrhosis in homozygous deficient children. Conceivably, a lack or near lack of antienzyme in the blood renders liver tissue more vulnerable to damage by noxious stimuli of one sort or another. Ordinarily, hepatic inflammation from invasion by toxic agents or infection causes formation of scar tissue that is subsequently removed by normal degradation of collagen and fibrin. Perhaps the collagen or fibrin formed in response to inflammation may be inadequately resolved in the absence of sufficient alpha-1-antitrypsin.

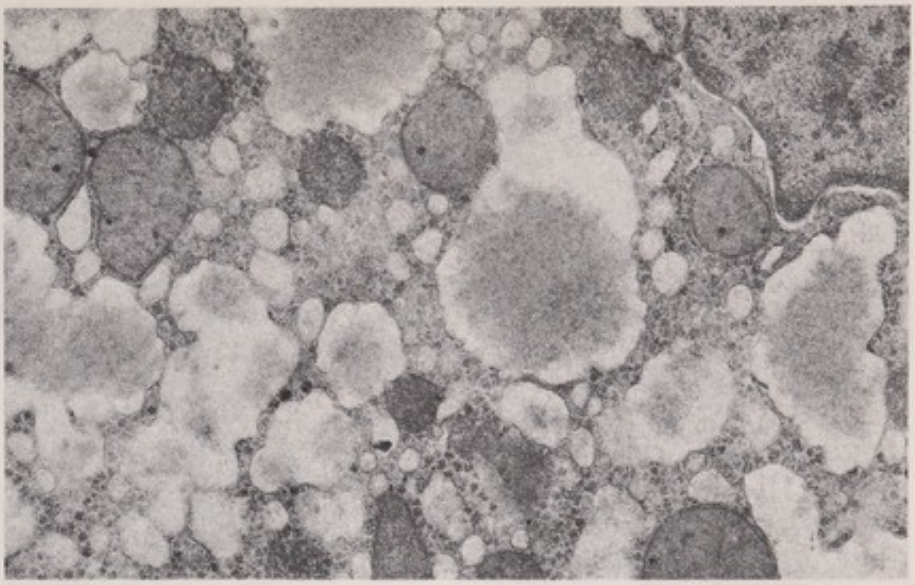
Whatever the cause of the cirrhosis, as of now the clinical outlook for these children is uniformly bleak. Treatment of esophageal varices with portacaval or splenorenal shunts effectively reduces the possibility of uncontrolled hemorrhage in the upper gastrointestinal tract, but life expectancy is not significantly improved.

Obviously, when liver transplantation becomes a more routinely practical possibility, these patients would be among prime candidates, since removal of the diseased liver and its replacement by an organ from a donor without the genetic defect should provide a source of normal alpha-1-antitrypsin, thus preventing redevelopment of hepatic cirrhosis. In our series, liver transplant has been attempted in two patients. One was a teenager whose problems were compounded by a 50% right-to-left shunt in his lungs apparently secondary to the hepatic disease. His hypoxia and cyanosis caused us to consider liver transplant; past clinical experience had suggested that pulmonary A-V fistulas have the capacity to close up upon improvement of liver function. Within two days following the transplant procedure, serum trypsin inhibitory capacity increased to normal; it remained so until the patient's death from respiratory complications. The second patient also retained normal levels until he died from complications of the transplantation.

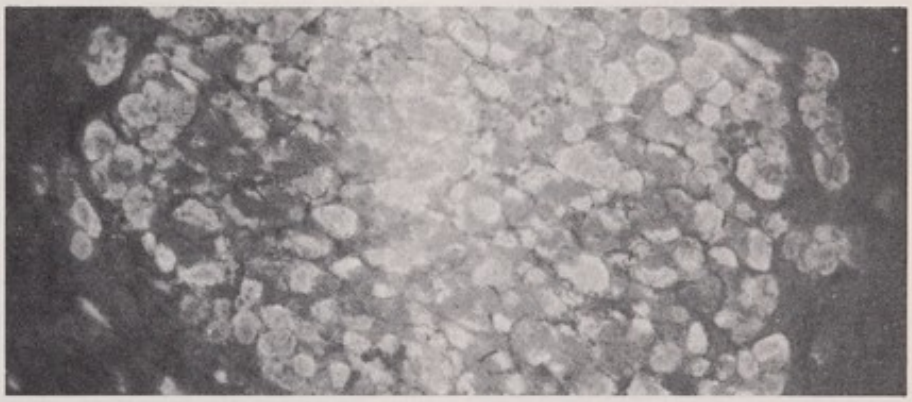
The question of a relationship between familial liver cirrhosis and familial emphysema remains to be clarified. In Eriksson's initial work he discounted the possibility of liver disease in the patients he studied. As already indicated, in our series one family consisted of individuals with either cirrhosis or emphysema associated with homozygous alpha-1-antitrypsin deficiency. A few patients have now been reported with both emphysema and cirrhosis. Homozygous deficient individuals in our other 17 pedigrees had normal pulmonary function, with two exceptions. One was the teen-age boy with the pulmonary A-V shunt already referred to; the other was a girl homozygous both for alpha-1-antitrypsin deficiency and cystic fibrosis. At postmortem her liver showed pathologic changes typically associated with alpha-1-antitrypsin defi-

ciency rather than the eosinophilic concretions within the bile duct lumen that are pathognomonic for cystic fibrosis. The abnormal liver material seen with serum alpha-1-antitrypsin deficiency was demonstrated by both light and electron microscopy. In this case, also, it seems probable that a familial cirrhosis and cystic fibrosis occurred in the same family by chance, since we have searched for the antienzyme deficiency in many patients with cystic fibrosis and found it only this one time. Other workers who have sought to link these two disorders have not succeeded either.

Although both the childhood cirrhosis and the adult emphysema are associated with a plasma deficiency of alpha-1-antitrypsin, similar or different mechanisms may be responsible. Important information has been obtained through liver biopsies of patients with emphysema associated with the antienzyme deficiency. These specimens tell us that the same abnormal material is present within their livers as in the children with cirrhosis.



Electron microscopy of thin sections discloses an unusual amorphous material confined to the lumen of rough endoplasmic reticulum, as in portion of one cell shown here.

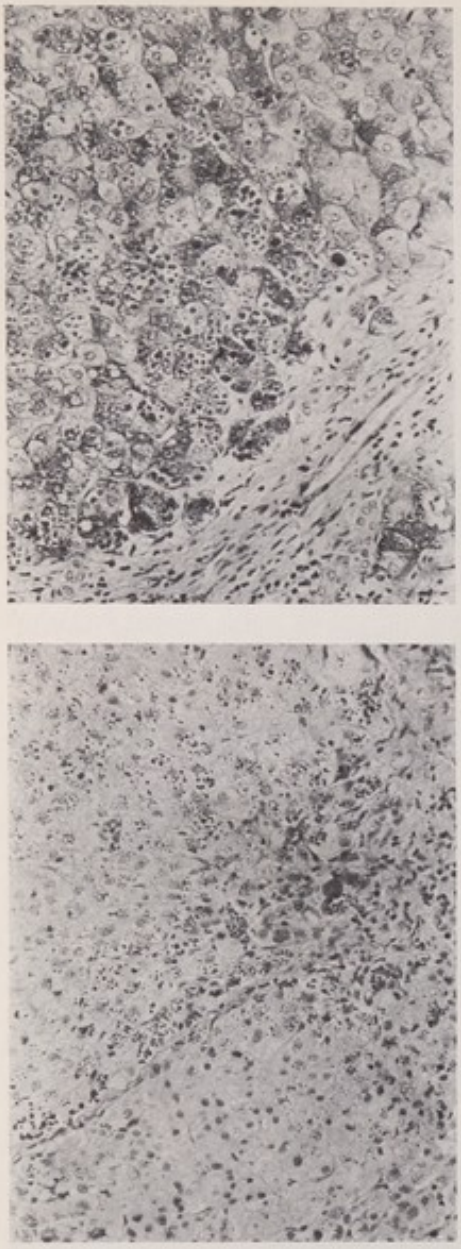


Marked positive fluorescence present in the hepatic cell cytoplasm after application of fluorescent alpha-1-antitrypsin antibody to a serum-deficient patient, shown above at lower and higher magnifications, is completely blocked with use of alpha-1-antitrypsin antigen, confirming the specificity of the antibody.

As is the case with the familial cirrhosis, pathogenesis of the pulmonary emphysema is still to be delineated. Eriksson's original postulate was that the pulmonary disease might relate to lung damage caused by the action of proteolytic enzymes. Kueppers and Bearn have reported evidence that leukocyte proteinases and other enzymes – released as a result of inflammation and in the absence of the antienzyme not duly inhibited – can lead to digestion and destruction of lung tissue. An alternative hypothesis associates the normal aging processes with metabolic turnover of connective tissue fibers in the lung. By somehow acting to hasten metabolic turnover, a deficiency of serum alpha-1-antitrypsin might accelerate aging changes, causing premature enlargement of air sacs in particular, and hence be a factor in early-onset emphysema.

More recently, studies of lung function in patients with this type of emphysema have shown severe limitation of forced expiratory flow due mainly to loss of lung recoil. Workers in the field have therefore focused on the elastase-inhibiting properties of alpha-1-antitrypsin, seeking to demonstrate possible effects on elastic tissues of the lung in the absence of the serum inhibitor. Just recently, James Adamson and workers in his laboratory studied patients with advanced emphysema at the stage of pulmonary failure; a marked deficiency in serum alpha-1-antitrypsin was demonstrated in all their cases. Collagen and elastin measurements on the right middle lung lobe were compared in these patients and in normal subjects. The mean ratio of collagen to elastin was 4.5:1 for those with emphysema and 3.3:1 for the controls, indicating a decreased quantity of pulmonary elastin in the former.

However, as Adamson pointed out,



Distinct cytoplasmic material in PAS-stained sections examined by light microscopy (photo at top) persists despite the use of diastase treatment to eliminate PAS-staining of glycogen (bottom).

whether this is a constant or important feature of the familial emphysema remains to be determined. Whether it occurs in other types of emphysema as well is also not known. If elastic tissue is shown to be diminished to a greater degree in the alpha-1-antitrypsin deficient than in patients with other types of emphysema, the question still to be answered would be how the damage is initiated within the lung. In reference to both diseases it seems almost self-evident now that they must have a multifactorial etiology, and that a severe alpha-1-antitrypsin deficiency, like that found in the homozygous z z individual, must combine with other unfavorable factors for overt disease to result. This seems the only explanation for the finding that some individuals homozygous for the serum protein defect have been protected from developing either the liver or the lung disease.

It is not yet certain how much clinically evident pulmonary disease in adults or liver disease in children occurs in association with homozygous alpha-1-antitrypsin deficiency. Eriksson estimated that probably 1% of cases of emphysema are of this type; others report estimates as high as 5% to 10%. Interest in the familial liver disease associated with alpha-1-antitrypsin deficiency is still too new for meaningful incidence figures to be available. Next to biliary atresia this is presently the most common type of infantile liver disease that we encounter in the pediatric age range, but ours is a referral hospital and admittedly the population is highly selected. We are also finding it more common than other storage diseases that are better known to the clinician.

Until now the most common inborn metabolic error has of course been cystic fibrosis; disease associated with severe alpha-1-antitrypsin deficiency may prove equally important, since both defects appear to occur at about the same frequency. According to current estimates the frequency of homozygous alpha-1-antitrypsin deficiency is 1 in 1,500 to 1 in 2,500, the same as that estimated for cystic fibrosis. A 5% incidence of the heterozygous state also appears to apply to both.

Certainly in a child with liver disease of undetermined cause the possibility of a genetic deficiency in alpha-1-antitrypsin should be carefully excluded, as it should in an adult showing characteristic signs of familial pulmonary emphysema. Screening for the possibility of homozygous alpha-1-antitrypsin deficiency is readily performed at any hospital by visual examination of serum electrophoretic patterns; in the absence of the alpha-1-globulin band a test of trypsin inhibitory capacity should follow. Other procedures such as the alpha-1-antitrypsin determination by immunochemical techniques or testing for elastase inhibitory capacity can be substituted for the trypsin inhibitory capacity test because the results have excellent correlation. All these tests can usually clearly identify homozygotes but offer some problems in definitively separating heterozygotes from normals. Genetic typing is ideal but not essential, particularly when other family members are available for antienzyme testing.

When a child is suspected of being genetically deficient in alpha-1-antitrypsin, family studies should indicate that both parents carry an abnormal gene. As our experience showed, the presence of this factor may be masked by pregnancy or use of oral contraceptives. In either circumstance the mother's serum level of antitrypsin may be pushed to a higher range; misinterpretation can result if this is not taken into account. In counseling the family of a genetically homozygous child, the recurrence risk is probably best presented on the basis of autosomal recessive inheritance. That is, with both parents heterozygous, there is a 25% risk that a child will be homozygous and deficient and a 50% risk that he will be heterozygous.

Screening for the heterozygote carrying the deficiency gene presents more of a practical problem than screening for homozygotes, since standard serum electrophoresis is unlikely to reveal the defect. More complex techniques, such as acid starch gel immunoelectrophoresis, are required, but these are not widely available. However, without firmer evidence that heterozygotes are predisposed to either of these familial diseases, there probably is no need for mass screening programs to detect carriers. That the homozygote can be readily identified early in life is of course important not only in terms of anticipating the possibility of one of two grave illnesses in the affected individual but in genetic counseling of the family as well.

That early knowledge of the serum glycoprotein defect will improve prospects for preventing the liver disease is unlikely until greater understanding of the pathogenic factors has been achieved. As for the lung disease, even without more knowledge, it seems only sensible that the homozygous deficient individual be advised against cigarette smoking and undue exposure to lung irritants.

14

Mucopolysaccharidoses: The Biochemical Approach

ELIZABETH F. NEUFELD
National Institute of Arthritis, Metabolism, and Digestive Diseases

As genetic diseases go, the mucopolysaccharidoses are not particularly common. Even considered collectively, the members of this closely related group of anomalies affect only one birth in perhaps 25,000, which puts them somewhere between phenylketonuria (one in 10,000 births) and galactosemia (one in more than 100,000 births). Yet these relatively rare conditions have turned out to possess a scientific and clinical significance well beyond that of many commoner complaints.

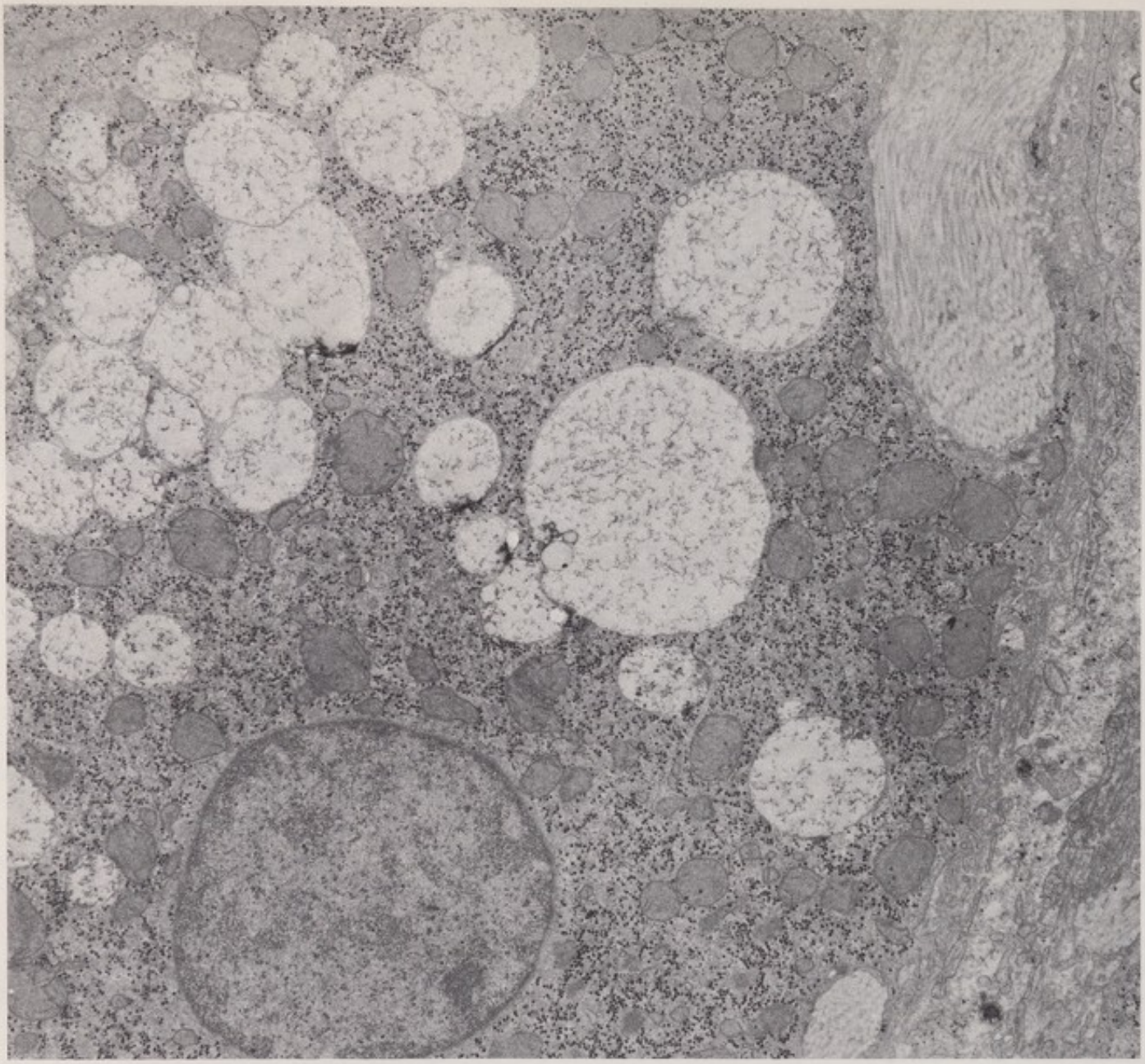
The study of these abnormalities has, of course, thrown considerable light on aspects of normal physiology, but this is no great distinction; the same could be said of most intensively studied genetic anomalies. In addition, however, the mucopolysaccharidoses were among the very first genetic diseases to be "diagnosed" by tissue culture techniques (see Mellman, Chapter 16). They also constitute a remarkable case study in disease classification, or nosology (see McKusick, Chapter 20), about which I shall have more to say later. Finally, though much work remains to be done, there is already good reason to think that the mucopolysaccharidoses will, before many years have passed, join the still sparse ranks of the genetic diseases that can be effectively treated.

The best known, and most severe, of the mucopolysaccharidoses is the Hurler syndrome, first reported in 1919 by the German pediatrician of that name. During the first year of postnatal life, the infant begins to show signs of both physical and mental deterioration, gradually acquiring an extraordinary appearance: large head, flat bridge of the nose, wideset eyes, large lips, coarse tongue. The Hurler patient suffers also from stunted growth, corneal clouding, impaired hearing, widespread skeletal involvement, hepatosplenomegaly, severe mental retardation, and from cardiovascular abnormalities as well. Under these circumstances, it is hardly surprising that few of these individuals survive their teens.

Even before Gertrud Hurler published her detailed and classic description, the English physician Charles Hunter had reported the related anomaly that bears his name. The Hunter syndrome usually follows a considerably milder physical course; several individuals have survived into their forties. The mental manifestations are variable: some patients are severely retarded, whereas others may be of normal or even above-average intelligence. Subsequently, other investigators described the Sanfilippo syndrome, in which the physical manifestations are mild but mental retardation severe, and the rarer Scheie, Morquio, and Maroteaux-Lamy syndromes, with mild to severe physical manifestations but little or no mental damage.

As their generic name suggests, all of these conditions involve abnormalities in the metabolism of certain mucopolysaccharides (MPS). Without at the moment attempting to characterize these substances in detail, we may note that they are normal constituents of various kinds of connective tissue. The most conspicuous abnormality in their metabolism is their level in the urine of affected individuals, where excretion may exceed 100 mg/day, as against less than 15 mg/day in normal persons. In addition, they accumulate within cells in many of the body's tissues, and there is every reason to believe that these widespread cellular deposits account for most or all of the clinical manifestations.

Elevated urinary MPS is the primary diagnostic sign of the mucopolysaccharidoses. Variations in the specific MPS present help in the differential diagnosis of individual anomalies. The substances in question (in all but one condition) are dermatan sulfate (DS), an important constituent in normal skin and blood vessels, and heparan



Bloated lysosomes (white bodies) crammed with particles of partly degraded mucopolysaccharide are seen in electron micrograph of liver cell from Hurler patient (Drs. F. van Hoof and H. G. Hers, Laboratoire de Chimie Physiologique, Louvain, Belgium).

sulfate (HS), found primarily in the blood vessels. In Hurler patients, DS and HS are present in a ratio of about two to one, while in Hunter patients the ratio approximates equality. Both ratios, however, are sufficiently variable to make them unreliable as an absolute guide to differential diagnosis. Here, however, the mode of inheritance is conclusive: the Hunter syndrome is a sex-linked recessive (manifesting itself entirely in males who inherit it from their carrier mothers), while the Hurler syndrome, like the rest of the mucopolysaccharidoses, is an autosomal recessive, requiring a gene from each carrier parent.

Sanfilippo urine contains HS almost exclusively, though a small amount of DS can usually be detected by refined techniques. Scheie, like Hurler, involves DS and HS, whereas Maroteaux-Lamy urine appears to contain DS exclusively. Morquio is unique in that neither DS nor HS is present, but an unrelated mucopolysaccharide, keratan sulfate, is found, which is a normal constituent of cartilage and the cornea. Not very surprisingly, the physical manifestations of this disease are quite different from those of the others.

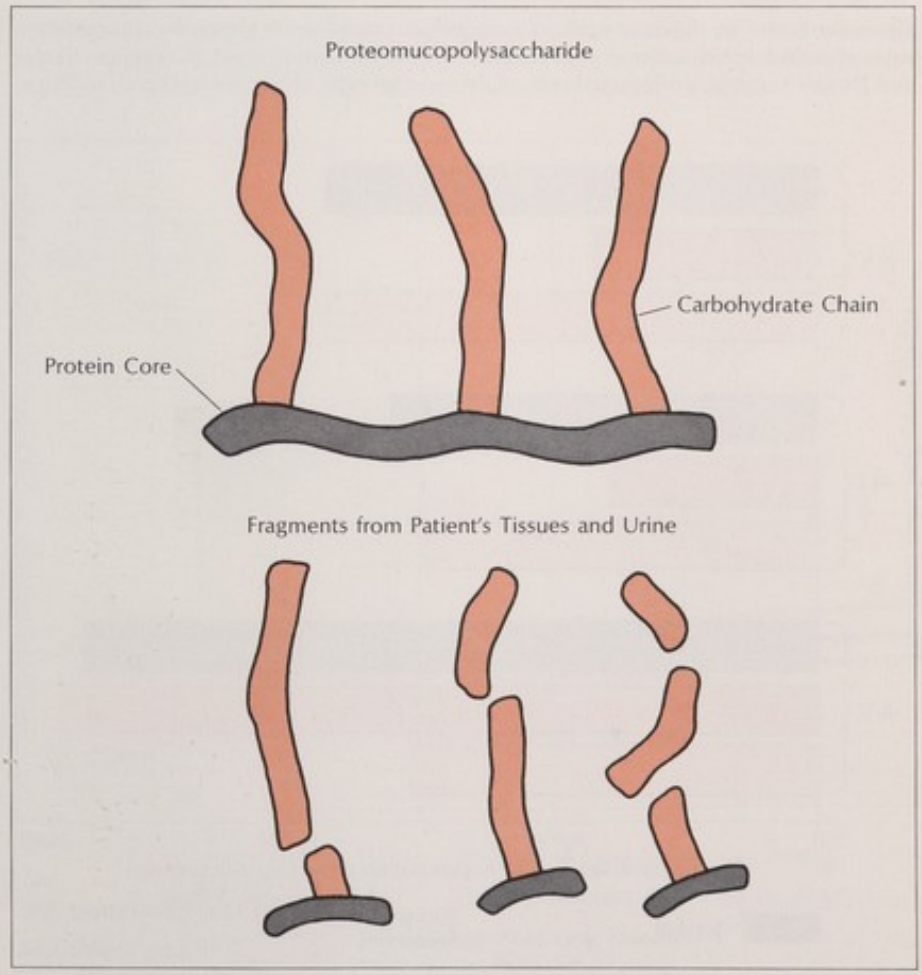
Based on the clinical signs and specific MPS anomalies, the six conditions were classified by Dr. Victor McKusick some years ago as MPS I through VI. My associates and I have demonstrated the general accuracy of this classification by means of biochemical findings. These were based on an earlier finding, by Drs. Danes and Bearn, that cultured fibroblasts from Hunter and Hurler patients can be distinguished from normal cells by their development of "metachromatic" (pink) granules when stained with toluidine blue dye. Under some conditions, this effect shows up even in

heterozygous individuals and is the basis of a test for carriers. Subsequently, our research group has determined that cells from different mucopolysaccharidoses can "cross-correct" one another. For example, if Hurler and Hunter cells are cultured together, neither will show the characteristic pink granules. A test more sensitive than metachromasia is the measurement of sulfated MPS, using radioactive sulfur for easy monitoring. Hurler and Hunter cells, as described below, accumulate much more radioactive MPS than do normal cells; when they are mixed, their accumulation is reduced to a normal level. The same type of correction occurs if one adds merely the medium in which cells of the other type have been cultured. Such cross-correction has been demonstrated in almost all members of the group (except Morquio, which we have not studied), though with certain interesting anomalies. For one thing, it appears that Sanfilippo, biochemically speaking, is not one disease but two; that is, cells from some Sanfilippo patients will cross-correct with cells from others. We have named these two groups, which are otherwise indistinguishable by present techniques, Sanfilippo A and B. For another thing, and rather surprisingly, Hurler and Scheie appear to be, in a certain sense, two forms of a single disease. Despite their marked clinical differences, their cells will not cross-correct. As a final point in this connection, I might mention that cells (or culture medium) from normal individuals will correct *any* of the group.

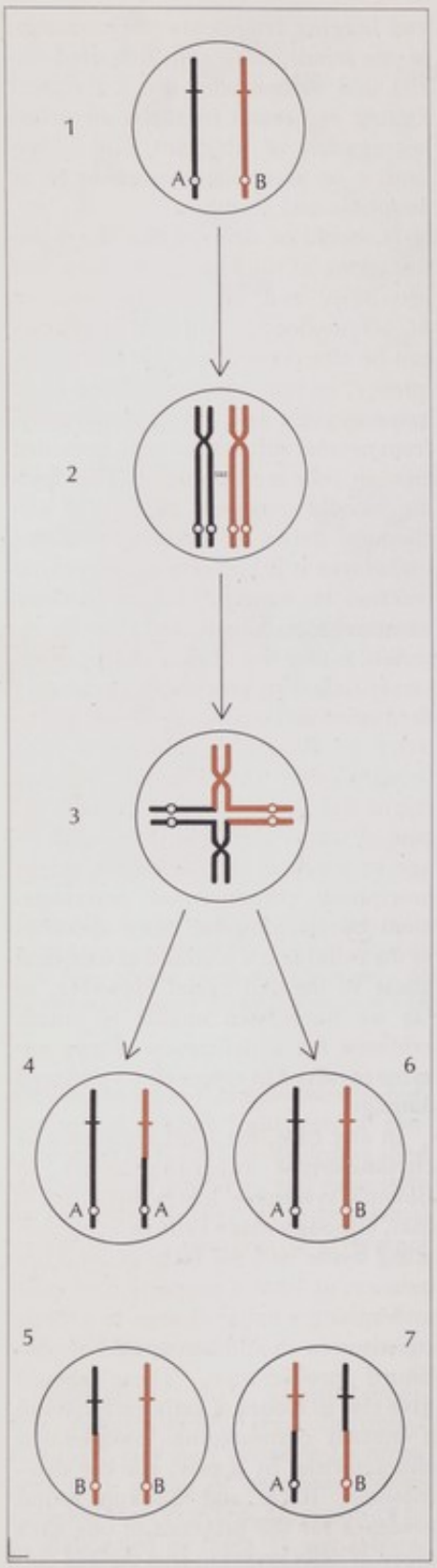
These cross-correction findings are an obvious aid to the accurate differential diagnosis of the mucopolysaccharidoses. In addition, however, they carry a number of implications concerning the deeper biochemical causes of these conditions. To explain what these are, however, I must first fill in some background on the biochemistry of MPS as such.

As they normally occur in the body, MPS are "protein polysaccharides" – long chains of amino-acid residues to which are attached at intervals chains of polysaccharide. Though we know little about the synthesis of these complicated molecules, it is generally believed that the protein core is synthesized first, after which the polysaccharide chains are constructed along this "skeleton" by the step-by-step addition of one sugar at a time. For example, it has been shown that if protein synthesis is inhibited (e.g., by addition of puromycin to the culture medium), this prevents formation of polysaccharide also, although under other circumstances the inhibitors in question have no effect on the polymerization of sugars. The details of the elongation of the polysaccharide chains are still obscure, but intensive research carried out in Sweden may soon clarify this aspect.

About the degradation of MPS even less is known. A probably significant fact, however, is that though in the connective tissues of normal as well as affected individuals, these substances occur in their "ordinary" macromolecular form, both the abnormal intracellular deposits and the abnormal urinary output of MPS consist rather of fragments of these large molecules. The protein core has almost disappeared; what remains are groups of a few amino acids linked to some bits of carbohydrate, plus other polysaccharide fragments without amino acids. Apparently most of the protein has been completely broken down, while the carbohydrate portions and attached protein fragments have been only partially degraded. These accumulations of fragmented MPS can be seen: Under the electron microscope, for example, cells from the livers of affected individuals are full of vacuoles containing very finely dispersed materials. The vacuoles are evidently lysosomes – the cellular organelles normally responsible for the breakdown of macromolecules – pathologically gorged with undigested mucopolysaccharide. Significantly, they resemble the bloated liver lysosomes of rats that have been injected with unmetabolizable molecules such as those of the detergent Triton WR-1339.



Normal MPS (top) consists of carbohydrate chains attached to a protein core. In urine and tissue deposits of patients, most of the protein core is degraded into its constituent amino acids, but the polysaccharide chains can only be partially broken down (below).



Diagrams suggest sequence whereby chromatid interchange (2, dashed line) between homologous chromosomes (1) resulting in formation of symmetrical Qr figure (3) could yield progeny genetically different from parent cell strain and homozygous for all genes distal to the point of exchange (4, 5). Alternatively, the progeny might be like the parent cell (6) or merely show new position of alleles (7).

become homozygous for all genes of the affected chromosome distal to the point of exchange. Though retaining a balanced and complete human genome, each daughter cell, and its progeny, would lack certain alleles present in all the other cells. If this is actually the case in vivo in Bloom's syndrome, evidence for antigenic or enzymatic mosaicism should be demonstrable, if an adequate search were made.

The cytogenetic findings described in reference to Bloom's syndrome cells in tissue culture do not necessarily apply to cells proliferating in vivo, but if they do, this could account for some of its clinical features. For example, poor survival among daughter cells receiving the more abnormal chromosomal complements could relate to the severe growth retardation characteristic of Bloom's syndrome. Evidence that chromosomal instability does occur in vivo was obtained when we examined cells aspirated from the bone marrow of a *bl/bl* seven-year-old boy and found increased numbers of cells with dicentric chromosomes, triradial configurations, and other complex exchange figures. (The one other marrow we have been able to examine appeared normal.)

In the course of our cytogenetic studies of Bloom's syndrome, similar chromosomal aberrations in dividing lymphocytes have been identified in dividing lymphocytes from all those examined, i.e., from most of the known affected individuals, including three of those who went on to develop cancer. This probably means that the chromosomal instability has little to do with the cancer itself, but does it not provide a predisposing background?

Fanconi's Anemia

The situation is, in some ways, similar in Fanconi's anemia, another rare genetic disorder, which combines a functional defect in the bone marrow with major anomalies of solid structures including the skeleton, heart, and kidneys. Actually there are not many clinical similarities between Bloom's syndrome and Fanconi's anemia. The latter may be associated with growth retardation, but not to the same degree. Brownish skin pigmentation is prominent but more extensive and of different distribution than the café au lait spots often observed in Bloom's syndrome.

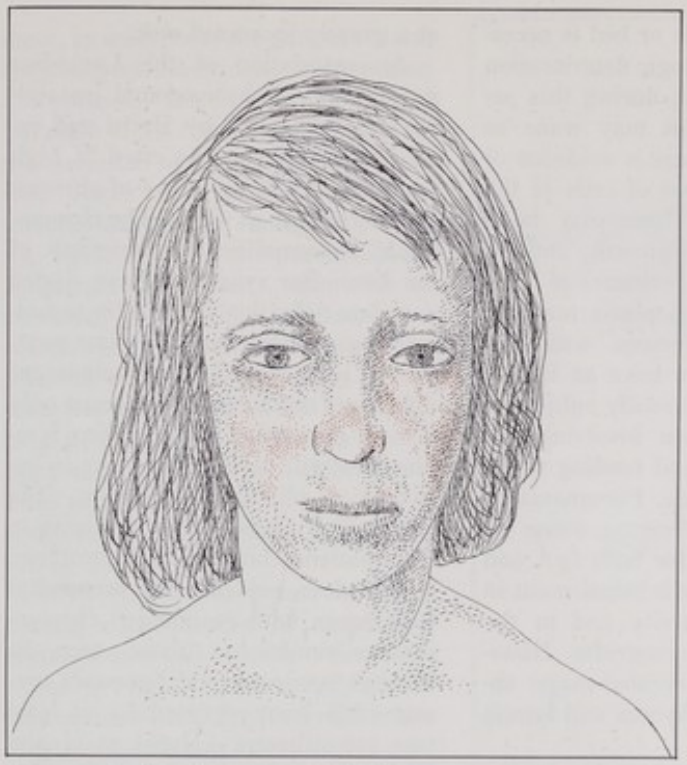
Since Fanconi's anemia has been recognized far longer (since 1927), more cases, probably around 200, have been studied and described. As in Bloom's syndrome, an increased incidence of parental consanguinity and of affected sibs of both sexes points to a recessive transmission. The genetics has not been well studied, however.

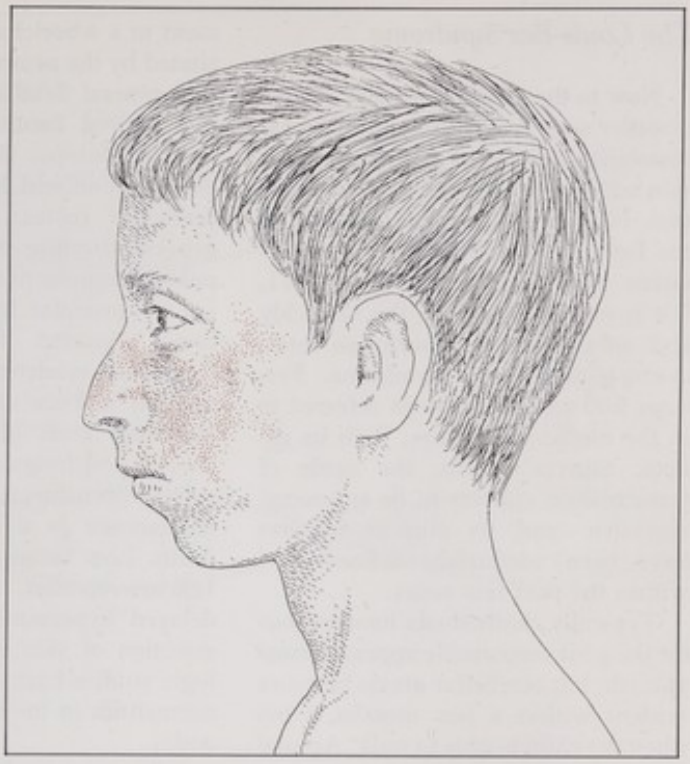
The hematologic dysfunction associated with Fanconi's anemia is severe and affects all elements of the bone marrow. Pancytopenia and bone marrow hypoplasia usually become symptomatic between the ages of 4 and 12 and tend to progress until the child's death from hemorrhage or other manifestations of bone marrow failure. Recently, hormone therapy and other interventions have been used to ameliorate the hematologic disorder.

Although some detailed reviews of Fanconi's anemia, including one by Fanconi himself, do not refer to an increased incidence of leukemia, the generally accepted view now is that the risk is definitely increased. While this view is based on relatively few cases, considering the infrequency both of leukemia and of the syndrome itself, the number of affected persons with leukemia appears sufficient to indicate an association. An important possibility that has long intrigued clinicians observing this disease is that cancer risk may also be increased in relatives of affected individuals, presumably heterozygous for a Fanconi's anemia gene, and supporting evidence was recently provided by Swift and coworkers.

Given the severity of the hematologic abnormality in Fanconi's anemia, any predisposition to leukemia might not be fully expressed because, at least in the past, few patients have survived for long after the appearance of panmyelopathy. If a neoplastic cell line is to emerge and replenish the marrow, obviously it would have to do so before the patient succumbs to hemorrhage or overwhelming infection. Will prolongation of life in Fanconi's anemia by hormonal therapy be associated with increasing numbers of leukemics?

Some cytogenetic comparisons bear mentioning in reference to Fanconi's anemia and Bloom's syndrome. The increased chromosomal instability in





Because of their characteristic facial appearance as well as the presence of telangiectatic lesions, unrelated individuals with Bloom's syndrome bear a close physical resemblance to one another. A key feature shared is their small stature, resulting from growth retardation beginning in utero and continuing throughout life. Maximum height almost never exceeds five feet.

Fanconi's anemia was first reported some eight years ago by Schroeder, on the basis of her studies of cultured blood lymphocytes. We have observed this in skin fibroblast cultures as well. On first inspection the chromosomal abnormalities seem similar to those in Bloom's syndrome; there certainly are many breaks and rearrangements. Recently, Schroeder and I jointly evaluated the cytogenetic disturbance in the two disorders and concluded that the patterns of chromosomal disruption and rearrangement are different. Whereas chromatid interchanges and rearrangements, chiefly the equal symmetrical Qr, predominate in Bloom's syndrome, in Fanconi's anemia a relatively higher proportion of cultured cells show chromatid gaps and breaks. When Qr's occur in Fanconi's anemia, nonhomologous chromosomal regions are usually affected; moreover, sites of lesions are fairly randomly distributed throughout the chromosomal complement. Genes responsible for the two disorders obviously have different effects clinically, and a difference in the pattern of chromosomal instability is not unexpected.

Using bone marrow aspiration studies to evaluate the degree of chromosomal instability occurring in vivo, Schroeder found an increase in mitotic and nuclear aberrations during metaphase in a few patients as well as anaphase bridges, fragmented chromosomes, and micronuclei. Using serial cytogenetic analysis of aspirated bone marrow, she also detected and followed the proliferation of a pseudodiploid clone, i.e., a cell population with 46 chromosomes but with some chromosomal rearrangement. In examining skin fibroblast cultures derived from a case of Fanconi's anemia, Young detected two clonal subpopulations, each with a mutant chromosomal complement, growing among cells with normal complements. In these patients there was no association with leukemia. All that can be said at present is that clones of cells containing a clearly aberrant chromosomal complement but exhibiting no malignant characteristics are associated with a disorder carrying an increased risk of cancer. Moreover, like the stem lines of established cancers, these clones evidently originate from a single cell containing a stable chromosomal rearrangement.

In attempting to elucidate the relationship between Fanconi's anemia and leukemia, Todaro used simian virus (SV) 40 to infect diploid fibroblastic cell lines either homozygous or heterozygous for a Fanconi's anemia gene. In such cell lines, as compared with control cultures, monolayers of cells planted on the surface of a culture dish showed a greater number of localized areas undergoing rapid and disarrayed proliferation. Whereas in control cultures 1.6 to 5.1 areas of "transformation" appeared for every 10,000 cells, the number increased to 20.1 to 28.2 per 10,000 among cells carrying one gene for Fanconi's anemia; in cells containing two Fanconi's anemia genes it increased to 41.4 to 79.7 per 10,000. Although these proliferative responses termed "transformation" may not always be assumed to represent malignant conversion, what is of interest is that while the changes occur most readily in cells from the affected homozygote, they are also seen in cells of normal individuals. Response in the heterozygote is intermediate between the two. Thus, the difference is not in the kind of interaction between the virus and the cell but in the frequency with which it occurs.

Summary of Mucopolysaccharidoses

Name	*Transmission*	*Physical Anomalies*	*Mental Deficiency*	*Urinary MPS*	*Deficient Factor*
Hurler	AR	+++	++	DS, HS (~2:1)	α-L-iduronidase
Hunter	X-linked	++	+/–	DS, HS (~1:1)	sulfoiduronate sulfatase
Sanfilippo A	AR	+	+++	HS	HS sulfatase
Sanfilippo B	AR	+	+++	HS	N-acetyl-α-glucosaminidase
Morquio	AR	+++	–	KS	Not studied
Scheie	AR	+	–	DS, HS (~2:1)	Same as Hurler
Maroteaux-Lamy	AR	++	–	DS	Purified, NEC
"Atypical"	AR	++	–	?	β-glucuronidase

AR *Autosomal recessive* DS *Dermatan sulfate* KS *Keratan sulfate* HS *Heparan sulfate*
+ + + *Present, marked* + *Present, mild* + + *Present, moderate* – *Not present*
NEC *Not enzymatically characterized*

most invariably, rather, it *is* further metabolized, along some normally little-used pathway. Actually, in some diseases it has been determined that the pathologic features are not produced by the intermediate substrate as such but by abnormal metabolites derived from it. Thus it is perfectly possible that two different MPS breakdown intermediates, corresponding to two different blockages in the enzymatic chain, might be transformed through abnormal channels into identical end products.

As against these quite inconclusive objections to the enzyme theory, we have much additional, and I believe conclusive, evidence on the other side. For one thing, we have succeeded in isolating in varying degrees of purity five substances corresponding to the five hypothesized missing enzymes.

Our source was human urine – a most useful and usually neglected source of human proteins. These substances (we like to call them factors) are all proteins, as all enzymes are. They range in molecular weight from about 70,000 to as high as 300,000. They are highly specific: Hunter factor, for example, will correct the defect in Hunter cells, without affecting in any way normal cells or cells derived from patients with other mucopolysaccharidoses. This is to be expected, as these cells already have Hunter factor in adequate amounts, and their deficit is in a different factor.

In all cases, the factors correct by accelerating MPS degradation in the appropriate recipient cells. This implies that they are lysosomal enzymes. On the theory that MPS accumulated in abnormal cells may represent intermediates in the breakdown process, we have mixed radiosulfur-labeled MPS obtained from such cells with relatively pure preparations of Hunter, Hurler, and Sanfilippo A factors. This approach proved to be particularly successful in the case of the last. Sanfilippo A factor released inorganic ("free") sulfate into the medium, suggesting that it is a sulfatase, whose biochemical function is to detach the sulfate groups from the heparan sulfate molecule. With the other factors, no sulfate was released, but this is not to say that nothing happened – merely that whatever happened was not detectable in this labeling system. By other methods, however, it was found that the Hurler factor is the enzyme alpha-iduronidase. Its function is to cleave iduronic acid residues – the major constituents of dermatan sulfate and minor constituents of heparan sulfate. The Sanfilippo B factor was identified as the enzyme N-acetyl-alpha-glucosaminidase, which is needed for the breakdown of heparan sulfate. The very rapid identification of factors as enzymes involved in mucopolysaccharide degradation came as the result of efforts in several laboratories – in Chicago, La Jolla, and Münster – as well as of our own work.

Meantime, we have come across another and apparently distinct type of mucopolysaccharidosis, in a child being treated at a St. Louis hospital. The patient was clinically "atypical" in having some but not all the signs of the Hurler syndrome. Fibroblasts taken from a skin biopsy turned out to be even more atypical: they were not corrected by our purified factors. Instead, the missing factor turned out to be a known lysosomal enzyme, β-glucuronidase, preparations of which (from beef liver) reduced the cells' MPS accumulations to normal levels. Since there are other grounds for believing that this enzyme plays a role in MPS degradation, we may have identified a sixth step in the breakdown process. Nor does this exhaust the possibilities. Examination of the chemical structure of the affected mucopolysaccharides shows at least five linkages for which there are no enzymes yet known. One of these unknown enzymes may correspond to the Maroteaux-Lamy factor. This would still leave several enzymes "in search of a disease."

The fact that pathologic cells can be more or less routinely freed of their MPS accumulations, either by cross-correction or by the addition of purified factor, naturally raises the question of whether the same could be done in vivo. In most inborn errors of metabolism, the "enzyme replacement" approach is not considered very promising at present, mainly because of the difficulty of getting the exogenous enzyme into the cells. Lysosomal enzymes, however, seem to present a less intractable problem. Cells have long been known to take in proteins from outside the cell by pinocytosis. After the exogenous molecules are enclosed in an invagination of the cell wall, they are in a vacuole that eventually merges with lysosome. Our own experiments have shown that even abnormal cells take up proteins from the medium very easily by this process.

The presumption is that enough of the right enzyme injected into the bloodstream should permit the cells to purge themselves of MPS accumulations and, it is to be hoped, to resume

normal function. To be of any use, of course, such therapy would have to be continued lifelong, as with insulin in diabetes. It is likely, however, that replacement therapy in mucopolysaccharidosis would not need to be anywhere as frequent as in diabetes. It takes time for the MPS accumulations to build up to a serious level and, moreover, once the enzyme is in the cell, it does not leave rapidly. Its half-life can be as long as a week. Thus one would expect that the timing of enzyme injections would be on the order of every few weeks or longer.

The real problem with enzyme therapy is getting enough enzyme. Here, several approaches are under investigation. One of them, pioneered by Drs. Knudson and DiFerrante of Baylor University, involves transfusing leukocytes into the affected individual; these are believed to contain quantities of many enzymes, including the relevant ones, that could benefit the patient, even though the cells die off fairly quickly. Another, related, approach involves transfusing large quantities of cell-free plasma.

Since I am not a clinician, I cannot comment on these experiments in any detail. But I believe most of the clinicians involved in this work would agree that the results to date are unquestionably encouraging, though still preliminary. Concerning the plasma approach, however, there are certain biochemical considerations that make the results rather puzzling. We ourselves have assayed the amount of corrective factor in plasma, and find that one milliliter contains enough factor to produce 50% correction in a few milligrams of cells. A liter of plasma (which is the dosage level that has been used), then, would contain enough factor to partially correct only a few grams of cell mass, whereas even in a child we are talking about kilograms. On the face of it, the plasma technique should be insufficient by at least three orders of magnitude. Conceivably, of course, the factor may be more potent in vivo than in vitro (thus reversing an all too common phenomenon in pharmacology). Alternatively, we have speculated that the therapeutic effects of the plasma may be due to the presence of other enzymes, which may set up an abnormal, but efficient, pathway of MPS degradation. In any case, it is clear that our understanding of correction in vivo leaves much to be desired.

We ourselves have been working at the problem from a somewhat different angle. Instead of plasma or leukocytes, both of which depend on adding a host of enzymes (and other substances) of which only one is actually needed, we hope to employ pure preparations of individual factors. These can be precipitated quite easily from urine, as mentioned above, but one then has the problem of separating the factors from other urinary proteins in order to minimize the possibility of foreign protein reactions. These would probably not be a problem in the case of the factor itself, if we are correct in thinking that in the affected individual the factor is not "missing" but merely inactive. Its immunologic characteristics would then probably be similar to those of the exogenous, active factor.

Thus far, however, we have achieved about 40% purity in one factor and, with present methods, it has been a long, complicated, and expensive business. But we hope to develop more efficient methods of purification in order to accumulate sufficient quantities for a clinical trial.

Meanwhile, it is worth noting that both Hurler and Hunter syndromes are now among the genetic diseases that can be reliably diagnosed in utero by testing cultured fetal cells obtained by amniocentesis; it is possible that Sanfilippo may eventually be added to the list. Assuming that mothers and physicians are willing to follow up a positive diagnosis by terminating the pregnancy, we can hope that at least two of these rare diseases will become even rarer, while continued advances in therapy give the possibility of eventual treatment for those infants who escape the prenatal diagnostic net. Quite apart from these important clinical expectations, there can be little doubt that we will continue to expand our biochemical understanding of both the mucopolysaccharidoses and the normal metabolic processes of which they represent a derangement.

15

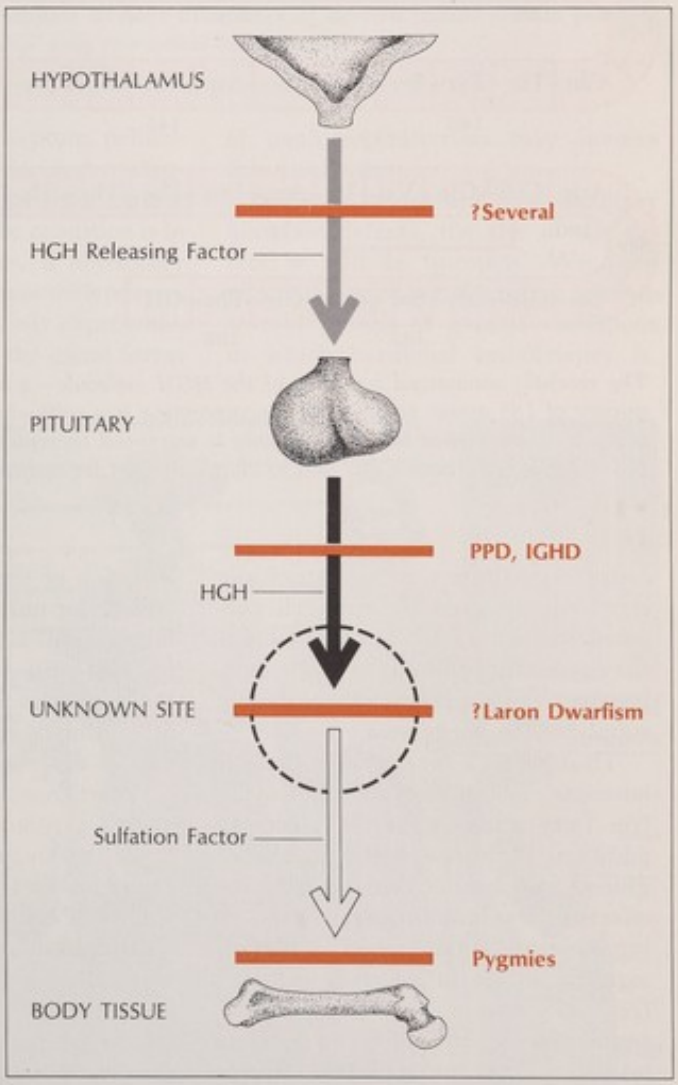
Genetic Defects of Growth Hormone

DAVID L. RIMOIN
University of California, Los Angeles

Six years ago, when he was summing up the current knowledge of growth hormone, John C. Beck of McGill University remarked that "the questions far outnumber the answers." To a large extent, his generalization remains true. Yet though we still have more questions than answers, some new answers have begun to emerge and some tentative old ones have been firmed up. In both instances, useful clues have come from the study of genetic defects involving growth hormone, which may affect either its manufacture and release by the pituitary or its actions elsewhere in the body.

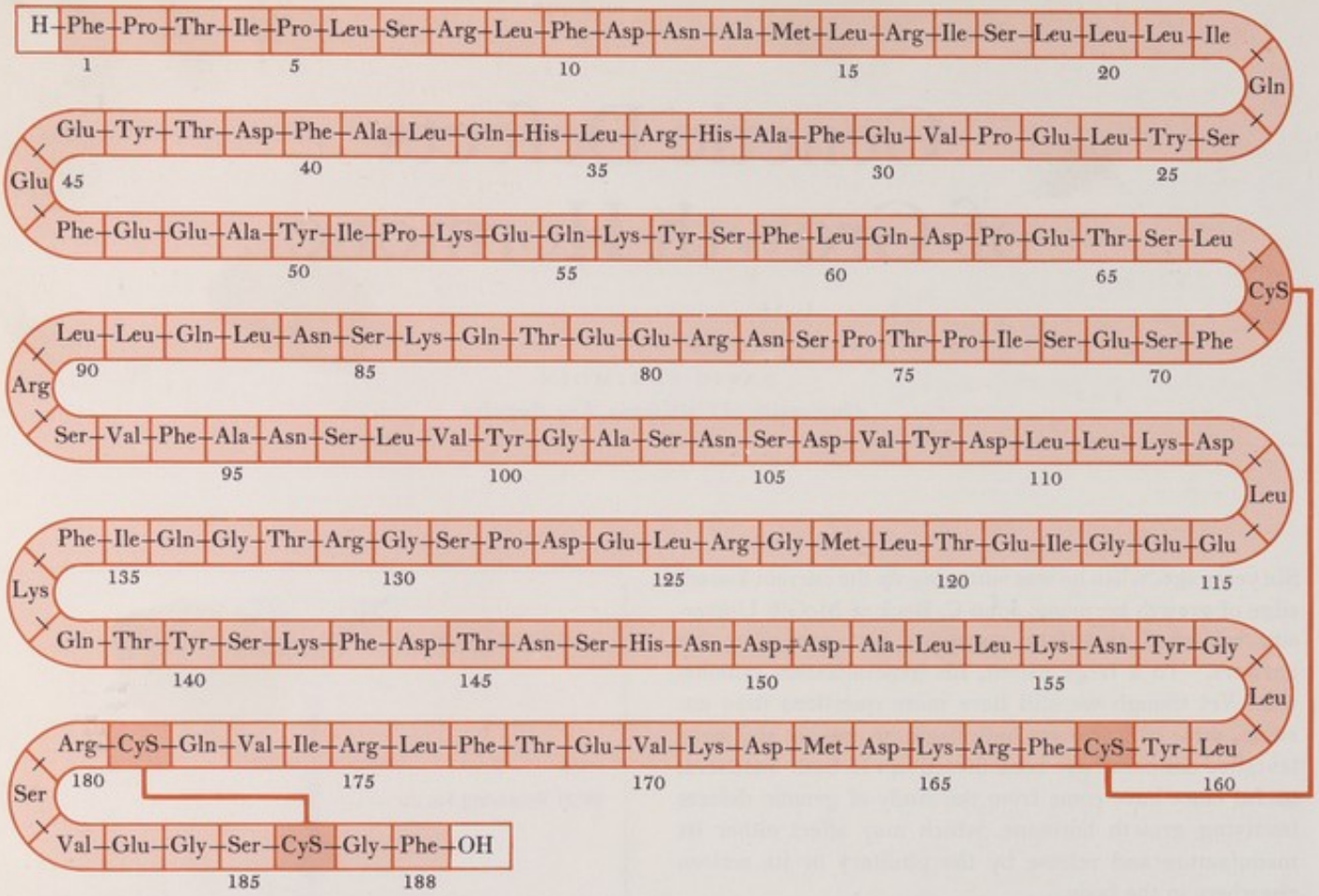
Before considering the findings relating to these rare but intriguing disorders, however, it might be well to review some of the basic facts about human growth hormone (HGH). It is produced, of course, by the anterior pituitary, and like the other pituitary hormones (thyrotropic hormone, follicle-stimulating hormone, ACTH, etc.) is a protein. Its molecule, whose composition was elucidated by Dr. C. H. Li in 1966, consists of 188 amino acid residues arranged in a single chain that is looped back on itself in two places by sulfhydryl bridges (see "The Growth Hormone Puzzle" by John C. Beck, HOSPITAL PRACTICE, May 1967).

Physiologically, the hormone's most obvious function is the promotion of growth, from infancy through adolescence, and it is this property that is chiefly important in genetic HGH disorders, all of which involve dwarfism as their most prominent feature. In addition, however, growth hormone enters into a number of other physiologic processes, both during childhood and in adult life. Some of these less conspicuous effects are of special importance for HGH studies in adults, where growth has ceased and therefore cannot serve as an index of HGH function, and often in children as well, when the assessment of growth response to HGH would require an impractically long time or prohibitive quantities of the hormone. (The substance, which thus far has been obtained only by extraction from



Metabolic chain incorporating HGH involves at least three distinct substances: the hypothalamic "releasing factor," the hormone itself, and the so-called sulfation factor. Genetic defects apparently may affect any point in the chain, or – as in the case of the pygmies – the tissues' response to it.

The HGH Molecule's Amino Acid Sequence



The recently announced synthesis of the HGH molecule – a sequence of 188 amino acid residues incorporating two sulfhydryl bridges – should open the way not only to improved therapy of HGH-deficiency diseases but also to clarification of its multiple physiologic roles, including the determination of height. Answers may also be expected sooner to such questions as HGH's participation in metabolism of fats and carbohydrates, in sexual maturation and lactation, in bone formation, and to other problems.

cadaver pituitaries, is in chronically short supply even for research purposes. In January 1971, Li announced the successful synthesis of HGH, paving the way for a potentially unlimited supply of the hormone.)

Thus we find, for example, that the hormone mobilizes fats, increasing free fatty acids (FFA) in blood. In addition, it decreases the uptake of glucose into muscle, counteracting the effects of insulin and thereby producing hyperglycemia. In fact, it plays a significant role in a rather complex feedback system involving insulin, serum glucose, and serum FFA. Thus insulin injected into a normal individual will not only reduce serum glucose and FFA but the resultant hypoglycemia will also increase serum HGH, which apparently then helps restore glucose and FFA levels to normal. An injection of the amino acid arginine, which (for unknown reasons) releases insulin, will also release HGH. Finally, and still in normal individuals, HGH appears to potentiate the release of insulin in response to hyperglycemia or other stimuli.

Apart from its effect on these parameters, serum HGH can also be measured directly, or almost so. In high concentrations (e.g., in pituitary extracts) it can be assayed chemically, but in physiologic concentrations one must employ the indirect technique known as radioimmunoassay. Briefly, this begins with the production of antibodies to HGH by injecting it into rabbits. A measured amount of antibody is then added to the serum sample, binding whatever HGH is present; then a measured amount of radioactively labeled HGH is added, which binds to any "unused" antibody binding sites. When the antibody is separated out, the proportion of bound to unbound radio-HGH provides a measure of how many binding sites were unoccupied and, inversely, of how much HGH was present in the serum.

To understand how all these physiologic processes are disordered in genetic HGH diseases, we must now consider the steps in HGH metabolism at which derangement might occur. The process begins with the hypothalamus, which triggers the manufacture and/or release of the hormone by the pituitary through the secretion of a small peptide known as growth hormone releasing hormone, or GHRH (for simplicity, I am ignoring the complicated feedback loops – neurologic, humoral, and hormonal – which govern the hypothalamus). Even when

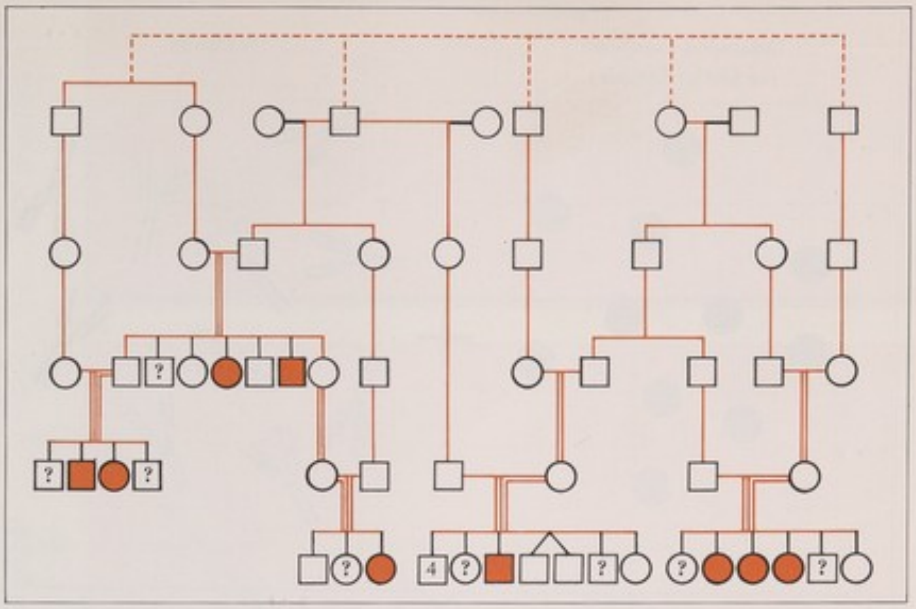
released into the bloodstream, however, HGH appears to produce little effect on the cartilage.

Slices of cartilage or isolated cartilage cells that will "grow" (i.e., incorporate metabolites, including radioactive sulfate) in normal human plasma will not do so in plasma from a hypopituitary human or animal as actively as in that from normal individuals *even when exogenous growth hormone is added*. Yet when the same hypopituitary individual is treated with growth hormone for several days, the serum will then induce normal "sulfation" in the cartilage. We conclude, therefore, that growth stimulation is not induced by the hormone per se but rather by what is identified as "sulfation factor." This may be either the hormone itself, modified in some way that activates it physiologically, or a quite different substance produced somewhere in the body in response to HGH stimulation.

In any case, and whatever its nature or source, sulfation factor appears to be the actual substance that acts on cartilage to produce what, for simplicity, are usually called HGH growth effects – always assuming, of course, that the tissues are sensitive to its influence. The other effects of HGH (e.g., insulin release, fat mobilization) may not act through this intermediate.

One of the more interesting aspects of growth hormone disorders is that among them they seem to cover the entire sequence of HGH metabolism described above. That is to say, hereditary conditions have been identified that apparently interfere with the hypothalamic releasing factors, with the pituitary manufacture and/or release of HGH (several types), with formation of sulfation factor, and with the tissues' response to it.

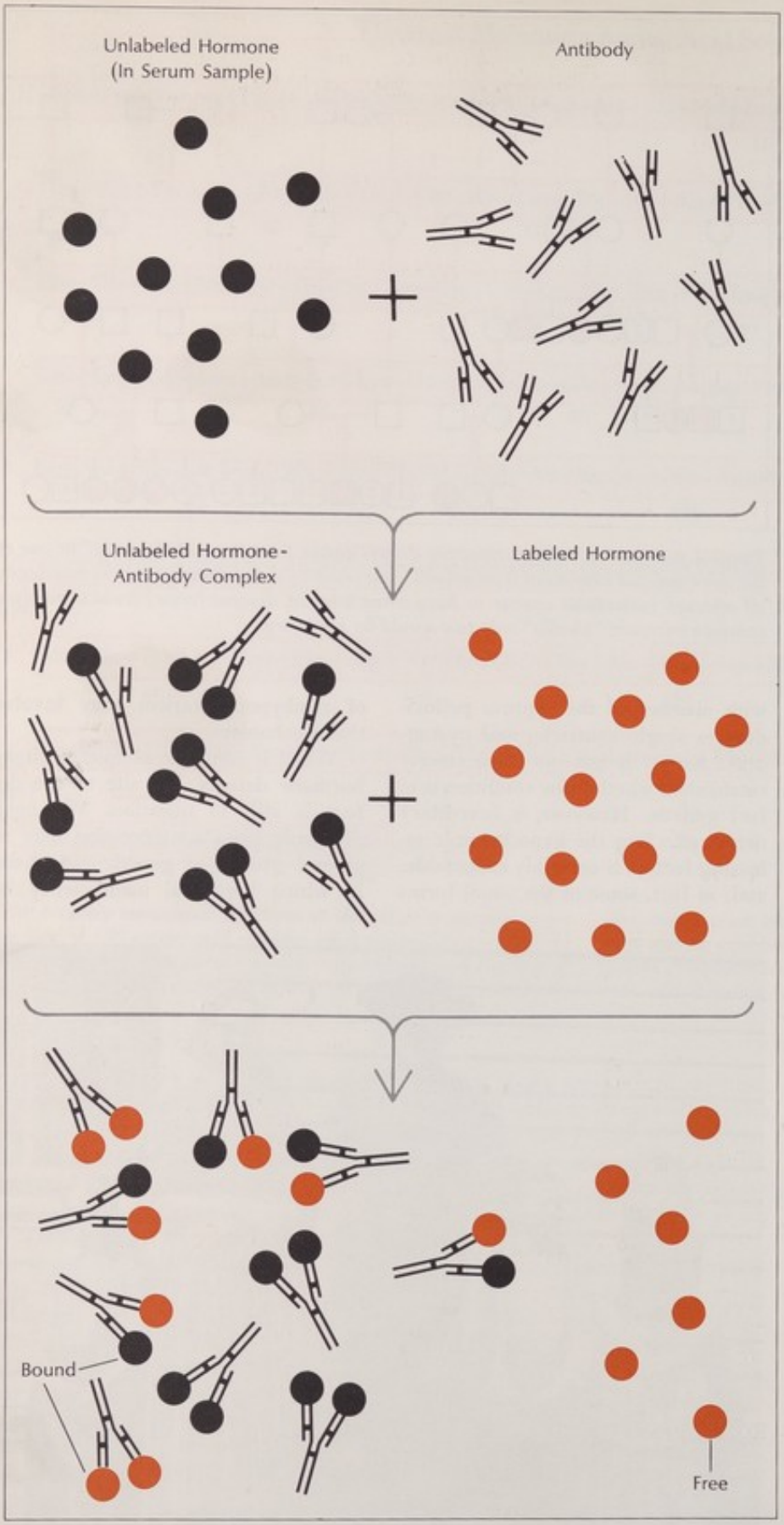
The hypothalamic abnormalities are the most diversified and, in the present context, of minor significance. In many of them (e.g., anencephaly, holoprosencephaly) growth failure does not develop because of death at an early age, while in others pituitary insufficiency does not occur until late childhood or adult life, so that growth is affected only minimally. Recently, Grumbach and Kaplan in San Francisco described a hypothalamic abnormality in pituitary dwarfs, coupled with absence of the septum pellucidum, a single ventricle, and nystagmus. As yet it has not been clearly established whether the condition is in fact genetic. However, a hereditary defect affecting the hypothalamic releasing factors is certainly expectable, and, in fact, some of the usual forms of panhypopituitarism may involve this mechanism.



Tangled pedigree of panhypopituitary dwarfs comes from "extended family" in one of the more isolated cantons of Switzerland. As a result of many consanguineous marriages, all affected individuals appear to have three lines of descent (color) from a putative common ancestor; "ideally" only two would be necessary.

When it comes to simple pituitary hormone defects, the site of the defect is still in question. We need give only brief consideration here to several groups of genetic conditions in which hormonal insufficiency is,

Three panhypopituitary dwarfs, aged 32, 26, and 20, are shown here with their two normal siblings. All were born to a single family of Canadian Hutterites, an isolated, inbred community of the sort where recessive anomalies frequently crop up.



Radioimmunologic assay is used to measure serum HGH. Antibody, added to serum sample, complexes with endogenous HGH; labeled HGH is then added to occupy "free" antibody sites. After bound and free labeled HGH are separated, their relative radioactivity shows amount of free sites and, inversely, of endogenous hormone.

in effect, secondary to absence of or gross malformation of the pituitary. These include a recently described syndrome in which a very small sella turcica is located in a morphologically abnormal sphenoid bone, producing generalized deficiency of all the pituitary hormones, growth failure, hypoglycemia, and motor and mental retardation; it appears to be inherited as an autosomal recessive.

The subtler pituitary abnormalities occur in two sharply distinct forms, the second of which may itself be divisible into at least two different conditions.

The first of these was originally called asexual ateliotic dwarfism, but is now more commonly known as panhypopituitary dwarfism (PPD). Even this name is not wholly accurate, since it does not always affect all the pituitary hormones in all cases. Invariably, however, there is a severe deficiency in both HGH and the gonadotropins, producing not only growth failure but also sexual immaturity (primary amenorrhea in females, small penis and testes in males, absence of secondary sexual characteristics in both). Increased subcutaneous fat and high voice are also present, but curiously these seem unrelated to the sexual hormone deficiency, since, as we shall see, they are present even in cases of isolated HGH deficiency.

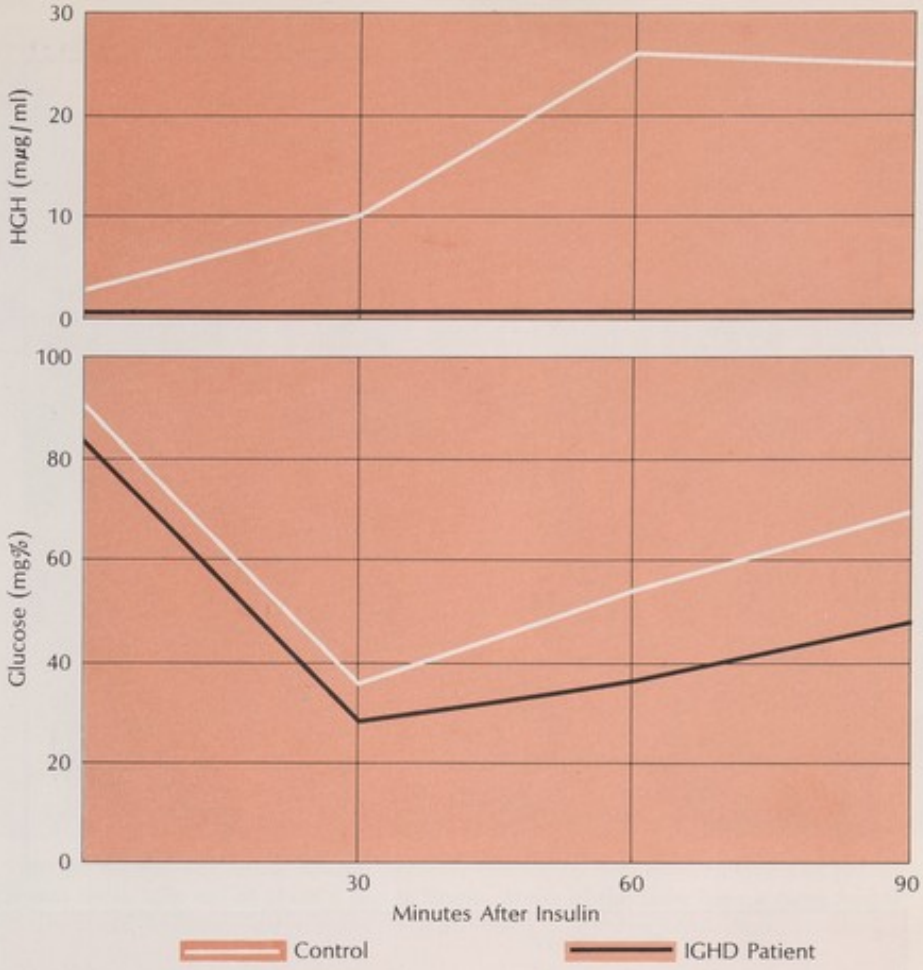
It should be stressed that the great majority of PPD cases are sporadic rather than demonstrably familial, and most of these probably involve no genetic defect. Some are probably due to pituitary disease, e.g., cyst formation, and several cases are known in which the disease affected only one of two identical twins, making clear that no genetic component is involved. I might also note that in most PPD cases, hereditary or otherwise, in which no pituitary abnormality can be demonstrated, we cannot with certainty exclude an absence of the hypothalamic releasing factors as the cause of the trouble – rather than failure of the gland to manufacture hormones. It has recently been demonstrated that in the majority of cases of sporadic PPD it is the hypothalamus rather than the pituitary that is at fault.

Though most PPD cases are not familial, well over 20 families are known with multiple affected members. The presumption that these reflect a ge-

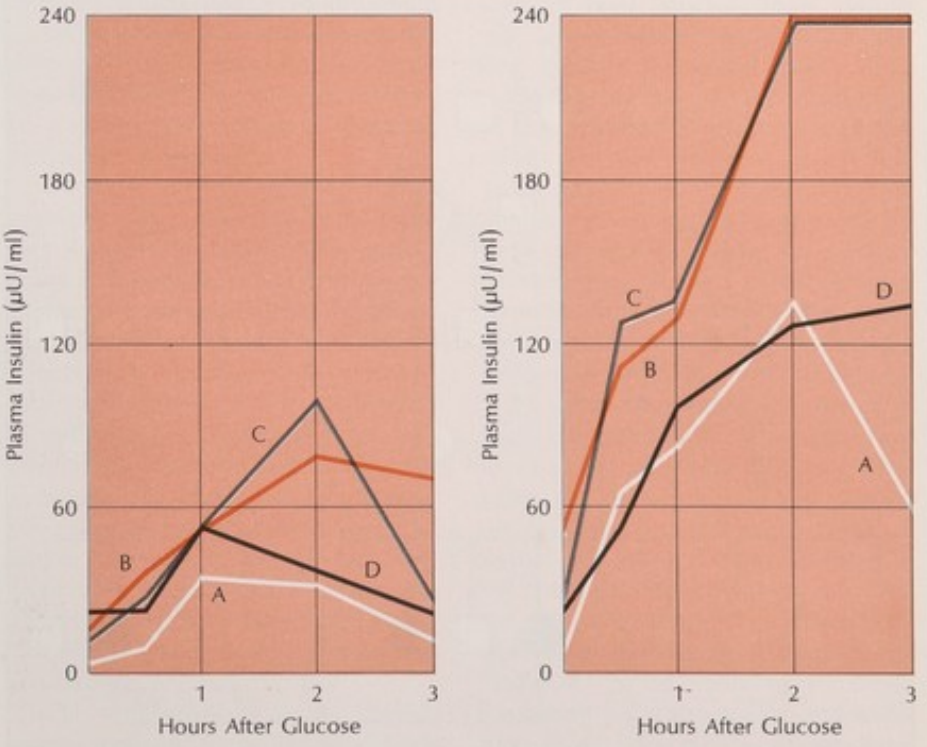
netic anomaly is strengthened to near-certainty when we find that most of them have originated in inbred communities – notably in Switzerland, on the Yugoslav island of Krk, and in Hutterite communities of western Canada. The same genetic disease has also been reported in families where there is no evidence of parental consanguinity. In both groups, the hereditary patterns make clear that we are dealing with an autosomal recessive disorder. At least two families are known, however, in which the pedigree suggests that the disease may also exist in a sex-linked recessive form.

For a long time it has been known that proportionate dwarfism can exist without abnormalities of the gonadotropic hormones. One of the most famous pituitary dwarfs, P. T. Barnum's "General Tom Thumb" (Charles Stratton) married another dwarf, Lavinia Bump, and the lady is said to have become pregnant, though she suffered a miscarriage. And her sister, also a dwarf (or midget), is definitely known to have died in childbirth. We have been able to document isolated growth hormone deficiency in a dwarfed cousin of these sisters.

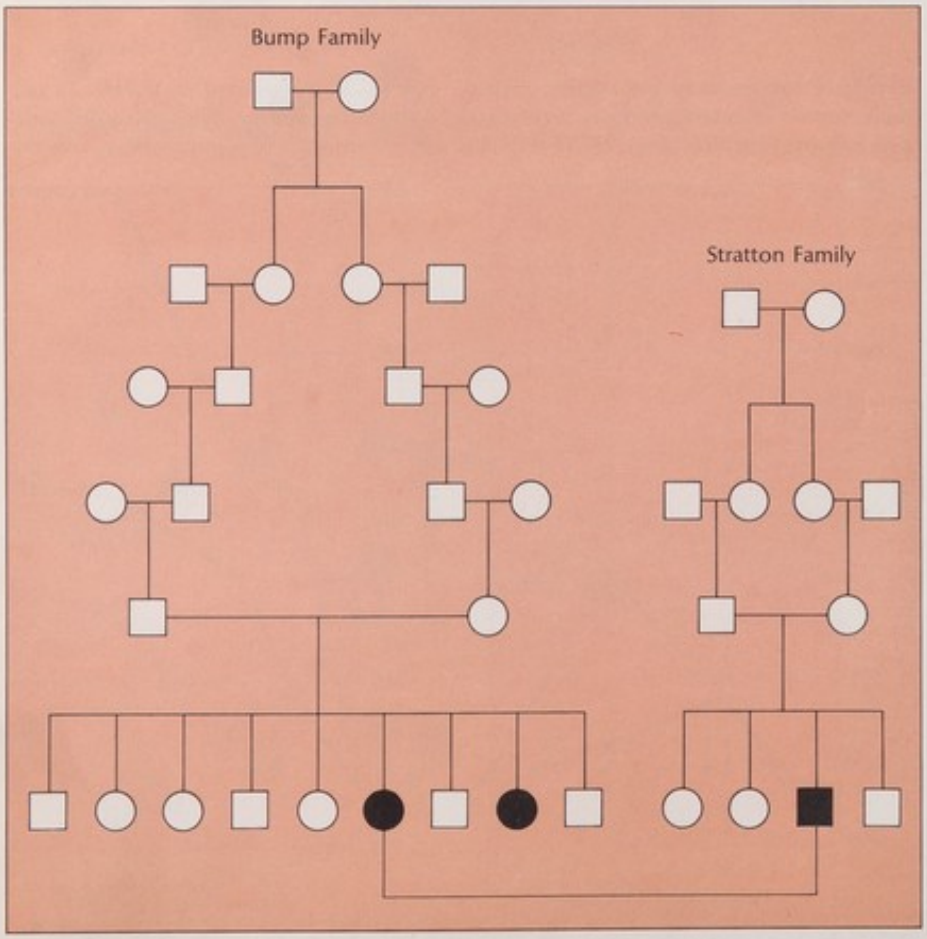
Clinically, cases of "sexual ateliotic dwarfism" or isolated growth hormone deficiency (IGHD) are indistinguishable from those of PPD until puberty, although the new immunoreactive gonadotropic hormone assays may allow this distinction. HGH deficiency in childhood coupled with clear evidence of thyroid-stimulating hormone (TSH) deficiency, e.g., myxedema, or ACTH deficiency could safely be presumed to be the panhypopituitary type. Even the nonoccurrence of puberty at the normal age is not diagnostic, since its average age of onset in IGHD patients is about 16 years in females and 19 years in males (in one case, as late as 27). Significantly, when such individuals are treated with HGH, puberty is often accelerated, suggesting yet another physiologic involvement of the hormone: stimulation of gonadotropin production and/or increased tissue response to the gonadotropins (or to the gonadal hormones themselves). Untreated males undergo a sudden "growth spurt" during puberty, showing that the well-known growth-promoting effects of



Feedback system involving insulin, glucose, and HGH is deranged in IGHD. In normals, insulin injection produces hypoglycemia, releasing HGH, which stimulates glucose release; in IGHD cases, HGH stays low and recovery from hypoglycemia is slower.



HGH also potentiates the release of endogenous insulin in response to hyperglycemia. Four IGHD individuals showed a decreased insulin response to injections of glucose (left), but responded at normal levels after three to five days' pretreatment with HGH.



Wedding of "General Tom Thumb" (Charles Stratton) and Lavinia Bump was a famous P. T. Barnum "stunt"; the bride was attended by her equally diminutive sister. Bump and Stratton pedigrees show that the three were offspring of consanguineous marriages.

testosterone do not require the presence of HGH.

Other interesting facts arise from consideration of the reproductive physiology of these individuals. As suggested above, they appear to be normally fertile without hormonal treatment. Moreover, their infants are of normal size (they must, of course, be delivered by cesarean) even if they ultimately prove to have HGH deficiencies. This indicates that growth in utero is not controlled by growth hormone, either fetal or maternal – a conclusion that had already been suggested by other findings. Finally, the IGHD mother lactates normally. Also, in humans and monkeys, though not in other mammals, growth hormone and prolactin (the milk-producing hormone) could not be distinguished chemically until very recently and it had been speculated that they were one and the same substance; it is now clear, however, since chemical separation of the two hormones has been accomplished, that they are not, though the two molecules are evidently very similar in primates.

Like PPD, IGHD occurs both sporadically and familially, but unlike PPD the majority of cases seem to be inherited – in most cases, in an autosomal recessive pattern. However, a number of patients are on record who differ from "typical" IGHD patients in various physiologic parameters, e.g., they may be insulin resistant, and in at least some of whom autosomal dominance may be the inheritance pattern.

Thus far, I have said little about therapy, mainly because the average physician is unlikely to be able to treat a case of pituitary dwarfism because of scarcity of HGH. It is worth noting, however, that both PPD and IGHD can be treated successfully. In PPD, HGH injections are coupled with (where necessary) oral thyroxin and/or corticosteroids, plus, at the appropriate age, male or female hormones. The HGH treatment is discontinued when adequate stature is obtained, but lifelong replacement therapy of the other substances is necessary. With IGHD, of course, only the HGH treatment is required. GHRH is a small peptide whose actual structure should soon be discovered; this molecule will probably be adaptable to inexpensive large-scale commercial synthesis so that GHRH may one day be

the treatment of choice in these cases of pituitary dwarfism secondary to a hypothalamic abnormality.

Because of the difficulty in distinguishing the two conditions and the epiphyseal closure produced by sex hormones, it is preferable to delay sex hormone treatment until growth is completed. As mentioned earlier, in IGHD, HGH alone will often bring on puberty "in a rush," thereby incidentally clarifying the diagnosis, while in PPD, premature administration of sex hormones can damp the growth response to HGH by triggering closure of the epiphyses. (The lack of epiphyseal closure in untreated PPD individuals means that they may continue to grow, albeit very slowly, until well into middle age. Several cases are known of circus midgets who eventually "grew out of their jobs.")

In the case of another type of familial dwarfism, however, there is apparently little or no response to HGH. And in fact when we assay serum HGH in these individuals, its levels are either normal or (in most cases) above normal. Yet clinically there is little to distinguish them from the types just described. More refined tests, however, show a number of differences, most notably a deficiency of sulfation factor even after treatment with exogenous HGH. This suggests that the basic defect is probably some problem in the formation of sulfation factor, but there is an alternative explanation: a minor alteration in the HGH molecule that is just sufficient to render it functionally inert while leaving its immunologic identity intact. (Alterations of this subtle type, the so-called CRM-positive mutations, have been observed in bacterial proteins.) The altered molecule would presumably still be able to bind to, and therefore inactivate, the normal receptor sites in the tissues, including those where sulfation factor is formed, thereby blocking the action of exogenous, normal HGH. Either mechanism would be consistent with the high serum levels of HGH in these patients, since either the missing sulfation factor or the (hypothetical) altered molecule might well fail to activate the feedback mechanisms that normally control pituitary hormone production, leading to a serum buildup of immunologically normal but ineffectual HGH.

Pygmies studied in Central African Republic are shown here with members of the research team, who are all about 6 feet in height. Males of this tribe have a mean height under 5 feet, females average under 4 feet 10 inches (153.6 and 144.7 cm, respectively).

This type of dwarfism (often called the Laron type, after its Israeli discoverer) is inherited as an autosomal recessive. Originally, it turned up only among Oriental Jews, but has since been described in non-Jewish families in a number of European countries and in North America.

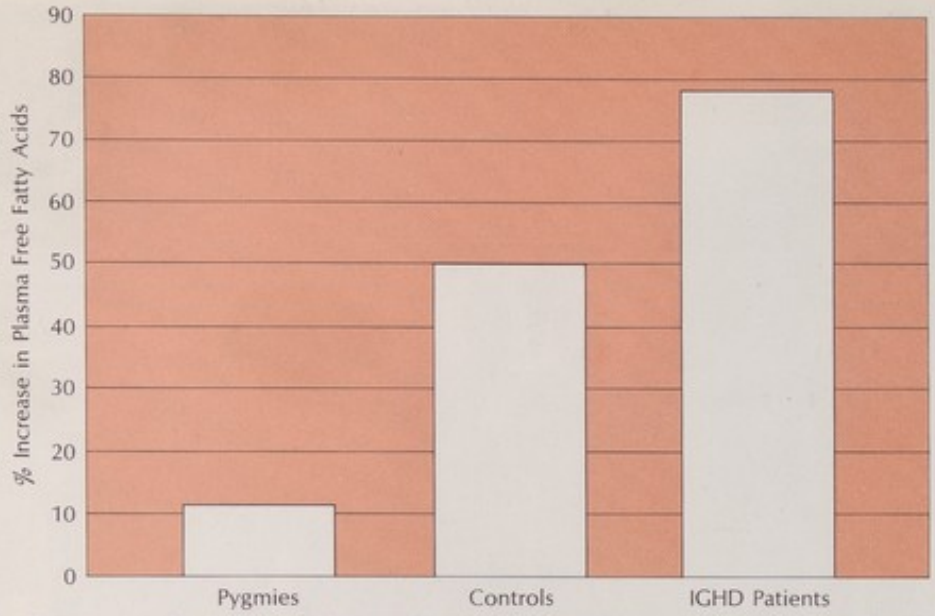
We have, then, identified, at least provisionally, types of inherited pituitary dwarfism in which the hypothalamus is abnormal, in which the hypothalamus is normal but the pituitary's manufacture of HGH is not, and in which both pituitary and hypothalamus are normal but the generation of sulfation factor is not. Logically, one would expect to find somewhere individuals in whom all the foregoing processes are normal but the tissue response to HGH is not. In fact they have been found, and they are not isolated individuals either but an entire ethnic group of several thousand people: the African pygmies.

The short stature of the various pygmy peoples (who are found not only in Africa but in New Guinea, the Philippines, and the Andaman Islands) has been something of an anthropologic puzzle. Some writers have suggested that in the dense forests most of them inhabit, short stature may have proved advantageous in getting through the underbrush.

Be that as it may, the physiology of the pygmies' short stature is certainly an interesting problem. A few years ago, I had the opportunity of joining an expedition, organized by Professor Cavalli-Sforza, to a rather isolated part of the Central African Republic, where we were able to carry out a number of revealing physiologic tests on pygmies and other indigenes.

In gross characteristics, the pygmies are not unlike other pituitary dwarfs, though they are somewhat taller on the average. They possess the childlike proportions of trunk to limbs – not when compared with Europeans but distinctly so in comparison with such neighboring black populations as the Bantu. However, the "childish" facies (bulging forehead, hypognathia), fatty trunk, and wrinkled skin characteristic of other pituitary dwarfs are missing.

Physiologically, their serum HGH levels turned out to be altogether nor-



That pygmies' small size is due to tissue insensitivity to growth hormone was demonstrated by studies such as the one graphed above. When rises in plasma free fatty acid concentrations following IV infusions of HGH were compared in pygmies, controls, and IGHD dwarfs, it was found that the pygmies manifested a mean rise of only 11% above basal levels while the rise in controls was 50% and that in the IGHD cases 78%.

mal. Injections of insulin produced the normal sharp rise in plasma HGH. The resulting hypoglycemia, however, was marked and prolonged – as it is in most pituitary dwarfs; evidently the released HGH was not doing its normal job of releasing glucose into the blood. Two explanations are possible for these findings: either they were manufacturing metabolically inert but immunologically normal HGH (as postulated in Laron-type dwarfism) or their peripheral tissues were insensitive to the actions of HGH.

The following year Dr. Thomas Merimee and I returned to Central Africa with Professor Cavalli-Sforza and carried out further experiments using exogenous HGH. In every parameter we measured – insulin response, drop in serum nitrogen levels, rise in plasma FFA – there was no metabolic response to normal HGH, even with enormous doses.

In a following expedition, environmental or nutritional factors were excluded as a cause of the pygmies' HGH resistance. It was also found that only the delayed effects of HGH (lipolysis, nitrogen retention, insulin release) were blunted, while the immediate effects of the hormone (transient decrease in glucose and FFA concentrations) were normal, suggesting that the defect lies at the level of DNA-mediated protein synthesis.

There seems no doubt, then, that the pygmies owe their short stature at least in part to a generalized tissue unresponsiveness to HGH. On the face of it, such unresponsiveness, involving many types of tissue and organ systems, might seem improbable, but in fact it might be produced by something as simple as an abnormal cellular receptor site that failed to "recognize" the HGH molecule. Similar generalized peripheral insensitivity is thought to be responsible for a number of well-known diseases, all of them simply inherited, including pseudohypoparathyroidism, testicular feminization syndrome, and nephrogenic diabetes insipidus.

The mode of inheritance of pygmy "dwarfism" (obviously it is inherited) has not yet been elucidated. There is, indeed, no reason to suppose that HGH production, or tissue response to it, are the only genetic factors involved in stature differences, though they certainly play an important role in them (including, one suspects, in the differences between various ethnic groups, e.g., Swedes and Vietnamese). The second possibility is that the tissue insensitivity is, like many human characteristics including skin color, a polygenic trait. This would explain the hybrids' intermediate stature. Studies of variation among pygmy-Bantu hybrids and their offspring would help to resolve the uncertainty between one gene and many genes, but very few such individuals are known thus far. In two such individuals tested, however, HGH responsiveness was normal. This form of HGH resistance may well be the cause of a type of short stature in our society.

As we have seen, then, the genetic growth hormone disorders, obscure and rare though they are, are throwing light not merely on the nature of HGH itself but also on some more general problems of growth and of hormones, e.g., the relationship between HGH and human prolactin. These in themselves seem to me quite respectable dividends. In addition, however, there is a further fact worth keeping in mind: As the medical significance of infectious and nutritional disease continues to diminish – as we can expect it to do in our society, at least – the relative importance of the "pathologic residuum," of which genetic disease is a major component, will go up. The more we know about genetic diseases of any kind, the better prepared we will be to deal constructively with this shift in emphasis of medical practice.

16

Human Cell Culture Techniques in Genetic Studies

WILLIAM J. MELLMAN
The University of Pennsylvania

On theoretical grounds, culture of the human diploid cell represents an almost ideal system for the study of genetic diseases. By definition diploid cell lines possess the full chromosomal complement of the individual from which the cells are taken; if cultured cells of succeeding generations express a defect present in the cells used to initiate the culture, this is strong presumptive evidence that the defect has a genetic basis. One would not expect an acquired defect to be transmitted (unless the cell cultures carry an infective agent). Moreover, the genetic nature of a defect can often be confirmed by parallel studies on cells from other family members. For recessive conditions, the defect can be expressed by cells from both carriers and affected individuals.

Given these principles, it is not surprising that intense and continuing activity was stimulated among human cell geneticists when first Tjio and Puck in the late 1950's, and then Hayflick and Morehouse, demonstrated that one could culture human cells with relative ease and that the resulting cell lines would retain the chromosomal characteristics of the tissues from which they were derived. Now, 15 years later, it is fair to state that this activity has been amply rewarding in terms of increased understanding of the often subtle chromosomal alterations, and the even more subtle biochemical aberrations, that underlie genetic diseases. Some of the specific information obtained by cell culture techniques is described elsewhere in this volume, notably in the presentations on the Lesch-Nyhan syndrome by Seegmiller (Chapter 10), on mucopolysaccharidoses by Neufeld (Chapter 14), and on Tay-Sachs disease by Kaback and O'Brien (Chapter 25). This chapter, then, will be concerned with a more general discussion of the areas under study by cell culture methods, and on the approaches being used.

It is important to keep in mind that the starting point of most cell culture studies is a patient who is expressing a defect. The investigator recognizes the defect and the probability that it is an expression of some biochemical aberration, and from experience he may have some clue as to where to look in order to define the defect specifically. He begins by taking cells from the patient and culturing them, hoping that by using sophisticated biochemical methodology he can identify the actual anomaly. Cell culture studies are not usually critical to learning that the defect is hereditary, although they may be confirmatory. A good example in this context is work reported by Danes and Bearn in 1965. It had been known from family studies that Hurler's syndrome, like other mucopolysaccharidoses, was inherited. However, by using diploid fibroblast cultures, these investigators were able to suggest that cultured cells from Hurler's patients expressed the genetic defect by an increased degree of metachromasia when stained with certain dyes. The value of this observation in stimulating further diagnostic and therapeutic studies, and its potential value in monitoring and designing therapy for the mucopolysaccharidoses, is documented in the chapter by Neufeld, cited above.

When one talks about human cell cultures, it must be kept in mind that from any higher species, there are only a few cell types that can be cultured with routine success. In genetic studies, the standard for most investigations has been the fibroblast. More recently developed and certainly useful have been systems employing lymphoblasts. Finally, amniotic cells have been cultured in a number of laboratories and in certain situations they are used for antenatal diagnostic studies.

The failure to date to consistently develop viable cultures from cells derived from specific organs such as the liver probably reflects the fact that the more specialized

the cell, the more specialized the milieu it requires for survival and replication. And we have not yet learned to duplicate "organ-specific" environments in vitro. By an extension of this reasoning, one can also suggest a rationale for the preeminence of fibroblasts in diploid cell culture systems. These cells after all form the matrix of virtually all tissues and can be cultured from almost any organ. To be sure, the great majority of fibroblasts used by human geneticists are obtained from skin, largely for reasons of availability. However, fibroblasts can also be obtained from kidney, liver, or almost any organ that can be biopsied. In fact, their ubiquity contributes to the difficulty of culturing differentiating organ-specific cells such as hepatocytes. Generally speaking, the fibroblasts accompanying the specialized cells in the specimen will rapidly overgrow the latter in culture. Attempts to eliminate fibroblasts from organ-specific tissues at the time of culturing have not been effective to date in facilitating the culture of the specialized cells.

A second cell type now being used in genetic studies, the lymphoblast, appears to require some type of virus component or product to maintain the cell line in vitro, an inherent problem that introduces a variable which cannot be controlled completely. This difficulty can in part be circumvented by always employing the same virus under the same conditions in the cultures. Nevertheless, over the long periods required to obtain adequate genetic information, the investigator is handicapped by the problem of differentiating between viral effects and those specific to the cultured cells. An alternative to this type of lymphoblast culturing is the so-called short-term lymphocyte culture, obtained by stimulating cells to proliferate with phytohemagglutinin or other mitogenic agents. Unfortunately, this approach provides sufficient numbers of dividing cells for less than a week at best, a major limitation for many genetic studies.

The lymphoblasts do have certain advantages over fibroblasts. It must be remembered that in order to grow, survive, and function, the fibroblasts must be attached to a surface. In many situations, they must also grow side by side with a considerable density before they will function in various ways. These characteristics are certainly related to their in vivo behavior in solid tissues. On the other hand, lymphoblasts grow free in suspension in the culture medium, making it easier to study many aspects of their metabolic activity. For example, transport of different substances in and out of a cell is more readily observable if that cell is floating in a medium, than if it is one of many irregularly shaped cells crowded by and crowding its neighbors. A second significant advantage is that lymphoblasts can be readily cloned, and will give rise to definable cell populations all derived from a single parent and all normally manifesting characteristics identical with that parent. With lymphoblasts therefore, the investigator can apply techniques analogous to those used in bacterial genetics. One can, for example, look for mutations simply by seeking out inherent variations within a population of cells. The way to do this is to select individual cells from large masses, grow them, and then study their progeny as individual colonies.

While this can be done with lymphoblasts, such efforts, for the most part, with normal human fibroblasts have been largely unsuccessful. An important exception has been the successful use of fibroblast cultures to study the Lyon hypothesis (the concept that in any cell carrying the normal female complement of two X chromosomes, one of the chromosomes will have been inactivated on a random basis). An important consequence of this has been to show that for X-linked biochemical traits expressed in fibroblasts, if one takes cells from a female heterozygous for two alternate traits at the same genetic locus, and clones those cells, the populations derived will either express or not express one of the two traits. The clones will not be mixed with respect to the allele involved (see Seegmiller, Chapter 10).

However, if one wishes to take these clones and study them further for specific biochemical properties, he finds that the cells soon fail to divide. The human fibroblast cell line has a definitely limited number of generations in culture. In contrast, the lymphoblast seems to have an infinite reproductive capacity. This property appears to be related to the presence of the virus with which it is cultured. However, the limited life expectancy of the fibroblast culture lines suggests some interesting approaches to the study of the biologic determinants of aging (see table on this page).

This approach has validity if a correlation between the finite lifespan of cells in culture and events taking place in the intact organism can be documented. Certainly, cell cultures do provide us with a mechanism for learning about the longevity and survival of proteins and enzymes, some of which are involved in inborn errors. It is possible to tag an enzyme

Some Metabolic Properties of Diploid Cells Studied During Aging

Parameter	*Variation with Cell Age*
Glycolysis	=
Glycolytic enzyme	=, ↓
Pentose phosphate shunt	↓
Permeability to glucose	=
Glycogen content	↑
Mucopolysaccharide synthesis	↓
Respiration	=
Respiratory enzymes	=
Lipid content	↑
Lipid synthesis	↑
Protein content	=, ↑
Permeability to amino acids	=
Transaminases	=, ↓
Glutamic dehydrogenase	=
Nucleohistone content	=
Collagen synthesis	↓
DNA content	=, ↓
RNA content	↑
Nucleic acid synthesis	↓
RNA turnover	↑
Lysosomes and lysosomal enzymes	↑
Alkaline phosphatase	=

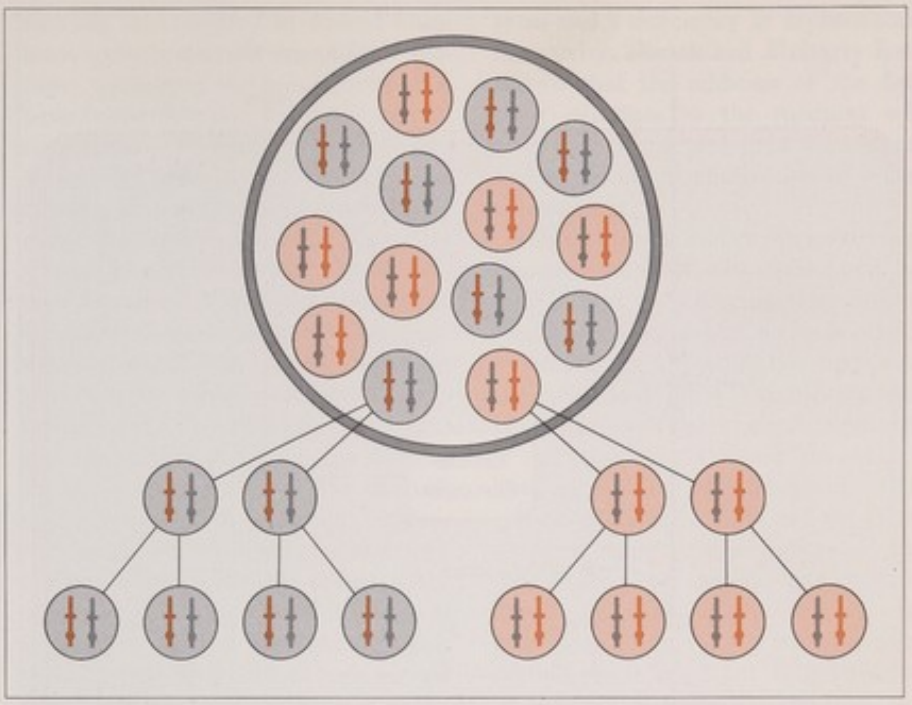
↑, *increase with age;* ↓, *decrease with age,* =, *no change.*
Reprinted with permission from V J Cristofalo, Advan Gerontol Res *4:45, 1972*

radioactively. From precise knowledge of when the tagging is done, measurement of the half-life of the enzyme within the cell becomes feasible.

As noted, the third source of human diploid cells for culture is the amniotic fluid. The obvious advantage of amniotic cells is their now clearly established fetal origin, the basis of their usefulness in antenatal diagnosis. However, standardization of cultures of amniotic cells has been difficult, largely because of the considerable variability in the source of these cells. It is currently believed that cells found in amniotic fluid can derive from fetal bladder, fetal trachea, fetal skin, and from the amnion itself. Starting with so heterogeneous a population, it is not surprising that the cell growth in culture tends to reflect this variability; growth rates can be very different from culture to culture, as can morphology and often biochemical behavior. Moreover, one cannot always draw analogies between the behavior of fetal cells and that of cells obtained from individuals postnatally. Kaback et al have shown that fetal cells obtained at the time of abortion, cells from the amniotic fluid of that fetus, and cells from the mother all behave very differently when compared on the basis of certain enzyme activities. Despite these differences, amniotic cells have already proven valuable in a number of situations involving antenatal diagnosis (see Dancis, Chapter 24).

It can be seen that the diploid fibroblast cultures are still in a position of preeminence as tools of genetic study. This remains true, at least temporarily, even though only a fraction of the inherited biochemical disorders express themselves in fibroblast cultures, and among the many that do not are such quantitatively important disorders as phenylketonuria. It seems almost certain that to obtain a valid in vitro model of PKU it will be necessary to culture liver cells directly and, as noted, it has not yet been possible to establish cell lines from such highly differentiated tissues. However, since fibroblasts presumably contain the genetic information present in liver cells, we should not despair of finding ways to "turn on" the gene responsible for PKU in fibroblast cultures.

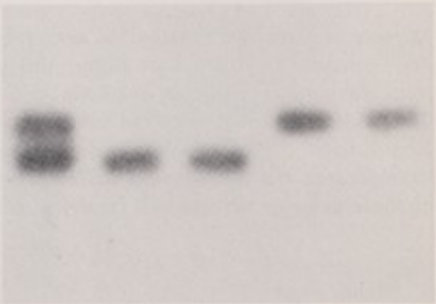
This then fairly well sums up the status of human cell cultures, except for the omission of tumor cell lines. Many such have been developed but in the main they have little value for research in genetic disorders. Let us turn now to the techniques that have been and are being employed by tissue culture workers. For purposes of this review, one can assume that the most common type of culture is one employing skin fibroblasts.



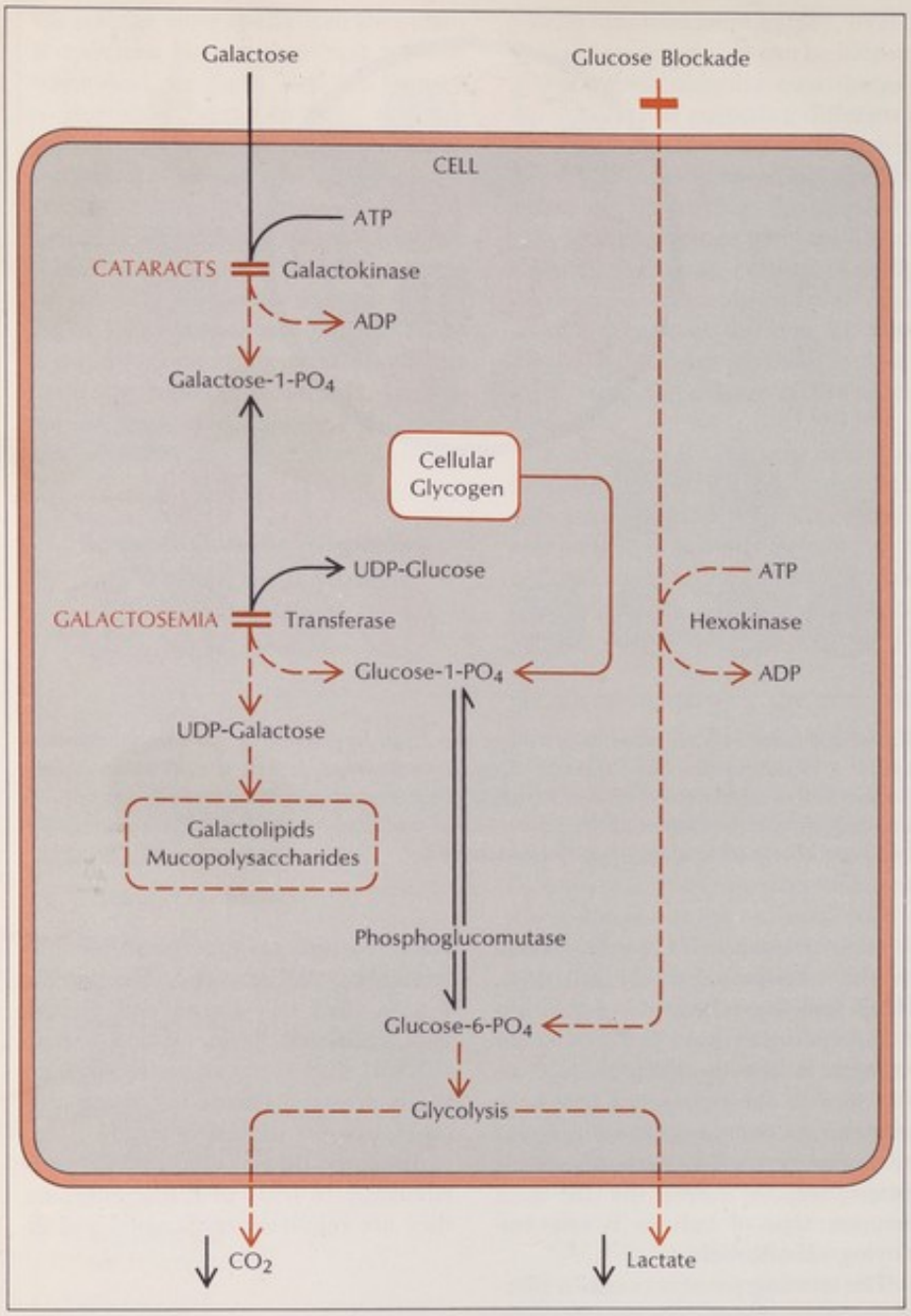
Fibroblast cultures have been used to test the Lyon hypothesis of random inactivation of the X chromosome. The rationale, diagrammed above, is that if cells taken from a female heterozygous for an X-linked allele are cultured and then cloned, only one of the two alternative traits will be expressed in any particular clone. Such segregation has been observed in a number of experiments.

The starting point is to take a biopsy specimen two to four mm in diameter from the dermal layer, in which one can expect to obtain some fibrous tissue. The tissue specimen is very finely minced in a small quantity of tissue culture medium, which includes some serum in addition to one of several kinds of prepared medium (preferences here are quite individual). The minced tissue-in-medium is placed in an incubator under an atmosphere of carbon dioxide and left for about a week. At this point, if the culture is likely to be successful, one sees adherence of the tissue to the plastic (or glass) surface of the container. A few days to a week later, some amount of cellular proliferation can be observed. With proper nutrition, this continues at a predictable rate until the cells have reached an appropriate density (roughly a million cells). The population is then trypsinized and divided into additional flasks. When several million cells have been produced, some flasks are then frozen for future use; others are for immediate study.

Because these cells divide more efficiently in less crowded conditions, they are regularly trypsinized and di-



The ability of electrophoresis to distinguish phenotypic expressions in cell culture of genetically determined enzymatic activity is exemplified. Shown are the starch gel electrophoretic patterns for two G-6-PD types in homozygous and heterozygous individuals (from H. M. Nitowsky, R. G. Davidson, D. D. Soderman, and B. Childs, Bull. Johns Hopkins Hosp. 117:6:370,1965).



Defects in galactose metabolism are expressed in fibroblast cell culture when the cells are "forced" to utilize that sugar; this can be accomplished by eliminating glucose from the medium (glucose blockade) and permitting cellular glycogen to be exhausted. Two known genetic flaws are schematized above, a deficiency in galactokinase, associated with cataracts, and a deficiency in galactose-1-phosphate uridyltransferase (transferase), the biochemical hallmark of galactosemia. Cells in culture having either of these enzyme defects will be deficient in their metabolism of galactose.

luted into additional culture vessels when they reach a high population density. When this is done, a high percentage of cells will be dividing; when division is arrested in metaphase, the condensed chromosomes are readily visible under light microscopy. Arrest at this stage requires the addition of colchicine or of another mitotic poison; processing the cell for chromosome study is much the same as with standard testing using peripheral leukocytes. It should be stressed here that with fibroblasts, the chromosomes of the dividing cells remain essentially unaltered from generation to generation, a fact which has made these cultures particularly suitable for genetic studies. However, one can induce alterations in chromosomal morphology by infecting the cultures with viruses or mycoplasma or by exposing the cells to irradiation. This, of course, suggests a research lead, the use of these culture systems to study viral and other types of mutagenesis.

Turning now from cell culture methodology to the applications of that methodology, one can exemplify the value of these studies by citing some specific contributions to knowledge with respect to various categories of the inborn errors of metabolism. These categories include: a) disorders of carbohydrate metabolism; b) disorders of macromolecular degradation; c) amino acid defects; and d) defects in the synthesis of the nucleotides of DNA and RNA.

Carbohydrate Metabolism

A number of genetic defects in CHO metabolism are expressed in diploid fibroblast lines. Among the first to be described, and still a good exemplar, is deficient galactose metabolism. Actually two different enzymatic flaws that are expressed in cultures have been described; the first affects activity of galactose-1-phosphate uridyltransferase (transferase) and is the underlying defect in galactosemia where lack of transferase activity leads to the accumulation of galactose-1-phosphate in the cells of individuals on lactose-containing diets. The other is a deficiency in galactokinase, clinically expressed, in large part, by formation of cataracts.

The ability to differentiate cells from galactosemic patients from those obtained from normal individuals was actually among the first demonstrations of the application of cell cultures to the identification and study of genetically-controlled metabolic abnormalities. In 1960, Krooth and Weinberg at NIH showed that when radioactive galactose was fed to normal fibroblasts in culture, they would metabolize it, while fibroblasts from galactosemics would not.

Since then, galactosemia has proved to be an extremely valuable model in demonstrating the capability of cell cultures for investigating inborn errors of metabolism. In many instances, disorders of carbohydrate metabolism are masked by redundancy in the enzyme systems involved; that is, a genetic lack of one en-

zyme is compensated for by utilization of other enzymes. Such choice of alternate pathways appears to be less available for galactose metabolism. Moreover, while cells in culture do not require galactose in their medium, they do require a simple sugar, usually glucose. Under certain experimental conditions, galactose can be substituted for glucose and the cells in culture forced to utilize this sugar. In these circumstances, an inability to metabolize galactose will be manifested. Since the work of Krooth and Weinberg, who essentially did what has just been described, techniques have been developed to assay the specific enzymes involved in galactose metabolism, using cells from patients with various genetic states affecting these specific pathways.

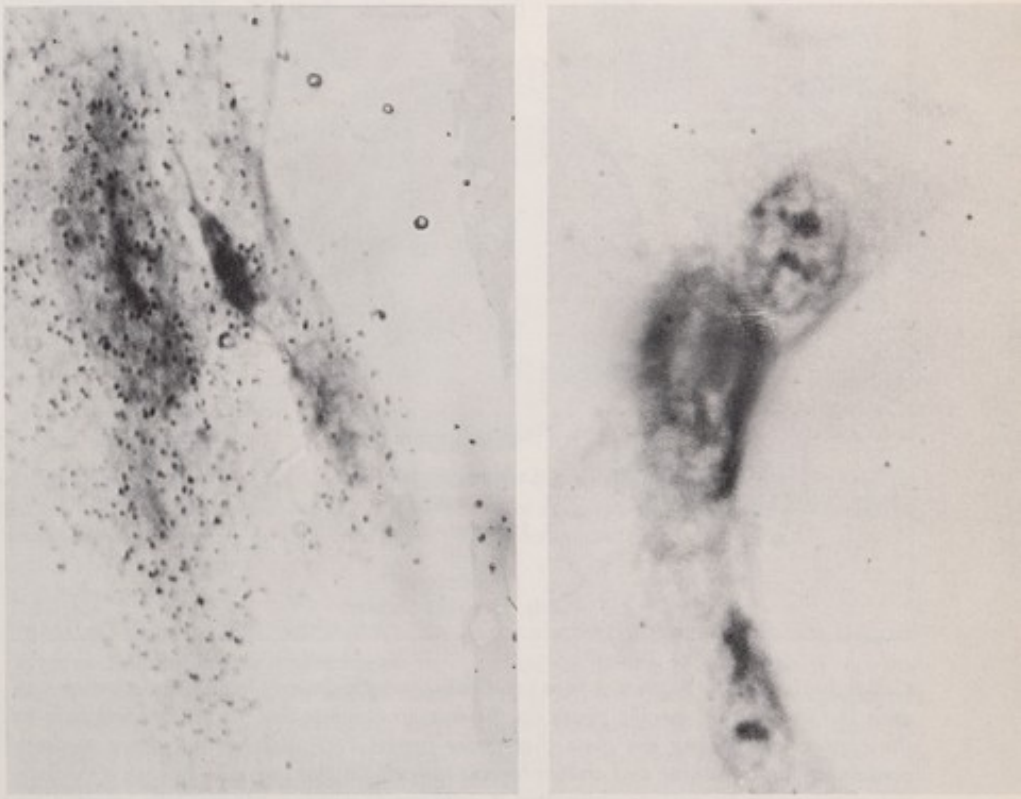
In broad outline, what is done is to measure the number of cultured cells in a sample, either directly by counting or indirectly by measuring the amount of either DNA or protein in the sample. The cells are then extracted and assayed for galactokinase, or transferase, or any of several other enzymes involved in galactose metabolism. Generally this is done on a comparative basis with control cells from normal individuals, with cells from galactosemic patients, with cells from individuals heterozygous for galactosemia, etc.

Defects in Macromolecular Degradation

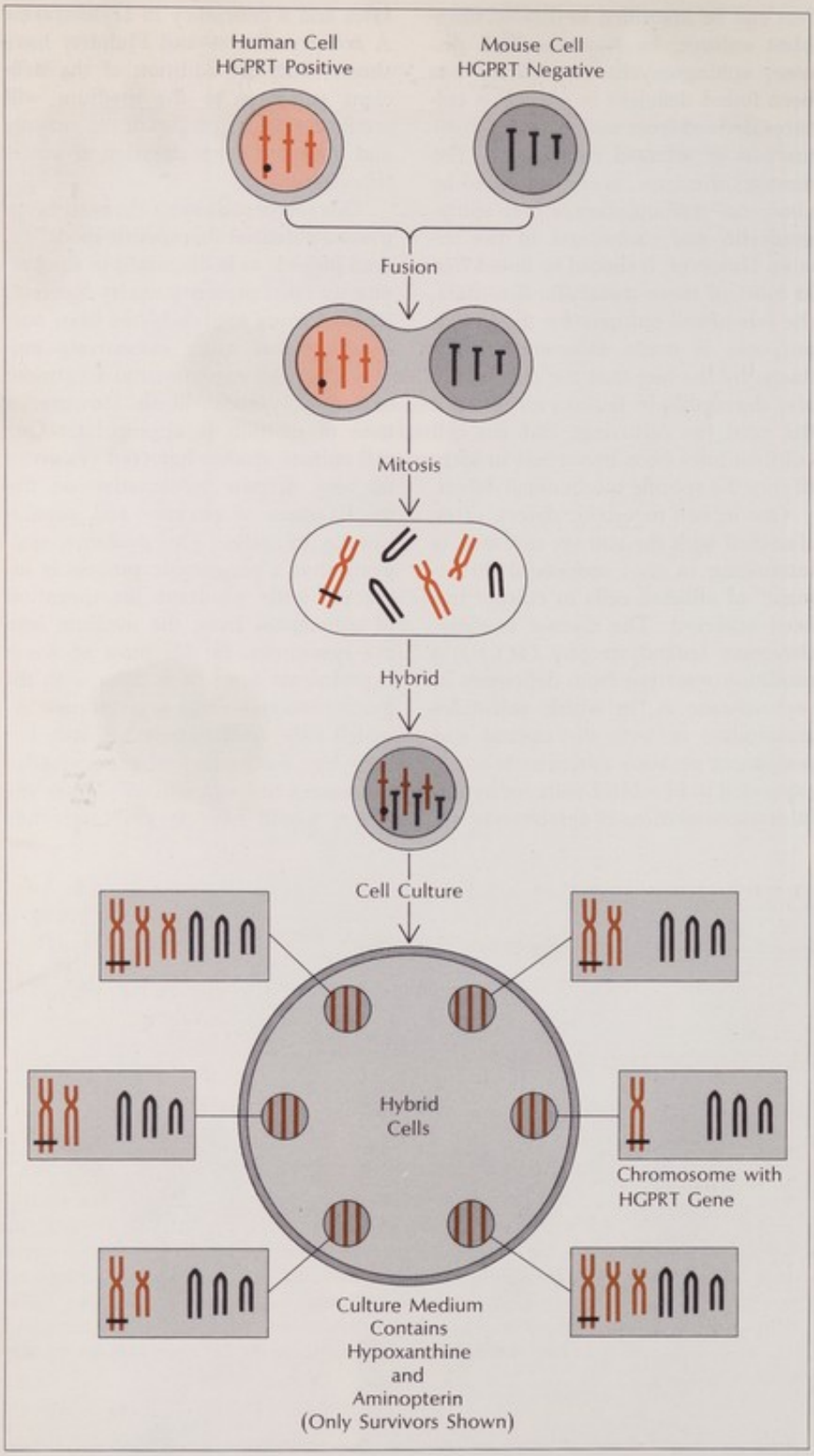
A number of the genetically-determined metabolic anomalies in lipid metabolism which are expressed in culture reflect diseases in which the defective enzyme's substrates are complex macromolecules. These diseases are storage diseases and we now realize they result from qualitative or quantitative deficiencies in various enzymes involved in the breakdown of macromolecules. In some cases, the relevant enzyme activity is directed against the carbohydrate moiety, in others against other components. To some degree, the situation with the mucopolysaccharidoses (Chapter 14) is typical. Certain glycogen storage diseases are also of this nature, as are the gangliosidoses – Krabbe's disease (galactocerebrosidosis), Gaucher's disease (glucocerebrosidosis), and Niemann-Pick disease. For most of these disorders, the specific biochemical defect can be identified in diploid fibroblast culture. In Niemann-Pick disease, sphingomyelinase activity has been found deficient in fibroblast cultures derived from both skin and bone marrow of affected individuals. The disease, of course, is characterized by abnormal accumulation of both sphingomyelin and cholesterol in the tissues. However, it should be noted that in most of these metabolic disorders, the role of cell cultures for diagnostic purposes is made somewhat superfluous by the fact that the defects are also detectable in leukocytes without the need for culturing. But the cell cultures have been important in identifying the specific biochemical defect.

One inborn metabolic defect, often classified with the storage diseases, is interesting in that successful "treatment" of affected cells in culture has been achieved. The disease is metachromatic leukodystrophy (MLD), a condition resulting from deficiency in arylsulfatase A, in which sulfatides accumulate in both the central and peripheral nervous systems. MLD is expressed in fibroblast cultures by cellular accumulations of cerebroside sulfates and a deficiency in arylsulfatase A activity. Porter and Fluharty have shown that the addition of the deficient sulfatase to the medium will result in cellular pickup of the enzyme and apparent normalization of sulfatide content.

This accomplishment obviously suggests a potential therapeutic modality. And indeed, as is discussed in the previously cited presentation by Neufeld, an analogous approach has been successfully and quite extensively employed in the experimental treatment of mucopolysaccharidosis. However, a note of caution is appropriate. Our cell culture studies have not yet given us very definite information on the mechanisms of enzyme and protein uptake by cells. The evidence suggests that a phagocytic process is involved, with resultant incorporation of substances from the medium into the lysosomes. So far most of these experiments have been done with digestive enzymes, like arylsulfatase A, which only need to be taken into the digestive mechanism of the cell (the lysosome) to be functional. Other enzymes would have to get into other



Cultured cells from normal individual (left) and from galactosemic individual (right) can be distinguished by the former's ability and latter's inability to incorporate tritiated galactose during incubation (reprinted with permission from H. Z. Hill and T. T. Puck, Science 179:1137, 1973).



Cell hybridization (e.g., mouse-man) and subsequent culturing have been employed to map the location of specific genes on the human chromosome, as exemplified here in the context of locating the gene that codes for HGPRT activity. A culture medium containing hypoxanthine and aminopterin is utilized for this purpose; in such a medium the presence of HGPRT is obligatory if cells are to survive. Mouse cells selected for absence of the enzyme are hybridized with human cells selected for the presence of HGPRT and placed in the medium. Only those fused cells that carry the chromosome with the human HGPRT gene will survive. Thus the common surviving human chromosome (in this case the X) can be identified as having the HGPRT gene locus.

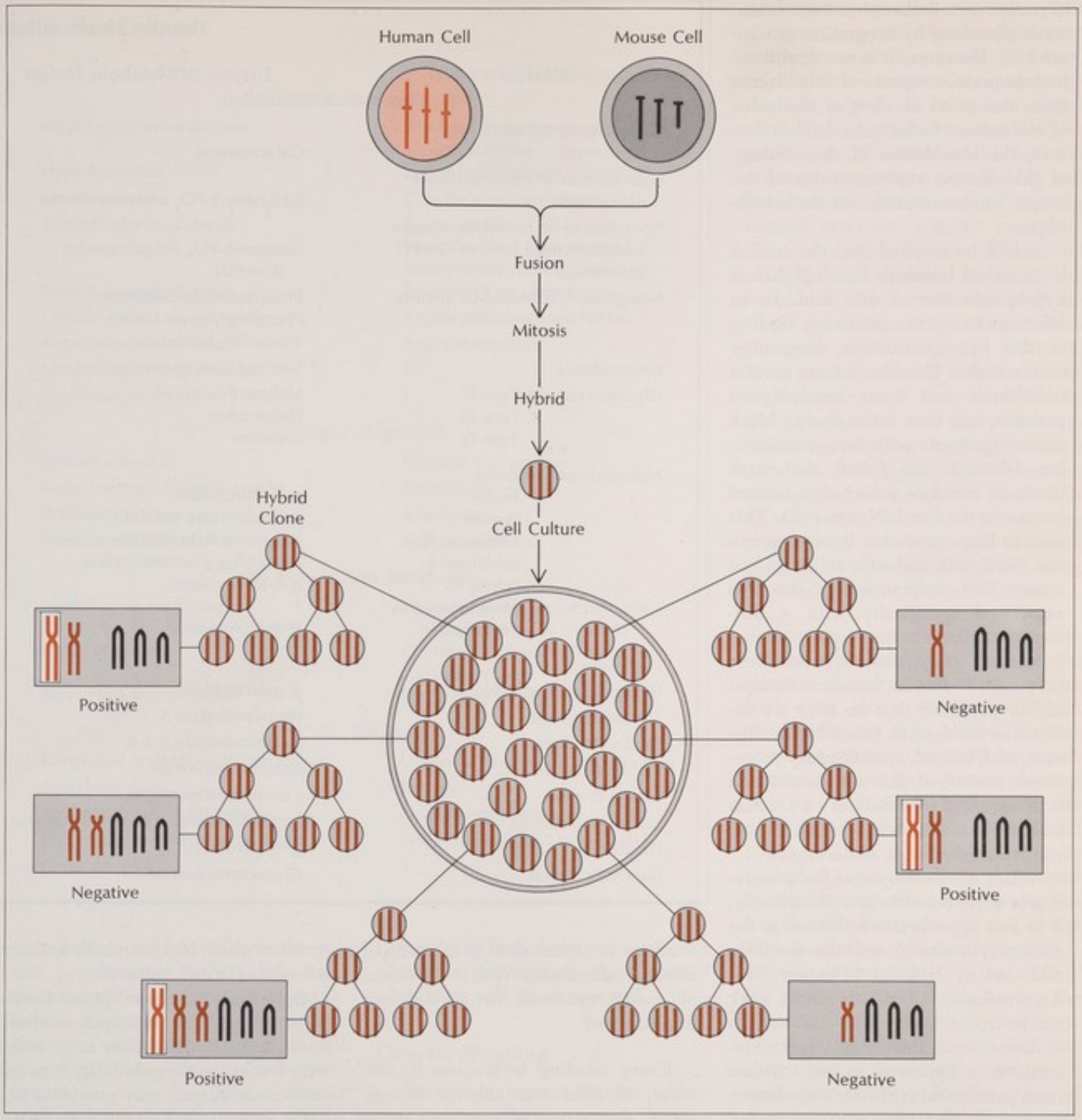
cytoplasmic organelles to be usefully active, and it is possible that this cannot be achieved.

Amino Acid Metabolic Defects

There are few inborn errors of amino acid metabolism in which the metabolic error is expressed in cell culture. One of these is citrullinemia. Why is citrullinemia so interesting as a genetic defect amenable to cell culture study?

A major portion of the answer to this question lies in the fact that many amino acids, but not citrulline, are essential to the growth of cells in culture. In fact, more amino acids are essential to cultured cells than to the intact individual. If essential amino acids are removed from the cultures, both normal and mutant cells will perish, and any possibility of the expression of impaired metabolism will be obviated. Citrulline, however, is not required; normally it is converted into arginine, a required amino acid. Experiments by Tedesco in my group at the University of Pennsylvania compared the effects in fibroblast cultures of the removal of arginine and subsequent replacement with citrulline in cells from a normal individual and those from a citrullinemic. It was found that arginine removal greatly reduced cell growth in both normal and mutant lines. But whereas the addition of citrulline normalized growth in the control line, it allowed the continued inhibition of growth in the citrullinemic cells. Correspondingly, ^{14}C-labeled citrulline was significantly taken up by protein from control cells but only minimally by citrullinemic cell protein. We also were able to measure the conversion of citrulline to arginine in the control cells and in those of the citrullinemic line and, somewhat paradoxically, we found no difference. However, when we varied the citrulline concentrations in the medium, the answer emerged. At high concentrations of citrulline, enzyme (argininosuccinate synthetase) activity was adequate to convert citrulline to arginine, but when the concentrations were brought into the physiologic range, the conversion was inhibited in the citrullinemic cells.

Here, then, was a good example of the use of cell culture not to identify the genetic etiology of citrullin-



A second approach to genetic mapping involves cloning of cultured hybrid cells and studying the karyotypes of surviving cells to determine which human chromosome persists in all cell populations that are positive for the phenotypic expression of the gene. If enough populations are studied, the particular chromosome (shown here with white background) can be identified.

emia – that had been established – but to pinpoint the specific phenotypic expression, in this case a low affinity for substrate by the enzyme which degrades citrulline; as a result there is an accumulation of citrulline in blood and tissues.

What about the disorders of amino acid metabolism that have not yet been amenable to study in cell cultures? The problem with PKU is typical. Among the amino acids essential to cells in culture but not to man is tyrosine. Tyrosine is normally metabolized from phenylalanine by phenylalanine hydroxylase, and all cell cultures lack this enzyme. Therefore, the defect in PKU cannot be expressed in culture. It should not be forgotten, however, that the genome required for the synthesis of phenylalanine hydroxylase is present in cultured normal cells, so we can expect development of techniques for production and detection of this enzyme in culture.

Nucleotide Metabolic Disorders

The metabolic defect underlying Lesch-Nyhan disease, a derangement

in purine metabolism, has been extensively discussed by Seegmiller (Chapter 10). However, it is worthwhile tc look at certain aspects of this disease from the point of view of the value of cell culture techniques. In that context, the elucidation of the etiology of this disease represents one of the major "success stories" for these techniques.

It will be recalled that the critical biochemical lesion in Lesch-Nyhan is hyperproduction of uric acid. In an effort to study the pathways leading to this hyperproduction, Seegmiller cultured skin fibroblasts from normal individuals and from Lesch-Nyhan patients, and then attempted to block purine synthesis with 6-mercaptopurine (6-MP). He found that such blockade could be achieved in normal but not in the Lesch-Nyhan cells. This was an important clue because previous work with leukemic cell lines had shown that resistance to 6-MP was associated specifically with a deficiency of the enzyme hypoxanthine-guanine phosphoribosyltransferase (HGPRT). It was therefore thought highly probable that an HGPRT defect was involved in Lesch-Nyhan disease, and, indeed, specific assay procedures confirmed this relationship. It is noteworthy that the HGPRT defect had not been suspected before the cell culture studies were undertaken.

While on the subject of the achievements made possible by cell culturing, it is also appropriate to return to the mucopolysaccharidoses, discussed in this text by Neufeld (Chapter 14). As previously noted, this line of work can be traced back to the observations of Danes and Bearn that metachromatic cells appeared in cell cultures from patients with Hurler's syndrome. Since these observations corresponded with the clinical findings of metachromatic cells in the skin of these patients, the investigators concluded that the cultured cells were expressing the basic disease defect, and that the cell culture system would provide an in vitro model for seeking to identify that defect, or, more precisely, for finding out what caused mucopolysaccharides to accumulate in the cells of affected individuals. Was it a problem of excess synthesis or of deficient degradation?

Human Biochemical Defects

Disorder	*Enzyme or Metabolic Defect*
A. *Carbohydrate Metabolism*	
Galactosemia (galactokinase deficiency)	Galactokinase
Galactosemia (classic transferase deficiency)	Galactose-1-PO_4 uridyltransferase
Nonspherocytic hemolytic anemia (Mediterranean form of G-6-PD deficiency)	Glucose-6-PO_4 dehydrogenase (G-6-PD)
Nonspherocytic hemolytic anemia	Phosphohexose isomerase Phosphoglycerate kinase Triose PO_4 isomerase
Ketoacidemia	Succinyl-CoA transferase
Glycogen storage, Type II	Maltase (lysosomal)
Type III	Debrancher
Type IV	Brancher
Mucopolysaccharidoses	
Hurler	α-L-iduronidase
Hunter	Sulfoiduronate sulfatase
Sanfilippo A	Heparan sulfate sulfatase
Sanfilippo B	N-acetyl-α-glucosaminidase
Scheie	α-L-iduronidase
Maroteaux-Lamy	?
"Atypical"	β-glucuronidase
I-cell disease	?
Generalized gangliosidosis	β-galactosidase
Tay-Sachs disease	Hexosaminidase A
Sandhoff's disease	Hexosaminidase A & B
Krabbe's disease	Galactocerebroside
Fabry's disease	Ceramidetrihexosidase
Lactosylceramidosis	Lactosylceramide galactosylhydrolase
Fucosidosis	α-1-fucosidase
Gaucher's disease	Glucocerebrosidase

Using labeling techniques in cultures, Neufeld was able to demonstrate that rates of synthesis were normal but degradation was impaired. Still employing cell cultures, Neufeld then mixed cell lines from patients with different but related diseases together, and with appropriate combinations she found that these cells elaborated factors that were mutually corrective. She then was able to isolate and enzymatically characterize the proteins which had corrective activity. All of this in cell culture!

Before leaving the subject of Lesch-Nyhan and the mucopolysaccharidoses, a brief discussion is in order with respect to the possibility that cell culture techniques may provide a valuable means of studying cell-cell interactions. This work is in its relatively early stages but some of the observations are interesting. For example, if fibroblast cultures are established from women heterozygous for Lesch-Nyhan disease, one has two cell populations in the cultures; in accordance with the Lyon hypothesis, half of the cells express the X-linked de-

Expressed in Skin Culture

Disorder	*Enzyme or Metabolic Defect*
B. *Amino Acid Metabolism*	
Maple syrup urine disease	Branched-chain amino acid decarboxylase
Hyperlysinemia	Lysine-ketoglutarate reductase
Homocystinuria	Cystathionine synthetase
Methylmalonic acidemia	B_{12} coenzyme Methylmalonyl CoA mutase Methylmalonyl CoA racemase
Ketotic hyperglycinemia	Propionyl CoA carboxylase
Citrullinemia	Argininosuccinate synthetase
Argininosuccinic aciduria	Argininosuccinase
Nonketotic hyperglycinemia	?
Cystinosis	?
C. *Lipid Metabolism*	
Refsum's disease	Phytanic acid oxidase
Metachromatic leukodystrophy	Arylsulfatase A
Wolman's disease	Acid lipase
Niemann-Pick disease	Sphingomyelinase
D. *Nucleic Acid Metabolism*	
Lesch-Nyhan syndrome	Hypoxanthine-guanine phosphoribosyltransferase (HGPRT)
Oroticaciduria	Orotidylic pyrophosphorylase and decarboxylase
Xeroderma pigmentosum	DNA repair
E. *Others*	
Congenital erythropoietic porphyria	Uroporphyrinogen III cosynthetase
Acatalasia	Catalase
Acid phosphatase deficiency	Acid phosphatase
Hydroxy-lysine deficient collagen	Lysyl-protocollagen hydroxylase
Cystic fibrosis	?

fect, half do not. As long as the culture population is such that the cells do not come into contact, these populations maintain their individuality; however, when growth produces cell crowding that causes the cell membranes to come into direct contact with each other, then the cell population becomes homogeneous and the mutant cell no longer expresses the defect. Presumably, a diffusion of enzyme has occurred between the cells.

In contrast, cells from patients with mucopolysaccharide disease do not need such direct contact for the corrective factors to enter into the defective cells. In fact, if abnormal cells are placed in the medium from normal cells, they will be normalized.

Genetic Mapping

Once it is known that an inborn error of metabolism results from a specific mutant gene, a natural question is: "On which chromosome is the gene located?" Cell cultures have demonstrated their ability to answer this question in a number of cases through hybridization of human cells and cells from the mouse, hamster, or rat. All things being equal, such hybridization generally results in the preferential survival of the rodent chromosomes and the disappearance of the human chromosomes. However, in cell culture one can manipulate the medium to provide for selective survival of the human chromosomes. As an example, if a mouse cell is selected for an absence of HGPRT activity and fused with a human cell possessing normal HGPRT activity in a medium designed to permit survival of only cells having HGPRT, the fused cell will survive only if it retains the human chromosome with the gene that codes for HGPRT. Since most of the other human chromosomes will have been lost in the hybrid, one can apply chromosomal karyotyping techniques in a number of hybrids and determine which chromosome always survives in viable cells. By this method the chromosome carrying the HGPRT gene can be identified (in this case, of course, it is the X chromosome).

Another approach to mapping with cell culture techniques involves the cloning of hybridized cells; hybridized cells clone readily and such clones will lose their human chromosomes at random, but will keep perhaps one, two, or three human chromosomes in each clone. If the gene one is trying to locate is that which codes for, say, lactic dehydrogenase (LDH-A), one then tests each clone for the human LDH-A expression (distinguishable from the LDH-A of the animal species by electrophoresis). Once again the problem is that of finding enough hybrid clones having the human enzyme activity. By karyotyping such hybrids one can find the chromosome persistently associated with the activity. In the case of LDH-A, it proved to be chromosome 11.

In summary, then, human cell culture systems have already proved their value in the study and definition of many inborn errors of metabolism. As the techniques improve and the cellular resources are expanded, this value can be expected to increase exponentially.

17

Genetic Control of the Antibody Response

Section Four

Immunogenetics

17

Genetic Control of the Antibody Response

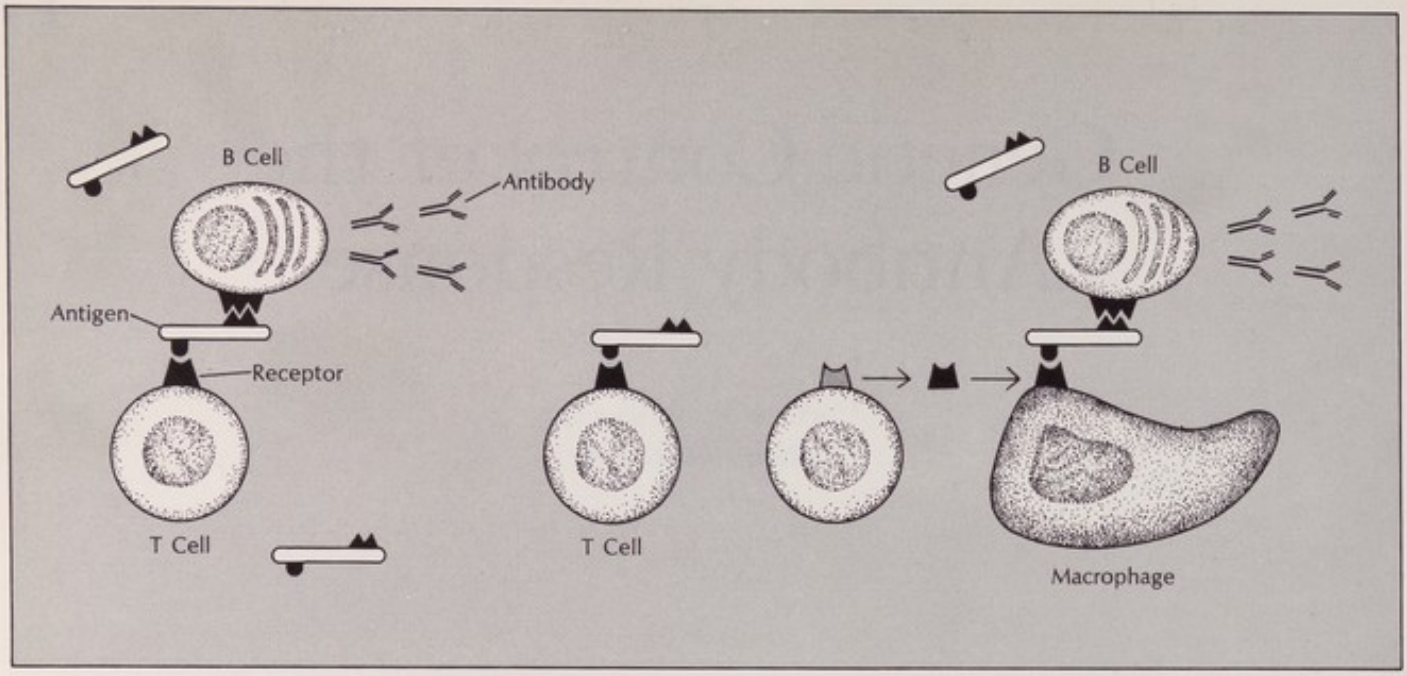
HUGH O. McDEVITT
Stanford University

Any analytic dissection of the humoral antibody response must take into account at least two, and probably three, cell populations: the B cells, derived from the bone marrow, the T cells of thymus origin, and the macrophages. The B cell manufactures antibody after it has come into contact with the haptenic determinant on the antigen. It is now well established that this event does not occur in isolation but is dependent on some kind of signal emitted only if there is simultaneous recognition of a carrier determinant on the antigen by the T cell or a T cell product. While the role of the macrophage has not been clearly defined, it has been suggested that it acts by acquiring a surface receptor from the T cell so that it can bind antigen for presentation to the B cell. Abundant evidence exists in some systems that the macrophage is an essential participant in the chain of events leading to an antibody response.

Clearly, there is considerable complexity in the cellular events that facilitate an antibody response to an antigen. Moreover, a tremendous amount of both diversity and specificity is involved in the immune response. Both the B and the T cells must be precommitted, in the sense that they must exist prior to the organism's encounter with the antigen and must have been genetically programmed to react with specific antigens that stimulate them to proliferate. Add to this the fact that every antigen has a large number of antigenic determinants and one is confronted with a system of impressive sophistication. It would be reasonable to assume that identifying the genetic controls involved for this system in the immunologically intact animal would be a Herculean task and that any geneticist attempting it would have very little expectation of success. With a very large number of functions requiring specific controls, one might well assume that a very large number of genes would be needed to do the controlling. To find an animal deficient in a particular immune response clearly might be an extremely difficult problem.

And so it proved. For many years this difficulty was not surmounted. There were a few observations in the literature, largely ignored, suggesting the existence of animal responders and nonresponders to particular antigens. However, such experiments were generally clouded by the imprecision of the assay systems employed. It was not until Baruj Benacerraf took note of an observation of Paul Maurer's that the logjam was broken. The latter had reported that when the very simplest of polypeptides, containing only two or three amino acids, were employed as antigens, immune response could be elicited from only about 40% of tested animals. Picking up this observation, Benacerraf put a hapten on poly-L-lysine (PLL) and injected the complex into outbred guinea pigs. He found that some animals were indeed responders, others absolute nonresponders. He then performed cross-breeding experiments and found that if he bred two responders, about 80% or 90% of the offspring were responders; with two nonresponders, the offspring were 100% nonresponsive. These results suggested that response was dependent on a single gene inherited as an autosomal dominant. Confirmation was obtained by experiments with inbred strains (2 and 13), in which it was seen that one strain (2) showed a 100% response to hapten-PLL conjugates and another (13) a 100% nonresponse.

Equally significant was Benacerraf's finding that the response to the poly-L-lysine was independent of the hapten conjugated with it. In other words, whether he used dinitrophenyl (DNP), benzylpenicilloyl, or picryl as the hapten, the pattern of response and nonresponse was identical so long as poly-L-lysine was the carrier; it was the ability of the animal to recognize the carrier that was the



Antibody response appears dependent on simultaneous recognition of carrier determinant on antigen by T cell and of hapten determinant by B cell. Signal may pass directly from T to B (left), or receptor transfer, possibly in form of T receptor, may go from T cell to macrophage; signal is then transmitted to B cell, which is "turned on" to elaborate specific antibody (right).

determining factor in responsiveness.

To establish the immunologic basis for this observation more firmly, Benacerraf turned to a molecule that had been unequivocally identified as a foreign protein antigen, bovine serum albumin (BSA). After complexing the DNP poly-L-lysine with acetylated BSA, he immunized guinea pigs that were completely unresponsive to the hapten poly-L-lysine antigens. Now these animals had clear-cut antibody responses to the immunogenic complexes; they produced anti-DNP antibody. He then made these animals tolerant to BSA and by doing so succeeded in abolishing the anti-DNP response. Here, then, was incontrovertible evidence not only that the genetic control was directed at the carrier but that the carrier had to be recognized by the host for the response to be evoked. This was one of the first suggestions that the T cell might be the primary site of the defect in the nonresponder. This could only be placed in context after Mitchell and Miller documented the function of the T cell in recognition of the carrier.

Studies of the antibody produced by the responder immunized with DNP-PLL alone and of the nonresponder immunized with DNP-PLL on the foreign carrier (BSA) strengthened this point still more. In terms of affinity for DNP, the antibodies appeared to be identical. Apparently both responders and nonresponders had the ability to make the same antibody. The defect in the nonresponder was at the T cell antigen recognition level.

These experiments created something of a paradox for the immunologic community, which had, in general, been proceeding on the hypothesis that the host recognized foreignness through the agency of antibody, and that, in the context established by Burnet, the clones of cells precommitted to recognizing a particular antigen did so via specific antibody either on their surface or in some latent or precursor form within their synthetic mechanism. If both nonresponder and responder animals have the same genes for antibody structure, but a nonresponder fails to recognize the antigen and produce specific antibody, it implies mediation by a different recognition system.

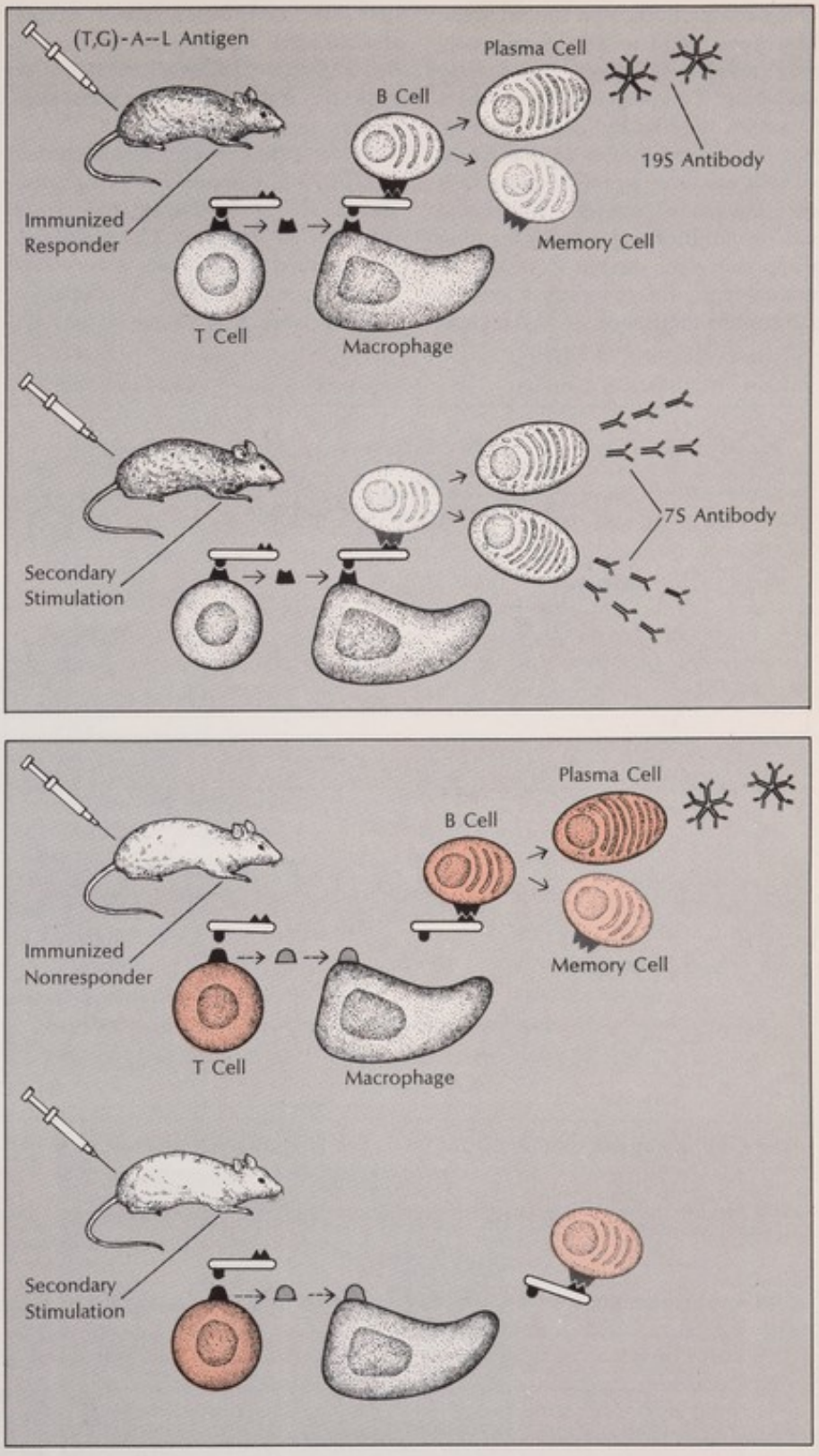
Benacerraf, whose contributions to the field cannot be overestimated, performed a number of other experiments that served to confirm the T cell as at least one site of expression of this type of genetic control. In summary, these experiments demonstrated that responder and nonresponder animals could be differentiated in terms of the capacity to manifest reactions thought to be pure T cell functions. Thus, nonresponder animals immunized with DNP-PLL on a carrier failed to exhibit delayed hypersensitivity to the hapten conjugate although they made antibody to DNP. However, the responder immunized with DNP-PLL alone had completely normal delayed hypersensitivity to the hapten conjugate. Another parameter of T cell function absent in the nonresponder and normal in the responder proved to be antigen-induced blast transformation. The evidence clearly indicated the genetic defect in the nonresponder affects T cell function.

At the outset of this discussion, it was noted that in the light of the complexity of cellular interactions involved in immune responses, one would not have expected to be able to identify responders and nonresponders to specific antigens. Yet such patterns were identified in the work of Benacerraf, in work in which I participated directly, and in experiments in several other laboratories. In the past few years, at least 30 examples of this type of genetic control have been reported, almost all arising out of chance observations. These apparently fortuitous findings shared a common element; in each instance, the antigen employed was either a syn-

thetic polypeptide having only a small number of antigenic determinants, or an isoantigen (analogous to the blood group antigens in which there is only a single foreign determinant), or a complex protein antigen presented to the immunized animal in an extremely low dose. The importance of dose was demonstrated with bovine gamma globulin, against which nonresponder animals did react when the dose was made large enough. While the significance of dose in detecting genetic differences in response has been shown, the explanation for the differential response has not yet really been proved experimentally. However, the explanation that seems entirely consistent with our other knowledge of the behavior of immune response genes is that even when an antigen has multiple determinants, the thresholds of response to these different determinants are different. With low doses, the threshold for only one determinant may be exceeded; if the animal is not genetically capable of responding to this determinant, it will fail to respond to the entire complex. As the dose is increased, the response threshold for other determinants may be reached and now a response will be elicited.

My own interest in the genetics of immune responses was first stirred about 10 years ago when I was working in John Humphrey's laboratory at Mill Hill in London. The work there was addressed toward obtaining experimental evidence about the mechanisms of antibody formation. We were employing a synthetic compound called (T,G)-A--L. The backbone of this compound is poly-L-lysine; each lysine residue has outwardly extending side chains consisting of nonantigenic poly-D-L-alanine. When tyrosine and glutamic acid [(T,G)] are added to the alanine side chains, the molecule [(T,G)-A--L] becomes immunogenic. When we first immunized rabbits with (T,G)-A--L, we got very poor responses. This was surprising, since Michael Sela in Israel, who had synthesized this compound, as well as many other polypeptide antigens, had reported excellent antibody responses.

The one explanation offering any hope of resolving this contradiction was that perhaps we were using the wrong rabbits. For this reason, the experiments were repeated using dif-

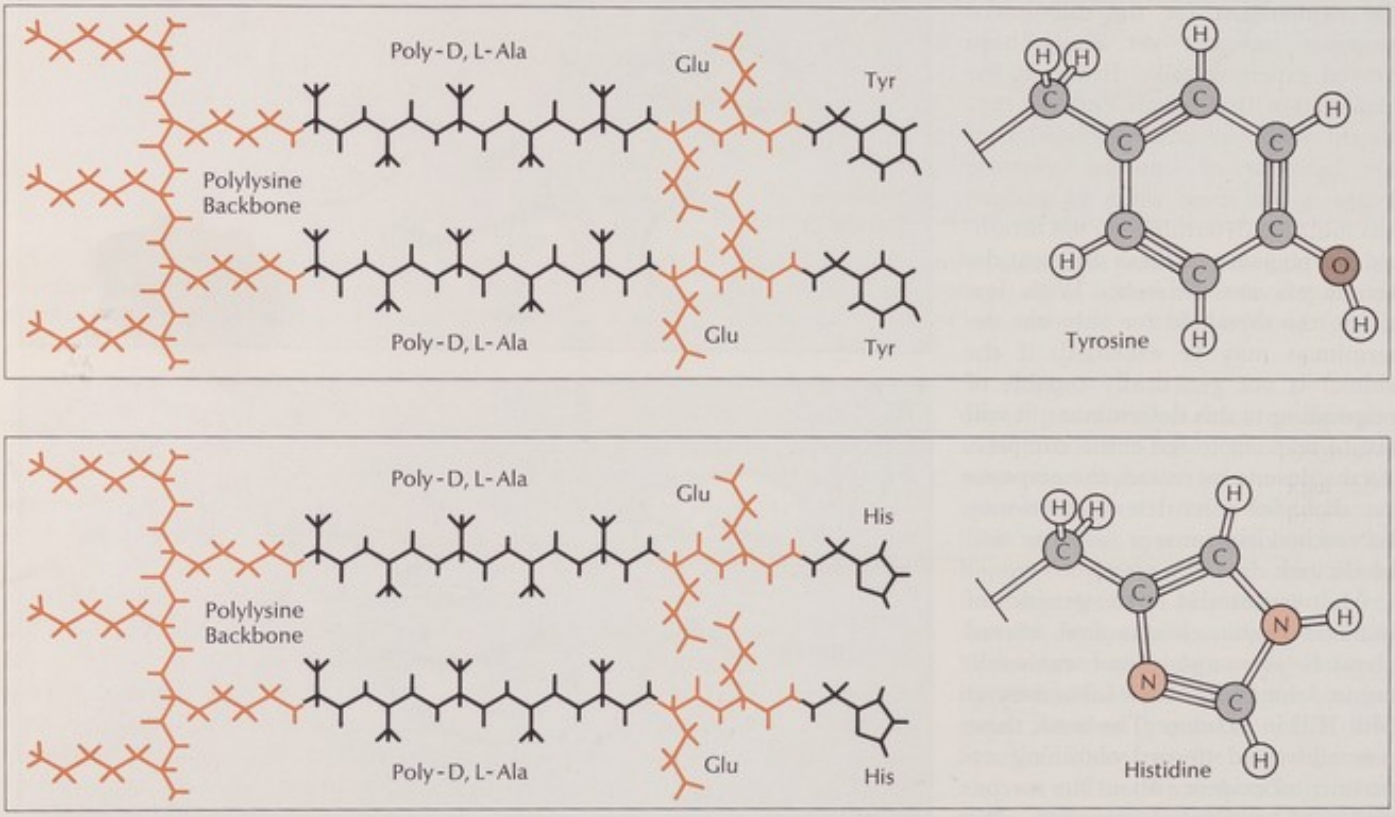


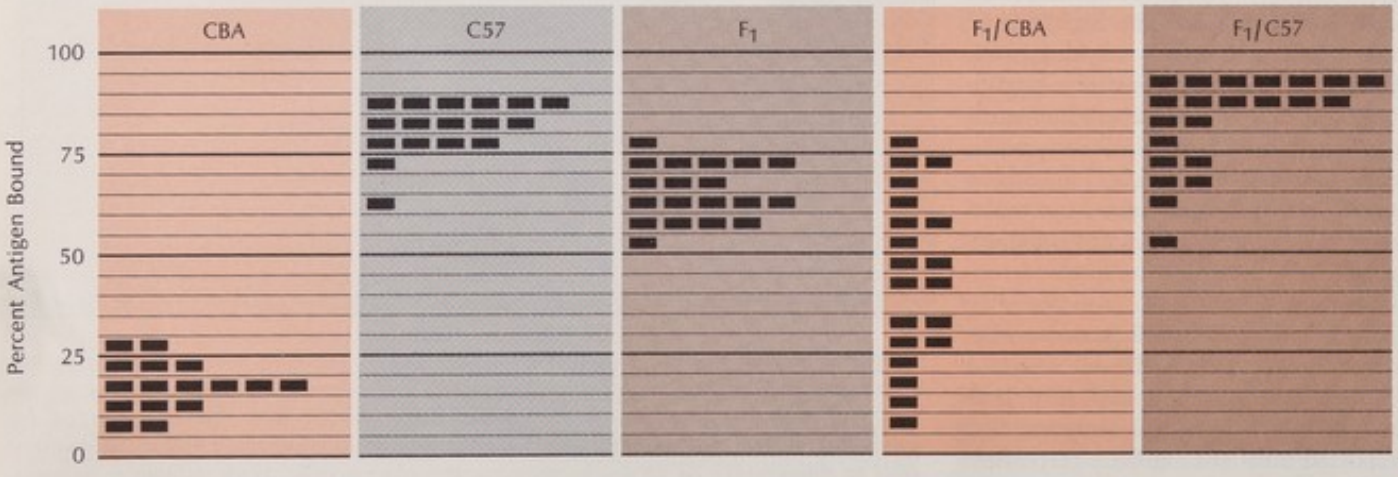
Diagrams, employing the T cell–macrophage–B cell receptor transfer model (facing page), suggest differences between responder and nonresponder animals are qualitative as well as quantitative. Responders (top panels) first react to antigen by production of 19S antibody and memory cells. Second immunization will cause memory cells to differentiate to 7S-antibody-producing plasma cells. With responder animals, 19S response is intact but there is no "switchover" to 7S. The implication is that the 19S response is partially thymus-independent, but that for 7S, the helper role of the T cell must be fulfilled. It should be noted that in this scheme, the 7S memory cells are depicted as descendants of 19S precursors. However, one cannot exclude the possibility that the two cell lines are independent and that the 19S cells mature first.

ferent rabbit lines, and indeed some lines were found to give consistently good responses, others consistently poor ones. This suggested that these responses might be under genetic control. Genetic studies were attempted with the rabbits, but because they were not inbred we were confronted with a multiplicity of segregating genes and were unable to relate the immunologic differences to a specific pattern of inheritance. It was logical, therefore, to turn to inbred animal strains; mice were chosen because of the availability of a large number of such strains developed over many decades by mouse geneticists.

In our initial experiments we tested (T,G)-A--L responses in CBA mice and in F_1 generation animals from CBA x C57 matings. The CBA animals turned out to have a very low antibody response, the F_1 animals a high response. C57's were tested next and they were very high responders to (T,G)-A--L. This suggested that a single gene governed response quantitatively and that the gene for high response was dominant; the CBA x C57 (F_1) animals showed responses much closer to the C57 than to the CBA. In this respect the mice differed from the Benacerraf guinea pigs, in which response was 100% or zero.

At this time we were aware of Benacerraf's observations with PLL and





Structures of "simple" synthetic antigens, (T,G)-A--L and (H,G)-A--L are diagrammed. Both have poly-L-lysine backbones with side chains of nonantigenic poly-D,L-alanine. Antigenicity depends on addition of either tyrosine and glutamic acid (T,G) or histidine and glutamic acid (H,G). Graphs show genetically determined difference in immune responses of two inbred strains to (T,G)-A--L: C57 mice are high responders, CBA low. Breeding studies show that the gene for high response expresses dominance. Although the data for (H,G)-A--L response are not given here, the situation was reversed, with CBA high and C57 low.

wondered whether we were dealing with some enzymatic peculiarity related to the handling of lysine, alanine, or both, rather than with some phenomenon of basic importance to the genetic control of antibody production. Fortunately, Sela was able to provide the tools with which to answer the question, since he had also synthesized two other compounds—(H,G)-A--L, in which histidine is substituted for tyrosine, and (Phe, G)-A--L, in which phenylalanine replaces the tyrosine. We repeated the experiments with (H,G)-A--L and found a response pattern that was exactly opposite. Now the CBA's were the high responders and the C57's the low. Once again, high response was dominant in the F_1 generation.

Although this was clear evidence that what was being expressed was genetic control and that the gene or genes involved were dominant for responsiveness, the evidence needed to answer three basic questions was either lacking or indirect. First, how many genes were involved? Second, were the controls involved related to a basic aspect of antibody structure, manufacture, or elaboration? Third, if, as was indirectly implied, the controls were expressed in the T cells, could this be unequivocally shown, and would this demonstration exclude concomitant expression of the same genes in the B cells?

We were able to attack the first question by employing conventional genetic maneuvers involving F_2 crosses and then subjecting the response patterns to mathematical analysis. The results clearly pointed either to a single gene, or possibly a major gene with modifiers, or, at most, to two genes. Later experiments, cited below, proved a single gene was involved. The second question was more difficult to approach, and the efforts to find an answer involved the testing of various possible relationships between the immune response genes and genes that might control other parameters of antibody structure and function.

Our first working hypothesis was of a relationship between the response gene(s) and those known to be involved in the determination of antibody structure. It should be explained that a number of workers in the 1960's had shown that antibody structure was controlled by genes coding for the heavy and light chains that form the immunoglobulin combining sites. Genes for the constant (C) portion of light chains and for the constant and variable (V) portions of the heavy chains had been defined first in the rabbit and were called immunoglobulin allotype genes. These genes constitute identifiable linkage groups. (Linkage groups will be discussed in detail later.) Similar genes for heavy chain constant regions have been identified for both mouse and man.

To test the possibility of association between the immune response gene(s) and antibody structure, we did crosses between animals that were high responders of one immunoglobulin allotype (A) and low responders of another allotype (B). Theoretically, if the two genes were associated, then A A allotype offspring would continue to express high responsiveness to the test antigen, while low

Histocompatibility-Linked Specific Immune Response Genes

	Antigens		Linkage
GUINEA PIG	PLL	(Probably same gene)	Strain 2, H specificity
	PLA		" " " "
	GL		" " " "
	DNP-PLL		" " " "
	GA		" " " "
	GT		Strain 13, H specificity
	BSA (low dose)	(Probably same gene)	Strain 2, H specificity
	DNP-BSA (low dose)		" " " "
	HSA (low dose)		" " " "
MOUSE	(T,G)-A--L		$H\text{-}2^{b,i}$
	(H,G)-A--L		$H\text{-}2^{a,k,h}$
	(Phe,G)-A--L		$H\text{-}2^{a,b,d,i,k,q}$
	GAT_{10}		$H\text{-}2^{a,b,d,k}$
	GAL_{10}		$H\text{-}2^{a,b,d,k,s}$
	GLA_5		H-2
	GLT		"
	(T,G)-Pro--L		Not H-2 linked
	Ovomucoid (low dose)		$H\text{-}2^{a,k}$
	Ovalbumin (low dose)		$H\text{-}2^{b,d,q}$
	Bovine gamma globulin (low dose)		$H\text{-}2^{a,k}$
	Tri-nitro-phenyl-hapten		$H\text{-}2^{b}$
	Mouse erythrocyte antigen Ea-$1^{a,b}$		H-3 or H-6
	Mouse male (Y) transplantation antigen		$H\text{-}2^{b,i}$
	H-2.2 specificity		Not known
	Mouse Ig A myeloma		H-2
	H-13 specificity		$H\text{-}3^{a}$
RAT	Porcine lactic dehydrogenase		Not known
	GLT		" "

Above is a partial list of the immune response (Ir) genes that have been identified thus far. With a few exceptions, when linkage has been established, it is to the histocompatibility system, and many Ir genes have been linked to specific H types.

responders would continue to be concentrated among animals of B B allotype. Our results were negative in that the offspring included a randomly distributed number of low responders who were A A and of high responders who were B B. In other words, no genetic association between immune response patterns and immunoglobulin allotype could be discerned.

With one possibility eliminated, we next sought evidence for an association with the process of antibody formation. Since immune response genes are dominant, we were able to transfer spleen cells from a responder × nonresponder (F_1) into an irradiated recipient of the nonresponder parental strain. In this situation, the F_1 spleen cells will be histocompatible with the parental recipient, and the spleen cell graft will not be rejected by the irradiated host. The chimeric animals thus created were high responders to (T,G)-A--L. This experiment is the primary evidence that immune response genes are expressed in immunocompetent cells, such as spleen cells. They are therefore directly implicated in controlling some aspect of the immune response. In further experiments, we tried to take into account both the evidence that this process involved several major cell types and the work in both Benacerraf's laboratory and our own, which pointed to the T cells as the expressors of the immune response genes.

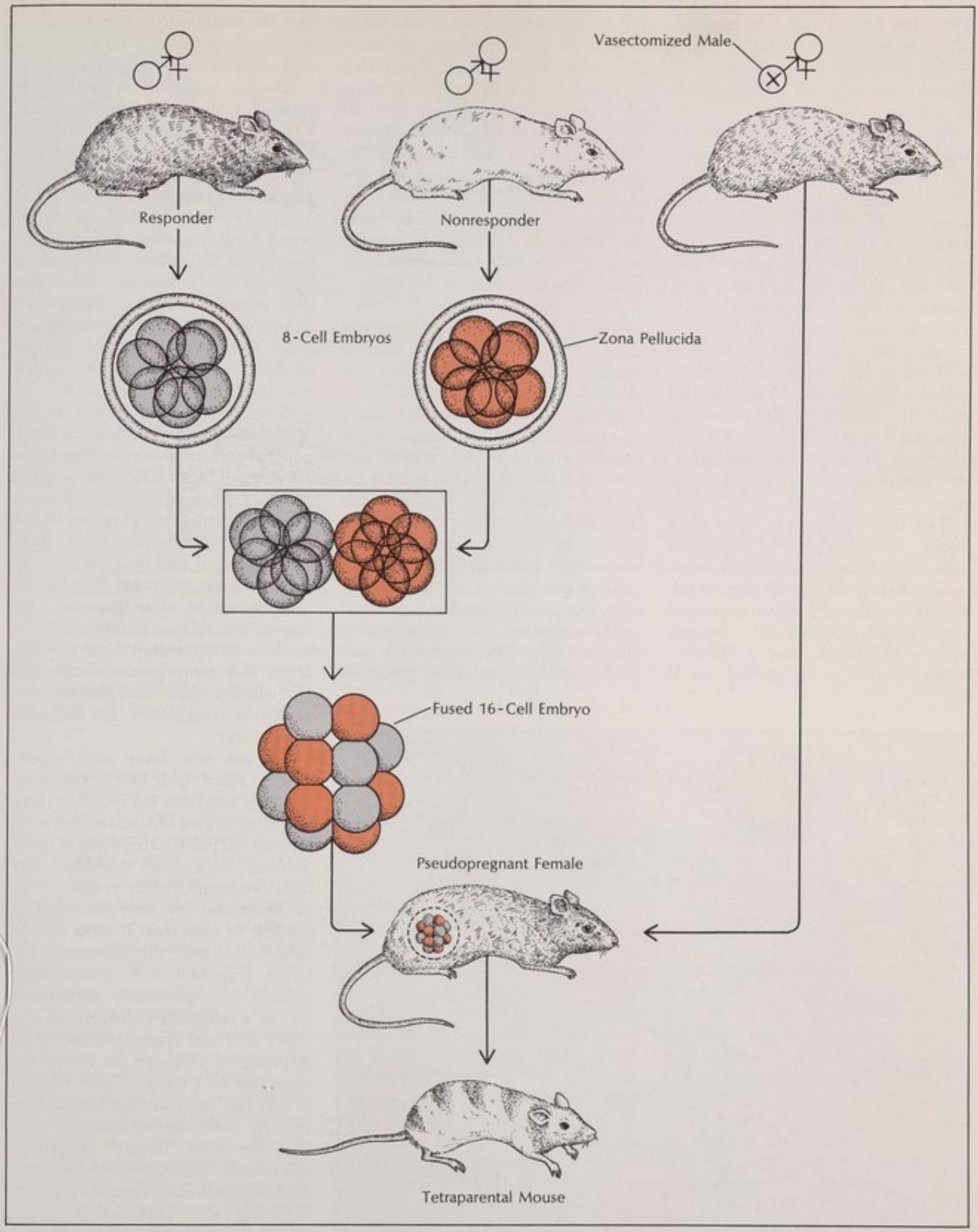
Also relevant to the experiments designed to pin down the relationship between the immune response genes and T cell behavior were findings by Grumet in the laboratory here at Stanford. He had shown that when responder and nonresponder animals were stimulated to produce anti-(T,G)-A--L antibody by single injections of (T,G)-A--L in aqueous solution, both produced an equal response in the form of 19S (IgM) antibody exclusively. But when a second antigenic stimulus was applied, the responder animals had a huge anamnestic response that was characterized by a prompt switchover from 19S to 7S (IgG) antibody. The nonresponder animals had no such secondary response. This strongly suggested that one function of the T cell controlled by the immune response gene was a switchover from 19S to 7S immunoglobulin, a suggestion that was strengthened when we irradiated and thymectomized the responder animals, thereby removing all their T cells. After replacing their bone marrow, we then challanged them with antigen. Their initial 19S response was unaffected, but secondary stimulation failed to produce the conversion to 7S. In other words, when thymectomized, the responders acted exactly as if they were genetically nonresponsive.

While these and other experiments gave ample proof that the immune response genes had major expression in T cells, we still could not exclude a concomitant expression in B cells. Indeed, Shearer and Mozes have presented evidence that suggests either expression of these genes in T and B cells or, if expressed only in T cells, they have a marked effect on T cell–B cell interaction. Our first attempt to find out about this involved the use of cell transfer systems. The working hypothesis was that if you mixed responder T cells with nonresponder B cells, then, if responsiveness was entirely a T cell function, the T cells would turn on the B cells to make antibody. The antibody would be recognizable as being of nonresponder origin by virtue of immunoglobulin allotype markers (i.e., of allotype B in the example cited above).

These experiments were tried in a number of different ways without success. The strains identified as responder and nonresponder all differed in their major transplantation antigen (H-2) type. All our attempts to develop chimeras with immunocompetent cells of both responder and nonresponder strains were frustrated by severe graft rejection reactions. Indeed, it was our attempt to identify strains that were responders and nonresponders of the same H-2 type that led to the discovery that the immune response genes were on the same chromosome with the genes for the major transplantation antigens, a finding of fundamental importance, to be discussed in detail later. Because of this association, attempts to mix responder and nonresponder animals were confronted with an obligate histocompatibility barrier. All of the animals succumbed to either graft-vs-host or host-vs-graft reactions. We then tried to create radiation chimeras by first lethally irradiating the animals and then grafting them with bone marrow of the opposite strain. In theory, this should produce animals with both responder and nonresponder T and B cells. These would be differentiable by their allotypic markers, and one therefore could study the cellular interactions and their immunoglobulin products. Conceptually this was fine but, as it turned out, the animals tended to "select" either responder or nonresponder cells on a mutually exclusive basis.

We needed a better way to get chimeras that were stable; if they could not be 50-50 responder and nonresponder, at least we required reasonably good mixtures. The solution was the tetraparental mouse system developed by Tarkowski in Europe and Mintz in Philadelphia. The specific details of the method employed to produce these remarkably useful experimental animals are diagrammed (facing page). Briefly, one starts with two mouse strains (for these purposes, responder and nonresponder) and many young females of each strain in estrus at the same time. The females are first given hormonal treatment to induce the production of a large number of ova (superovulation). They are then mated with normal males. At the appropriate time, the uterus is removed and the embryos flushed out of the oviduct with an injection of tissue culture medium. The protective zona pellucida is then digested off the eight-cell-stage embryos and one embryo from a responder is placed together with one from a nonresponder in a droplet of tissue culture medium under oil. After 18 hours in culture, many of the embryos fuse to form a larger embryo. Many of these go on to form blastocysts. These fused blastocysts are then transplanted into the uterus of recipient females that had been made pseudopregnant by a previous mating with vasectomized males. Approximately 18 days later, mice are born that carry cells from the two strains. These mice are called tetraparental because they carry four copies of each chromosome of the normal mouse karyotype.

In appropriate strains, these animals show all sorts of bizarre mosaicism, with coat colors that vary from all black to all white to patches or stripes of each. In our experiments, we used animals that did not differ in



Drawings show steps in the making of a tetraparental mouse. At the same time as animals of responder and nonresponder strains are mated, a third female is mated with a vasectomized male. Eight-cell embryos are removed from the two pregnant mice and placed together. Zonae pellucidae are digested off and the two embryos are nudged together in culture. If they combine into a single 16-cell embryo, this is then implanted into the pseudopregnant "foster mother," which has undergone the necessary hormonal changes that facilitate implantation. Here the implanted embryo can subsequently develop into a tetraparental mouse.

coat color and, in fact, were genetically identical except for chromosomes carrying H-2 and the immune response genes and the immunoglobulin allotype genes.

One of the most fortunate attributes of the tetraparentals, from our point of view, is that they have a remarkable consistency in their hematopoietic and immunocompetent cell populations. Cell distribution throughout the lymphoid-reticuloendothelial system is reflected in any part of that system, so that if there is a 60:40 responder : nonresponder distribution in the marrow, the same distribution is found in the spleen, the thymus, and the lymph nodes. Furthermore, these distributions appear to remain stable for long periods of time. Finally, cells from the two strains operationally tolerate each other immunologically.

This then appeared to be an ideal system for mixing nonresponder B cells and responder T cells in order to see whether the T cells turn on the B cells to make antibodies; or, to put it another way, to see whether one can pinpoint the genetic defect in the nonresponder animals at the level of a T cell recognition function. If this could be done, it would show that the nonresponder B cell is functional and capable of making antibody when given an opportunity to interact with the right T cell. In the tetraparental mice "made" in our laboratory by Bechtol and typed by Freed in terms of the origins of the specific anti–(T,G)-A--L immunoglobulins that they produced, we found a substantial number of animals with lymphocyte populations derived from both strains.

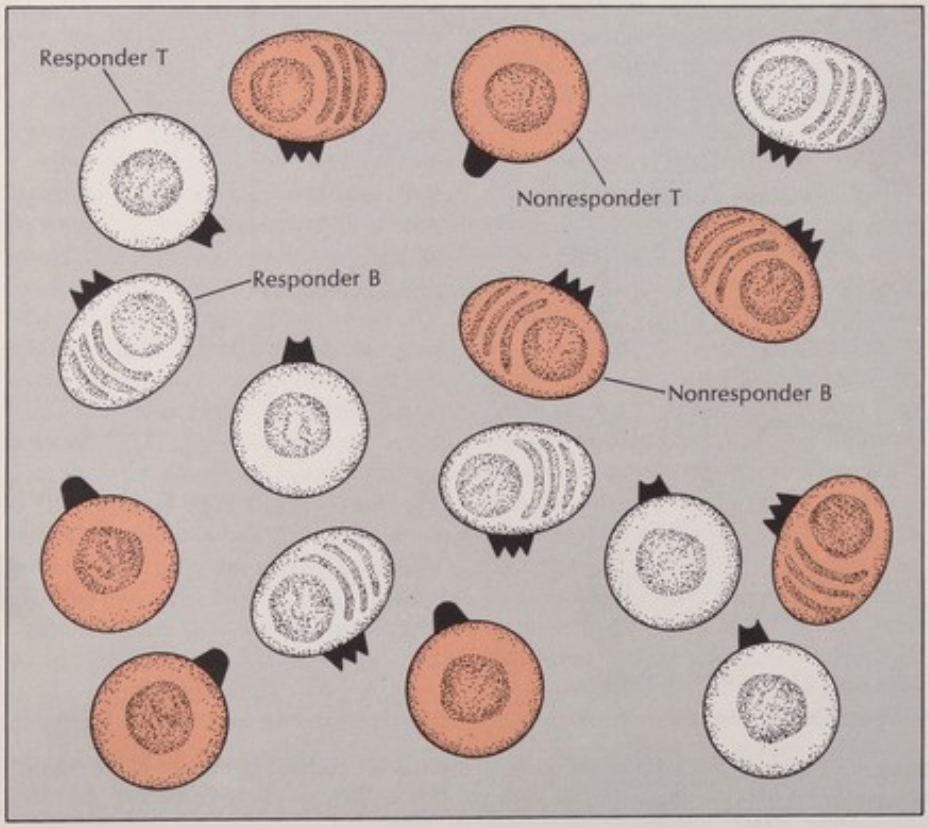
When we analyzed the anti–(T,G)-A--L antibody from these animals, it included large amounts of immunoglobulin produced by the nonresponder B cells. While the 7S antibody from the nonresponder strain was sometimes less than that from the responder cells, it still was readily detectable and, of course, greatly increased over the levels that one finds in the normal, intact nonresponder animal. In a strict sense, the result shows that something from the responder can turn on nonresponder B cells. Our assumption that it is the T cell is reasonable, since this is the only other cell type with immunologic specificity. However, before we can claim that the immune response genes are expressed entirely in T cells, two possible objections must be overcome.

The first is that within the chimeric system of the tetraparental mouse, information transfer might be taking place; that is, the nonresponder could be transferring its immunoglobulin genes to the responder, or some other gene's expression could be transferred from responder cells to nonresponder B cells. However, work at Stanford by Herzenberg and Hattis with tetraparental mice (being used in another context) has produced evidence against this possibility, and one would have to classify the information transfer hypothesis as unlikely.

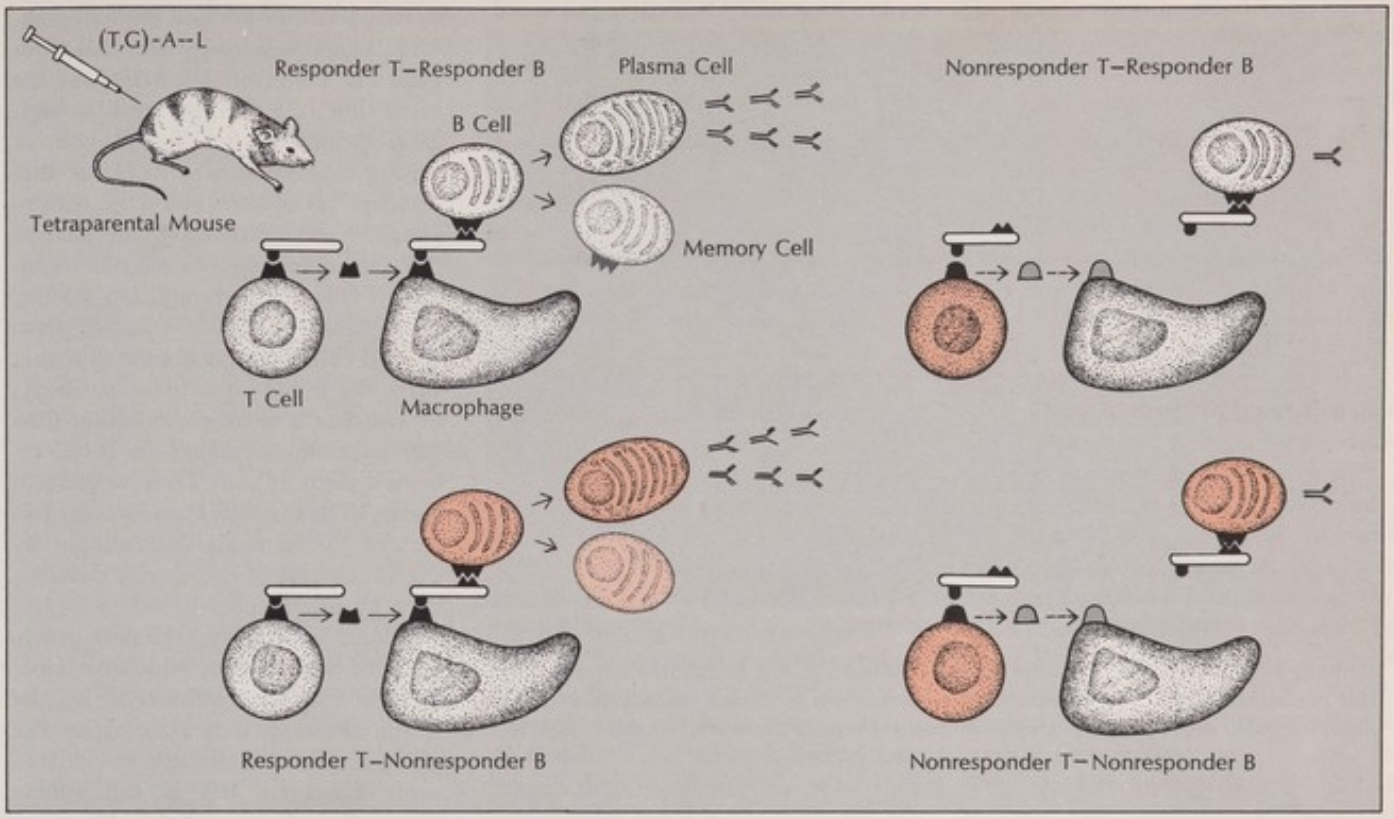
A more serious objection is the possibility that the T cells in the tetraparental animals are responding to the histocompatibility differences between the two cell populations in such a way that they emit signals that turn on the B cells to antibody production. Such nonspecific turning on of B cells by nonspecifically activated T cells has been shown in other systems. One way to test this possibility would be to make tetraparentals from two different H-2 nonresponder strains and see whether these animals are converted to responders in the nonspecific way.

We are now doing such experiments. From both the preliminary evidence and from the fact that there is a great degree of tolerance between the two cell lines contributing to tetraparental mice – and therefore very little histocompatibility reaction – we are reasonably sure that this objection too will be answered. If so, it will be established that the difference between responder and nonresponder animals is a genetically determined lack of a recognition unit, almost certainly of T cell origin, capable of responding to whatever the carrier determinant on (T,G)-A--L or PLL is. It will also be clear that both the responder and nonresponder B cells have the same "library" of genetic information and therefore the same antibody-producing capability.

Perhaps most important is the implication that we are dealing with two distinct systems of antigen recognition. Antibodies recognize antigens by virtue of having heavy and light chains through which they develop combining sites for particular anti-



"Ideal" chimera in tetraparental mouse will include approximately equal distribution of T and B cells from both responder and nonresponder strains.



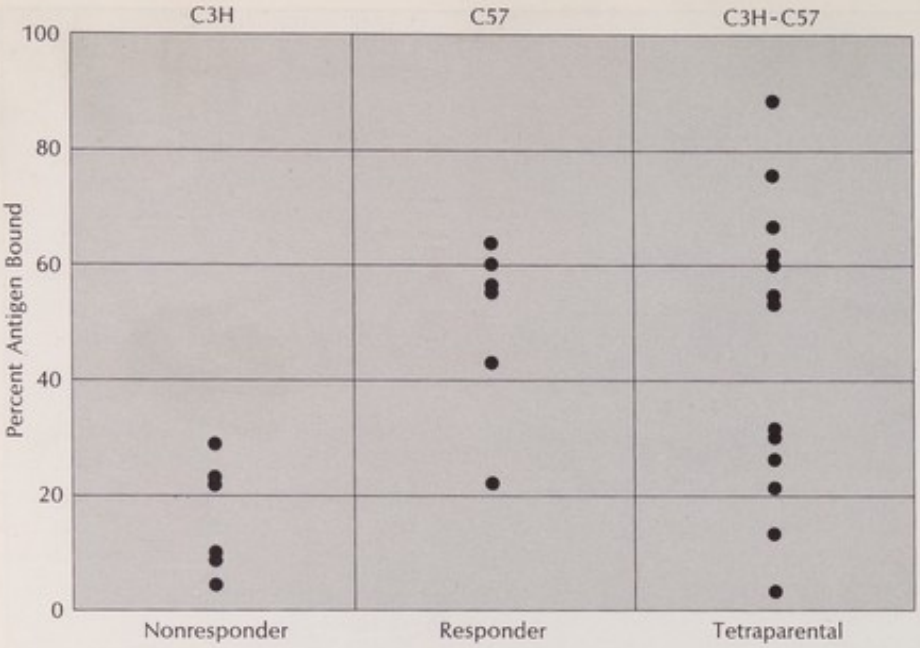
The extraordinary usefulness of tetraparental mice arises from the opportunity they provide to study cell-cell interactions for all possible combinations of responder and nonresponder T and B cells, with these interactions expressed in the form of antibody elaboration (or nonelaboration). Cells are identifiable by their immunoglobulin products. With responder T and responder B cells (top left) and responder T and nonresponder B cells, 7S antibody response is normal. With nonresponder T cells, there is little or no 7S response whether B cells are derived from responder or nonresponder. Key recognition role of T cells is in this way confirmed. Cell-cell interaction scheme involving a T cell receptor transposed to the surface of a macrophage is diagrammed here. This mechanism is supported by evidence from experiments by Feldman and others, but it is not fully proved.

gens. B cells appear to recognize antigens via antibodies on their surfaces. This antibody recognition system does not appear to operate in exactly the same way at the T cell level, since while both responders and nonresponders can make the same antibodies, the responder T cells can recognize the particular antigen, while the nonresponder T cells cannot. This hypothesis is supported by the fact that the immune response gene involved is linked to the H-2 locus and not to any of the known genes for immunoglobulin chain structure.

Most immunologists are understandably reluctant to accept the concept that recognition can operate through a system completely different from that of antibody. A number of "compromise" hypotheses have been put forth in an effort to reconcile the evidence for a T cell recognition system with the classic role of antibody. Perhaps the one that has the widest acceptance is that the immune response gene or genes really code for a new heavy chain class (IgV or IgT) never detected before because it is uniquely bound to the T cell surface and never circulates free in the serum. In fact, this heavy chain class may be the evolutionary precursor of other immunoglobulin linkage groups.

All of the available evidence indicates that variable regions of the heavy chains and the two types of light chains (kappa and lamda) are genetically different and the variable region genes are linked to the constant region genes for the heavy chain, the kappa light chain, and the lamda light chain. Each of these is in separate linkage groups, so that we already have evidence of three separate generators of immunoglobulin diversity. In light of this evidence, the existence of a fourth generator of diversity is not difficult to accept. Indeed, the evolutionary argument is extremely appealing if one conceives of a starting point in which a small gene product of, say, 110 amino acids is the progenitor. From that point on a whole series of immunoglobulins could have developed as the need of the host to recognize foreignness grew; in the course of evolution, gene duplication could well have led to an immunoglobulin with specific C and V regions always bound to cells. Subsequently, further gene duplication and translocation could have led to "migration" of the gene, the emergence of light and heavy chains, and the development of classical circulating immunoglobulins.

A second compromise hypothesis is that the immune response (Ir) gene products are cell surface antigens that interact with a typical immunoglobulin receptor and modify the function. The difficulty with this concept is that these genes seem to have a degree and range of specificity equal to antibody; the hypothesis then merely becomes another way of postulating a new "antibody" class. The most extreme interpretation is that Ir gene products are in fact T cell receptors but they have a structure completely



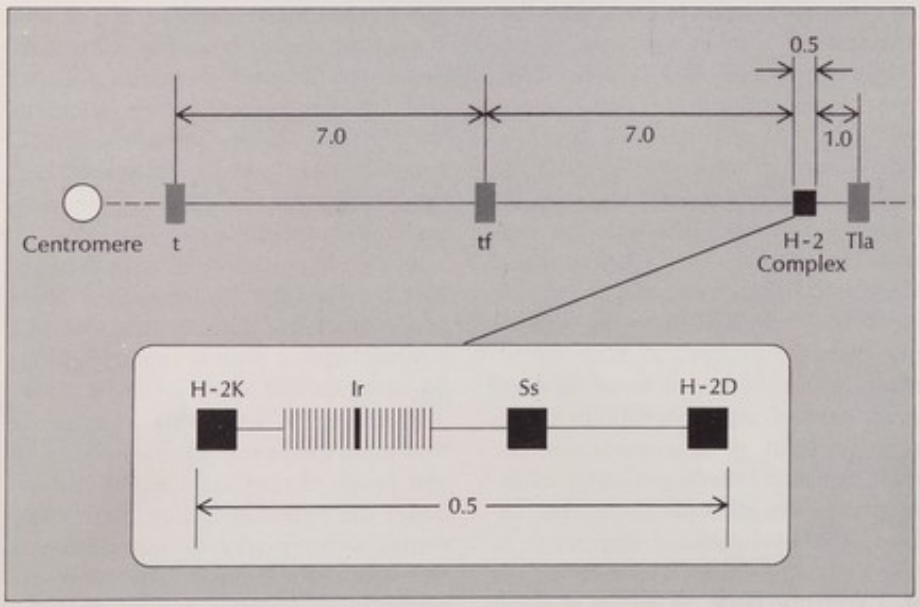
Immune responses in nonresponder (C3H), responder (C57), and C3H-C57 tetraparentals are shown. Note that tetraparental responses range almost linearly from low to high, probably depending on cell mixes. Antigen used in these studies was (T,G)-A--L.

unrelated to the known immunoglobulins.

All this is highly speculative, of course, and will remain so unless and until we learn how to isolate the specific gene product. Although many gaps remain, one can reasonably postulate that the immune response genes control a set of recognition units or receptors for antigen that are found only on T cells and are the T cell receptors; they are likely to be either a new heavy chain class found only on T cells and not circulating, or a different way of recognizing foreignness on a still-to-be elucidated molecular basis. This view of the Ir gene product can then be fitted to the accepted view of the cellular interactions that take place when a foreign antigen is presented to the organism, as follows:

The bone marrow cell "sees" an antigen in the form of the particular antigenic determinant complementary to its particular surface immunoglobulin. The B cell recognizes this antigen, say DNP, on a carrier. At the same time, a T cell, perhaps through close proximity to the B cell or through secretion of a molecule that attaches to another cell (the macrophage?), recognizes another part of the same antigen – another determinant. This dual recognition having taken place, a signal is passed from T to B cell, perhaps chemically, that turns the B cell on to make antibody, or switches it to make 7S rather than 19S antibody, or makes the B cell become a memory cell. The *sine qua non* seems to be *the simultaneous recognition* of the haptenic determinant by the B cell and of the carrier determinant by the T cell.

There are some established exceptions to this scheme, which certainly include functions subserved by the T cell alone, such as cytotoxicity and delayed hypersensitivity, as well as some functions the B cell fulfills alone. But on the whole, this is an extremely sophisticated biologic control system of cellular checks and balances. For example, in the course of the human antigenic experience, many endogenous substances are encountered that could cause elaboration of B cell antibodies capable of inappropriate response to self antigenic components. An illustration may be those streptococcal antigens that cross-react with heart valve or glomerular tissue (according to currently prevailing concepts of the pathogenesis of rheumatic heart disease or glomerulonephritis). The probability is that such autoimmunization will not occur unless the antigens can also "fool" the T cell recognition system.



Map of linkage group on 17th mouse chromosome that includes H-2 complex is at top; enlargement of the H-2 complex below includes Ir genes on H-2K side.

Linkage of Ir and Histocompatibility Genes

So far our discussion of the immune response genes has focused largely on their cellular expression. With this as background, it becomes appropriate to go a little more deeply into the experiments that have led to the identification and localization of these genes, particularly since much of the reasoning about their function has been derived from our knowledge of their chromosomal location. For example, the belief that the Ir genes are cell surface receptors got its origi-

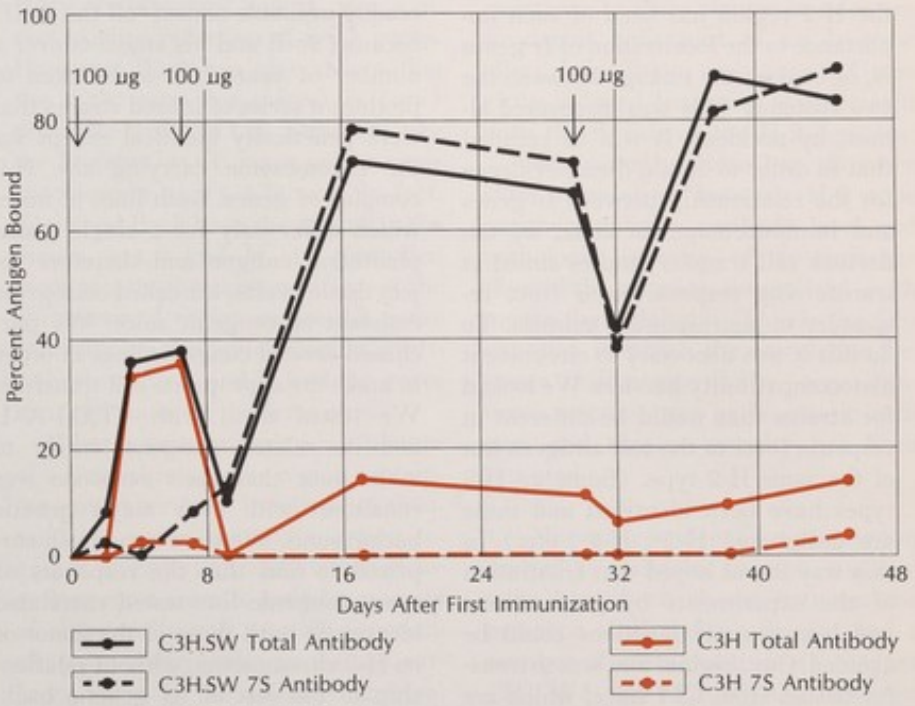
nal impetus in part from the finding that they were in a linkage group in which one finds a number of genes that code for cell surface proteins.

In the 1930's a number of scientists, particularly Gorer and Snell, began to study the mechanisms involved in tumor transplantation and rejection. For these studies, they used inbred strains of mice homozygous for every gene in the entire genome. Comparisons of genetic effects between inbred strains can therefore be done with the minimum of genetic complexity. It became apparent that there were no barriers to transplantation of tissue within an inbred strain, but that rejection of tissue was universal upon transplantation between strains. Systematic studies revealed that there were several genes, located on different chromosomes, that coded for cell surface transplantation antigens. It is now known that there are more than 30 transplantation-antigen genes in the mouse.

Before going further with this discussion, it is necessary to review the concept of linkage groups, genes located in close proximity on a single chromosome. Linkage groups are identified by the high frequency in which the genes involved are found to segregate together in crossbreeding studies of inbred strains. When two traits always segregate together, they are considered to be the effect of a single gene. However, breeding studies may result in the appearance of a recombinant, the result of crossing-over between the two chromosomes of a chromosome pair during meiosis. The frequency of crossing over by recombination is used as a measure of distance between two genes. A frequency of one recombinant per 200 offspring in an experimental cross is translated as a distance between the two genes of 0.5 map units.

By the time we embarked on our studies, it was obvious that the major barriers to grafting, those that caused rejection of allografted skin within 12 to 14 days, were a rather complex set of antigens coded for in the ninth mouse linkage group (see map, page 178). We now know that this group of genes is on the 17th chromosome. It was also clear that these H-2 genes (so called because they happened to be at the second histocompatibility locus to be studied) appeared not to be at a single-point locus, but rather to involve many genes spread over a considerable chromosome length. This was learned by cross-immunizing animals of different strains with spleen cells and producing an antiserum. An antiserum was defined by its ability to react with cells from certain strains but not from others. Thus, it is said to have a serologic specificity that is designated as 4. It may also react with another group of cells to which a specificity designation, 31, has been assigned. By crossbreeding experiments, one finds that specificities 4 and 31 segregate together with a frequency that could be accounted for either by a single gene or by two linked genes very close to each other. Similar findings linked specificities 2 and 33. However, Gorer then found that if a mouse from a 2-33 strain was crossed with one from a 4-31 strain, so that the F_1 generation had all four specificities, and the F_1 was then backcrossed with 4-31 mice, the offspring of the latter matings would sometimes lack the 33 specificity. This result indicated a crossing over in the F_1 parent between the genes coding for specificities 2 and 33. By calculating the frequency with which 2 and 33 were separable by recombination, Gorer was able to start the process of chromosome mapping. After many years of intensive work by many investigators, it became clear that the chromosome region involved in coding for H-2 histocompatibility antigens was at least 0.5 map units long and could accommodate 2,000 or more structural genes.

More recent work has shown that the H-2 genes are not spread throughout this region, as had been thought, but are concentrated at the two ends, the left hand or K end, and the right hand or D end. Parallel findings have been reported with respect to the HL-A histocompatibility antigens in man. For histocompatibility antigens in either mouse or man, it is now believed there are two clusters of gene coding for serologic specificities. In man, the loci analogous to K and D in the mouse are designated as 4 and LA, respectively. On the basis of recombination rates, there is a half map unit between these loci in both species. Recent discoveries of other ap-



Graph shows responses of C3H.SW (responder) and C3H (nonresponder) animals to three immunizing doses of (T,G)-A--L in aqueous solution. Initially the two strains respond in essentially the same way, with almost no 7S antibody detectable. Thereafter, the nonresponder animal manifests only a slight response and its antibody remains entirely 19S. The responder strain manifests marked responses to both the second and third immunizations, with all of its antibody accounted for by 7S immunoglobulin.

parently related genes in the chromosome region have served to focus attention on the question: What is going on in the rest of the chromosome region?

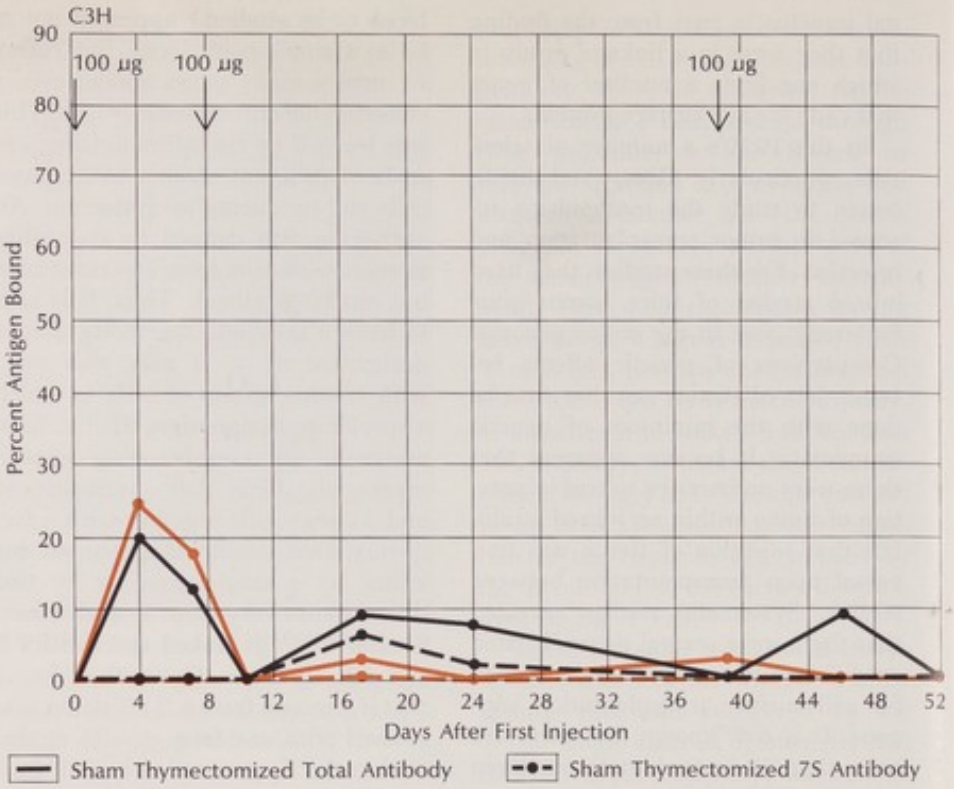
A short distance to the right of the D locus, there is a gene designated as Tla (thymus leukocyte antigen), which controls a series of cell surface antigens expressed only on thymocytes. On the left side of the chromosomal region there is another identified gene unrelated to histocompatibility; it codes for short tail (T). Perhaps most important in our context is the gene discovered by Shreffler that codes for the serum alpha globulin known as serum substance – the Ss gene. By a series of crosses between strains with low levels of serum substance (Ss^l) and those with high levels (Ss^h), it was found that the Ss gene sits on the chromosome virtually equidistant from the K and D loci. Practically, now, one has a three-point map of the H-2 region in the mouse, K to the left, D to the right, and Ss in the middle. Having such a clearly defined way to delineate the chromosome region and its two halves has proved a great convenience in efforts to locate the various Ir genes.

The reason that our knowledge of the H-2 region has been of such importance to the localization of Ir genes is, of course, the linkage between the two systems. This was discovered almost by accident. It will be recalled that in order to obtain direct evidence for the relationship between Ir genes and immunocompetent cells, we undertook cell transfer studies aimed at transferring responsiveness from responder to nonresponder animals. To do this it was necessary to circumvent histocompatibility barriers. We looked for strains that would be different in response level to the test antigens but of the same H-2 type. (Some 25 H-2 types have been identified and these are designated $H\text{-}2^a$, $H\text{-}2^b$, etc.) In this way it was hoped that frustration of the experiments by graft-vs-host and host-vs-graft reactions could be avoided. Our original aim was to transfer tissues from C57 mice, which are $H\text{-}2^b$ and responders, into C B A mice, which are $H\text{-}2^k$ and nonresponders.

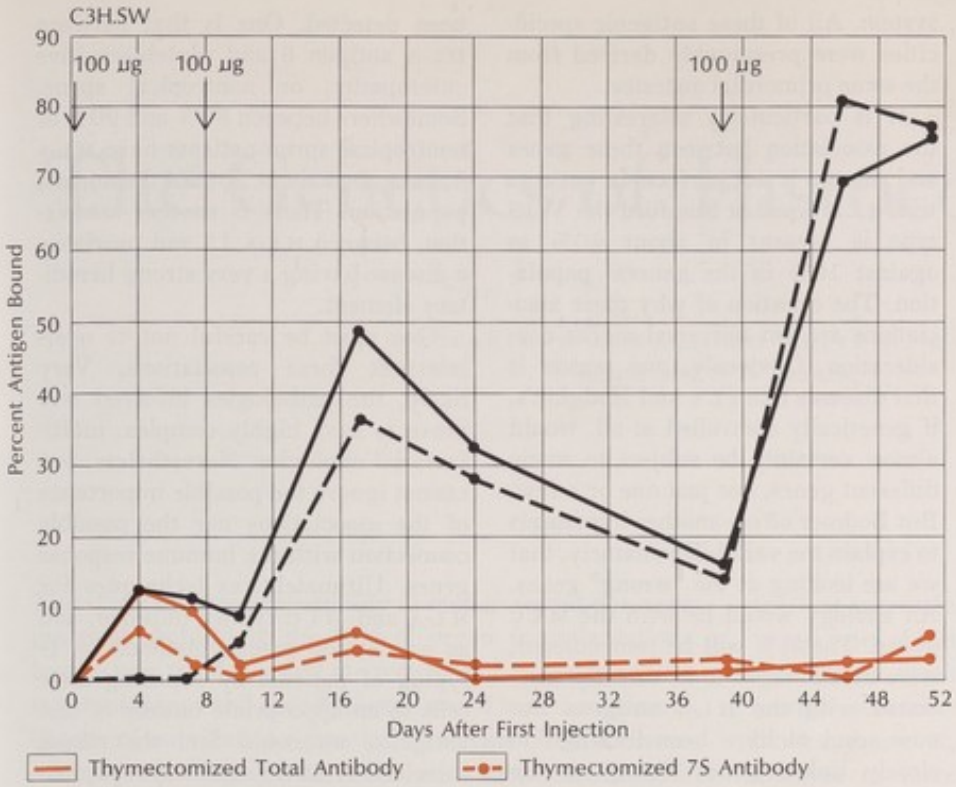
These efforts were frustrated by histoincompatibility. Therefore, we searched for inbred strains that were genetically similar to C57 but carried the $H\text{-}2^k$ type of the C B A mice. Fortunately, such mice were readily available almost "off the shelf" because Snell and his students over a number of years had undertaken to produce a series of inbred strains that were genetically identical except for the chromosome carrying the H-2 complex of genes. Such lines of mice, which differ only for a single transplantation antigen and therefore reject tissue grafts, are called coisogenic resistant or congenic mice. We purchased several congenic lines in order to make the appropriate cell transfers. We tested them with (T,G)-A--L and its related antigens simply to make sure that their response was consistent with their major genetic background. We were very much surprised to find that the responses of every congenic line tested correlated identically with those of the donor of its H-2 chromosome, without relationship to the rest of its genetic background. This suggested linkage between H-2 and Ir was confirmed by subsequent linkage studies.

Establishing the linkage between the Ir and H-2 was a piece of extreme good luck. Out of a vast number of possibilities, Ir was linked with an extremely well-studied chromosome region. We instantly became heirs to more than a half century of knowledge acquired by the geneticists who had been studying the inheritance of histocompatibility. The Ir–H-2 linkage has greatly facilitated not only the identification of some 30 specific Ir genes but also precise mapping of many of them on the chromosome.

Such mapping can be done by using animals carrying recombinant chromosomes. It will be recalled that the H-2 region has a K locus at one end and a D locus at the other, that the H-2 type is determined by genes associated with each of these loci, and finally that there is a predictable rate of crossing over in which animals will obtain the K locus from one H-2 chromosome and the D locus from another. If animals homozygous for such recombinant chromosomes are used for immunization with (T,G)-A--L, it becomes possible to determine whether the Ir gene under study is associated with the K or with the D side of H-2. By using other known points on the chromosome, such as T



The effects of thymectomy and sham thymectomy are recorded for nonresponder mice (above) and for responders (facing page). Thymectomy is combined with lethal irradiation to eliminate residual T cells. In effect, the nonresponder patterns are unaffected



by thymectomy, while those of the responder are converted to nonresponsive ones. Results are compatible with the postulate that Ir-1 gene is phenotypically expressed in T cells and that the gene controls antigen recognition by these cells.

and Ss, it is possible to localize the Ir genes with great precision. It has now been shown that the Ir genes are near the middle of the H-2 region to the right of K and to the left of Ss.

The finding by Lieberman of recombination between two genes among the relatively small number of antigens under this type of genetic control so far discovered speaks for a linear array of genes that control recognition of different foreign antigens. At the cellular level, these genes are probably expressed as receptor sites on the T cells. If the individual is a high responder to the antigen, it will recognize the carrier determinant on the antigen – say (T,G)-A--L – very well, because of a close complementarity between receptor and antigenic determinant. In the low responder, the gene is in some way mutant. It presumably still codes for a receptor on the T cell but this receptor is deficient in its complementarity for the (T,G)-A--L carrier determinant, although it may have great efficiency in recognizing another antigen.

The interest in this genetic system is growing. For example, until recently it had been assumed that the mixed leukocyte culture (MLC) test, widely used to help determine suitability of transplant donors, was a direct in vitro reflection of HL-A genes. However, Bach has now shown that the MLC gene (or genes) is (are) separable from HL-A, although like the Ir genes, MLC maps on the same chromosome as HL-A in man. In fact, Klein and Bach have shown that in the mouse, MLC and Ir are controlled from exactly the same position on the chromosomal region as are the genes controlling a third phenomenon, graft-vs-host reaction, and indeed the genes involved could be the same ones.

What is now emerging is a picture of a large chromosome region containing anywhere from hundreds to thousands of genes that have arisen together in evolution and remained together; most seem to code for a series of structurally and possibly functionally related molecules on the cell surface. What is the selective survival advantage of such a complex? Logically, it could be a sophisticated system for cell recognition and cell-cell interaction, related both to protection against foreignness and identification of self components.

Clinical Implications

Clinically, such a system has some important implications. Among them, of course, would be a possible role for genetically determined defects of the Ir system in the etiology of both autoimmune and malignant disease. If one recalls that the failure in the nonresponders is not in ability to produce antibody but in the necessary concomitant recognition system, it seems quite plausible that in autoimmune disease there is perhaps an inappropriate recognition of foreignness. Autoimmune reactions could then arise by a number of routes. For example, a number of workers have postulated that recognition by T cells of viral determinants on the surface of B cells could stimulate the latter to produce antibodies, some of which might be autoantibodies whose elaboration would ordinarily not occur. Similarly in cancer, defective T cells might fail to recognize tumor-specific antigens and thereby permit tumor growth.

There are also H-2 associations that have been found with respect to certain viruses, particularly the Gross leukemia virus and the lymphocytic choriomeningitis virus (LCM) in mice. Patterns of susceptibility and resistance to these viruses in mice are consistent with part of an Ir gene effect. In addition, resistance to Gross virus–induced leukemia is associated with the K side of the H-2 complex.

Very recently, Rose has shown a close correlation between H-2 type antithyroglobulin antibody level and pathologic severity of autoimmune thyroiditis in mice immunized with mouse thyroglobulin. This is the first situation in which the correlation between histocompatibility type, immune response to a pathogenetic antigen, and disease severity has been reported.

Nor is the evidence for association limited to animals. In man, associations have been shown between the presence of Hodgkin's disease and a number of cross-reacting HL-A types (5, W5, W15, W18), between systemic lupus erythematosus and HL-A type W15 in Caucasians and W5 in blacks, and between type W27 and ankylosing spondylitis in Caucasians. These associations are by no means absolute; in fact, in some

Frequency of HL-A Specificities

	HL-A Antigens	Percent Positive: Normal Controls	Percent Positive: SLE Patients
LA SERIES	HL-A1	27	41
	HL-A2	46	59
	W28 (Ba*)	11	24
	HL-A3	33	21
	HL-A9	22	21
	HL-A10	14	16
	HL-A11	6	20
	La-W	15	7
4 SERIES	HL-A5	11	16
	W5 (Te5)	27	16
	W18 (Te18)	9	12
	HL-A7	29	17
	W22 (AA)	5	0
	W27 (FJH)	5	0
	HL-A8	16	36
	W14 (Maki)	9	4
	HL-A12	27	21
	HL-A13	3	12
	W10 (BB)	11	28
	W15	10	36
	W17	5	8

A relationship may exist between HL-A specificities and SLE in humans. There is a statistically significant rise of HL-A8 and W15 specificities in SLE patients. Later study did not confirm the relationship of HL-A8 to SLE.

cases they seem to be rather weak and variable, but they do have statistical significance. Bodmer has presented evidence that HL-A types 5, W5, W15, and W18 are a set of cross-reacting antigenic specificities at the 4 locus, which seems to be the human analogue of the K locus in the H-2 system. All of these antigenic specificities were presumably derived from the same primordial ancestor.

It is particularly interesting that the association between these genes and disease is not perfect. In patients with SLE typed at Stanford the W15 type is present in about 40% as against 10% in the general population. The question of why these associations are not universal merits consideration. Obviously, one reason is that diseases like SLE and Hodgkin's, if genetically controlled at all, would almost certainly be subject to many different genes, not just one or a few. But Bodmer offers another hypothesis to explain the variability, namely, that we are looking at the "wrong" genes. An analogy would be with the MLC genes. These, it will be remembered, were first thought to be directly associated with the HL-A antigens but now seem to have been localized to closely linked genes sitting on the chromosome at or very close to the Ir genes. Could the Ir genes be the ones involved in coding susceptibility to these diseases? If so, this would explain the rather weak associations between HL-A and specific diseases, since crossing over in an outbred population such as man would have destroyed the association between HL-A and any possible human Ir genes.

Until recently, there was no evidence that H-linked Ir genes existed in man. However, Levine at New York University has now shown a close association between one HL-A chromosome (haplotype) and the level of IgE antibody response to ragweed allergen E and clinical symptoms of ragweed hayfever. This is the first evidence that H-linked Ir genes exist in man and have a role in disease pathogenesis.

Other striking associations have been detected. One is that between HL-A antigen 8 and gluten-sensitive enteropathy, or nontropical sprue. Somewhere between 80% and 90% of nontropical sprue patients have HL-A 8, as against about 20% of the normal population. There is another association, between HL-A 13 and psoriasis, a disease having a very strong hereditary element.

One must be careful not to overinterpret these associations. Very likely, the pathologies involved will prove to have highly complex, multifactorial etiologies. Nevertheless, one cannot ignore the possible importance of the associations nor the possible connection with the immune response genes. Ultimately, as techniques for HL-A and MLC typing improve, and as we develop precise methods for Ir typing of patients by exposing their cells to an appropriate battery of test antigens, we could find that these correlations have significant adjunctive diagnostic value as well as predictive value. For example, one can envision that if the association with nontropical sprue holds up, a built-in statistical bias will develop. Thus, if the gastroenterologist knows that a patient presenting with malabsorption is HL-A 8, he will recognize that the likelihood of gluten-sensitive enteropathy as the diagnosis is greatly enhanced.

We have only scratched the surface of clinical potential. Although the evidence that H-linked Ir genes exist in man is persuasive, we will need methods for determining human Ir genotypes, and we will need their application to large numbers of patients before the role of these genes in human disease becomes clear. The evidence from animal models and the preliminary evidence in man suggests that the game is worth the candle.

18

The Genetics of Histocompatibility

MARILYN L. BACH *and* FRITZ H. BACH
University of Wisconsin

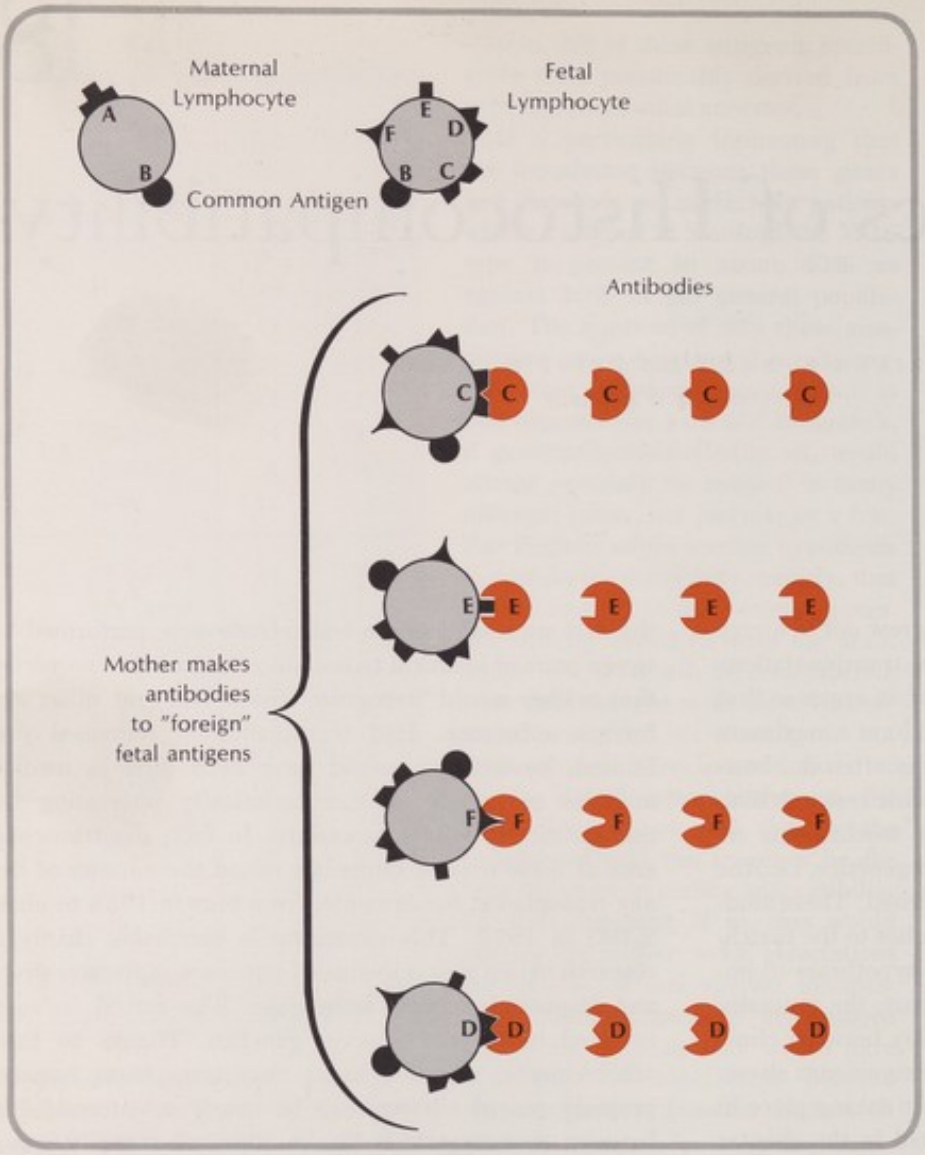
The study of histocompatibility in man grew out of a very practical problem in the field of organ transplantation: how to pair the donor and the recipient of an organ so that the transplant (and the recipient) would have a maximum chance of survival. As practical problems often do, however, this one has engendered extensive basic research both in immunology, i.e., the nature of the mechanisms responsible for transplant rejection, and in genetics, i.e., the way in which those mechanisms are inherited. These studies, in turn, have suggested new approaches to the matching problem, which have begotten new hypotheses in immunology and genetics, and so on. In part, the mutually beneficial results of this complex interplay between clinic and laboratory, and particularly the very significant alterations in our understanding that have been taking place in the past couple of years, are documented in the chapter immediately preceding this one.

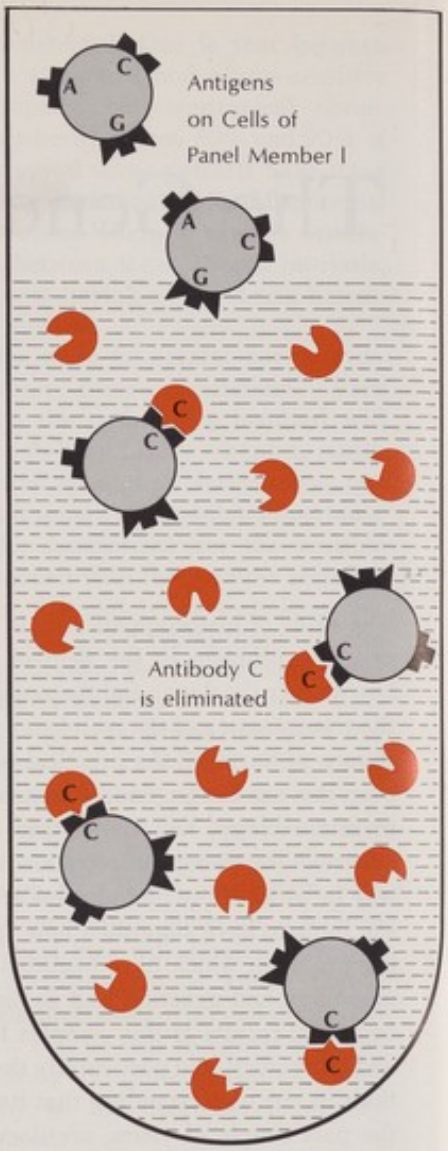
The work of McDevitt, Benacerraf, and others in defining the immune response gene system and associating it with the histocompatibility gene system certainly constitutes a major contribution in the evolution of our delineation of a single chromosomal region; on this region we can now locate a great many genes coding for a variety of related functions. It is also probable that the same chromosomal region contains a set of genes that govern the reactions seen by transplantation workers in the mixed leukocyte culture (MLC) system. These genes could either be the same as the immune response genes or they could be closely associated with them. But it now seems clear that the MLC genes are not – as was assumed until very recently – identical with the serologically identifiable genes of what might be called the "classical" major histocompatibility loci (the HL-A in man and the H-2 in mouse). In fact it would now seem appropriate to speak not of the major histocompatibility locus or loci but rather of a major histocompatibility complex (MHC).

Genetics has, of course, been deeply involved in the transplant problem from the beginning. As is well known, the first successful organ transplants were performed between pairs of identical twins, since it was held (correctly) that neither would "recognize" tissue from the other as a foreign substance. Had transplantation remained thus limited, however, it would have been filed in medical archives as merely another technically interesting but therapeutically trivial procedure. In fact, the transcendence of these narrow limits has raised the number of kidney transplants, for example, from four in 1958 to about 2,700 in 1972. This expansion is ascribable chiefly to research in two areas: the use of immunosuppressive drugs and improved pairing techniques. The latter, in turn, involved major advances in genetics. Thanks to these achievements, we now know that transplants between properly paired siblings can be nearly as successful as between monozygotic twins; in addition, there is reason to hope that these success ratios can eventually be approached in transplants even between genetically unrelated individuals.

The genetics of histocompatibility are remarkably complex, and for this reason they have been elucidated (partially) only by means of equally complex researches that have, moreover, approached the problem from several quite different angles. The following account, then, will necessarily be something less than ideally coherent; several of the disparate strands of research will have to be examined individually before they can be woven into a meaningful whole.

We begin with the central and generally accepted fact that transplant rejection in man, as well as in all other animals in which it has been studied, and its counterphenomenon, the so-called graft-vs-host reaction, are triggered by the antigenic differences between grafted tissue and recipient. These differences can be identified (to varying extents in different species) by their capacity to induce and react with antibodies both in the intact animal and in various biologic systems, some of which are to be discussed later. Having stated what for many readers must appear to be the obvious, let us now set forth two im-



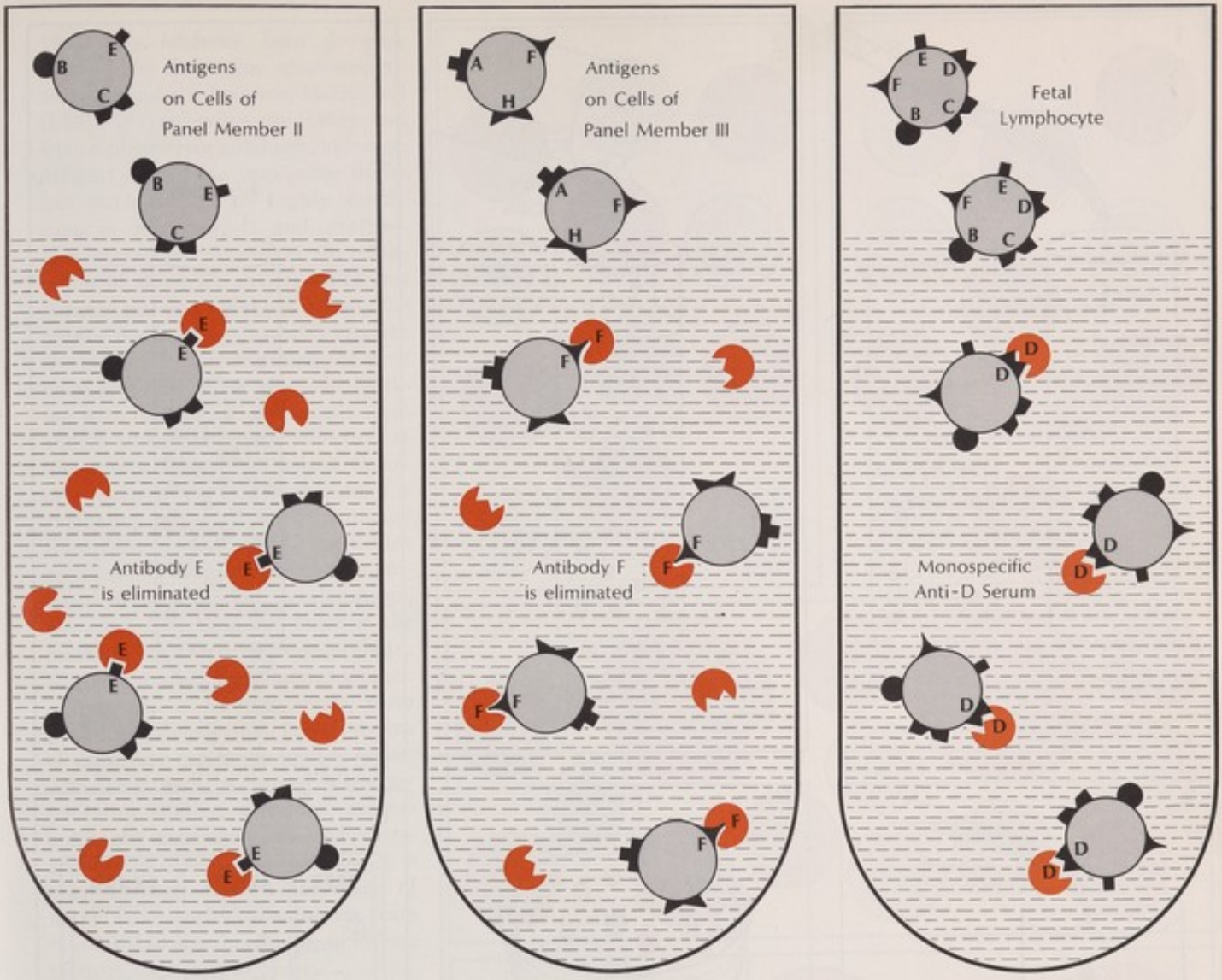


Diagrams illustrate how a monospecific serum for lymphocyte typing can be prepared from a polyspecific serum obtained from a multiparous woman. The mother's immune system can form antibodies to whatever fetal antigens she does not herself possess, e.g., C, D, E, and F. Her polyspecific serum is then successively reacted against lymphocytes from members of a panel; panel-member lymphocytes will in each case complex with and remove from serum whatever antibodies correspond to their own anti-

portant provisos, the first of which may perhaps be equally as obvious as the basic premise; that is, the actual rejection process does not involve the action of antibodies, at least to any large extent. Instead, it depends on the development of delayed hypersensitivity or cellular immunity, in which the host's lymphocytes first come to recognize certain foreign antigens as foreign substances and then attack and cause rejection of the tissues carrying those antigens. The second is one which we have just begun to appreciate, that there are certain histologic differences between transplant and host that have not been identified with antisera. Or to put it another way, we are now aware of two types of tissue differences, the serologically defined (SD) and those that can lead to a thymus-derived lymphocyte (T cell) response without causing elaboration of identifying antibodies; these we now call lymphocyte-defined (LD) differences. It is our ability to demonstrate a separability between the genetic controls of the SD and LD responses that has so greatly enlarged our knowledge of the scope of the major histocompatibility complex.

At the beginning of transplantation studies, it seemed at least possible that the rejection-triggering antigens represented a very large number of cell-surface components in the transplanted tissues, i.e., that they reflected the activity of scores of the donor's genes. Had this been true, of course, the problem of tissue matching would have remained impossibly complex except between monozygotic twins. As it turns out, however, in practical terms the genetic situation is a great deal simpler than that, though still more complex than is true of many other genetic traits that have been extensively studied.

Many of the basic concepts were developed by experiments with highly inbred strains of mice. Such strains can be refined to the point where any individual of a given strain will, for the most part, accept a graft from any other; by further manipulation, it is possible to develop strains that are

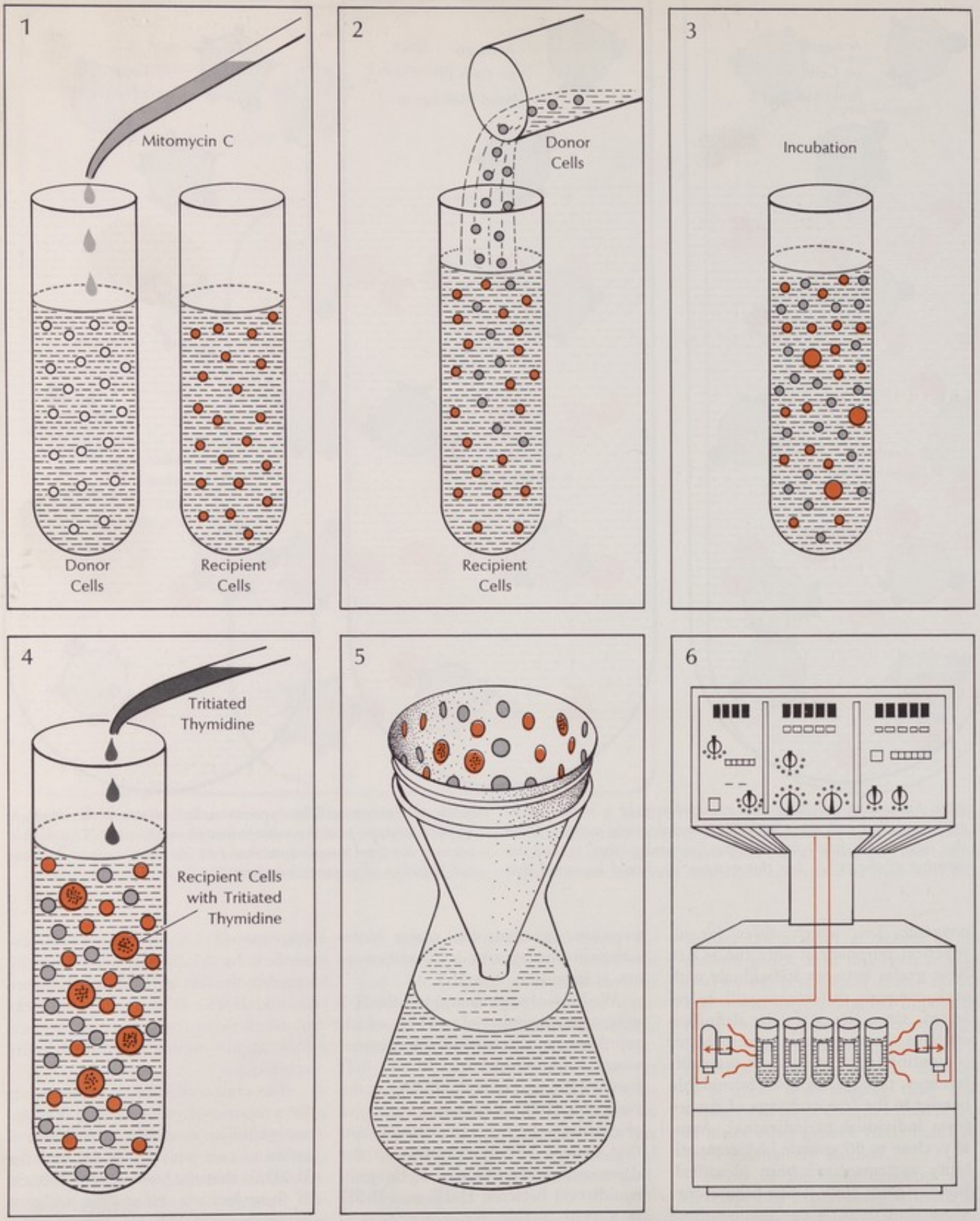


gens. Successive reactions can eventually yield a monospecific antiserum that will react with, for example, only antigen D on the infant's (ex-fetal) lymphocytes, assuming that it, but no member of the panel, has this antigen. It should be noted that while this serum will be "operationally" monospecific, it might in fact contain two (or more) distinct antibodies. This would occur if the fetal lymphocytes had two (or more) unique antigens not found in any panel member.

congenic, i.e., which differ for all practical purposes at only the MHC. Skin grafts between individuals with such congenic differences will be rejected, since the MHC by definition contains the genes that govern histocompatibility, or at least, in their differences, provide the most formidable barrier to the free exchange of tissues from individual to individual. Actually, close to 30 genetic histocompatibility systems have been identified. But for some time it has been recognized that there is one among these that is dominant in importance, and that differences between animals for this system are responsible for both rapid rejection of skin grafts and for lack of responsiveness to immunosuppression. In man, the major histocompatibility system is identified as HL-A; in mouse, it is H-2.

We have always thought of the H-2 system as consisting of two loci, which represent the extremes of the chromosomal region involved. At the left hand of the region is the H-2K, at the right the H-2D. (Left and right are oriented on the basis of a convention that defines left as closeness to the chromosomal centromere.) Roughly equidistant between H-2K and H-2D is a gene coding for a serum substance, designated Ss, which does not have any apparent functional relationship to histocompatibility but which serves as a convenient marker for mapping the region. The immune response (Ir) genes, previously described by McDevitt, have been located to the left of Ss, that is between Ss and H-2K. With these landmarks in mind, we can now take a systematic look at the major histocompatibility complex.

The reason that immunogeneticists for a number of years thought of histocompatibility control as a function of genes at two loci, the H-2K and the H-2D, is that the phenotypic products of these loci are cell-surface antigens detectable with antisera. For years the dogma has been that these SD products are the strong transplantation antigens principally responsible for both the T and B cell responses leading to rejection. Only since about

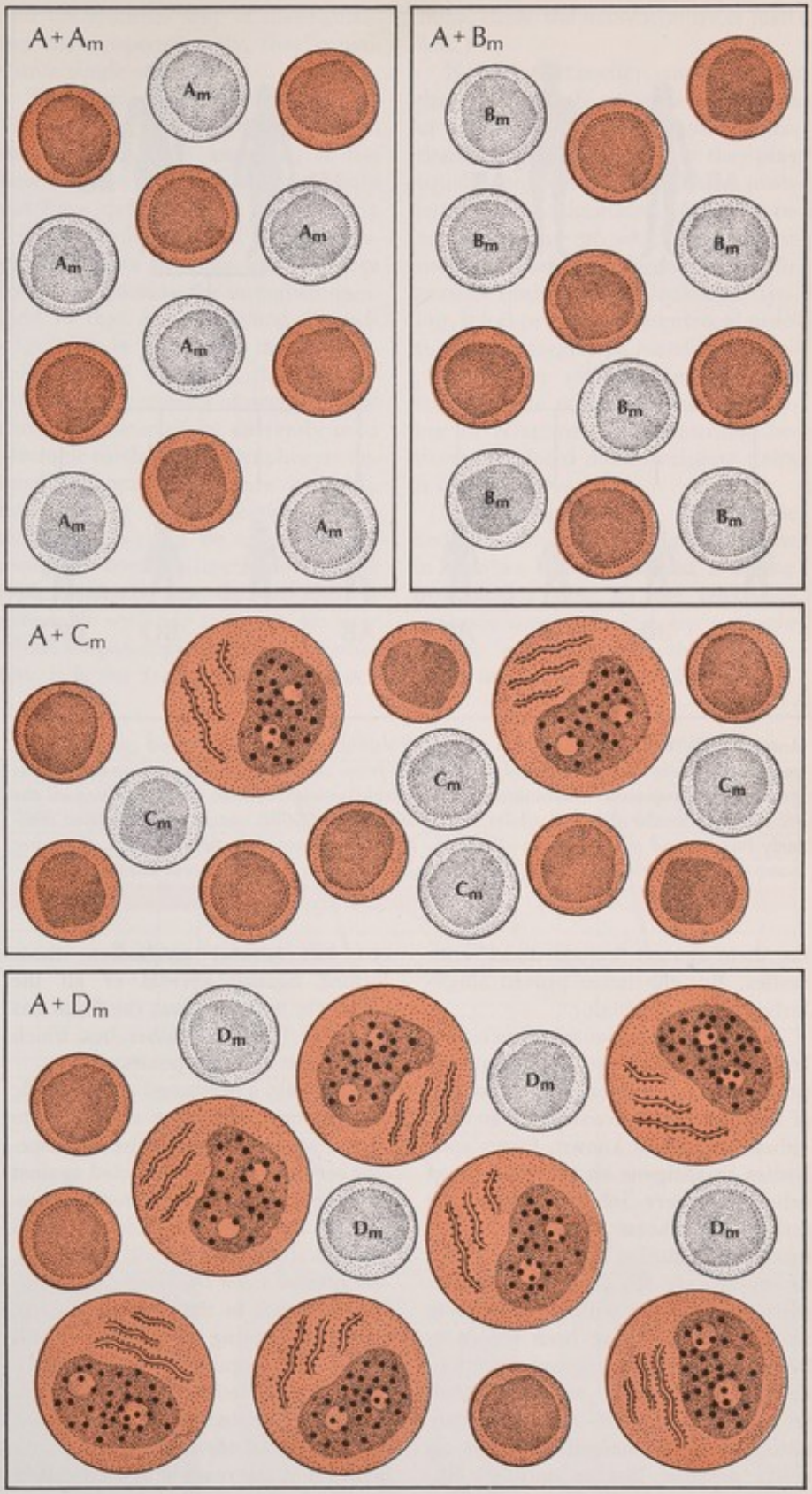


In "one-way" mixed leukocyte culture (MLC), donor leukocytes in suspension are inactivated with mitomycin C, which blocks DNA synthesis but leaves cells' antigens intact, enabling them to stimulate other cells (1). After inactivated cells are mixed with recipient leukocytes not so treated (2), the cells are cultured together for three days (3);tritiated thymidine is then added (4) and is subsequently incorporated into DNA of the recipient cells to the extent these are stimulated by the donor cells – i.e., to the extent donor cell antigens are not present on recipient cells. After a set period of time, cells are filtered and washed (5); stimulation is then measured by the radioactivity in the precipitate, which represents amount of thymidine taken up by recipient cells.

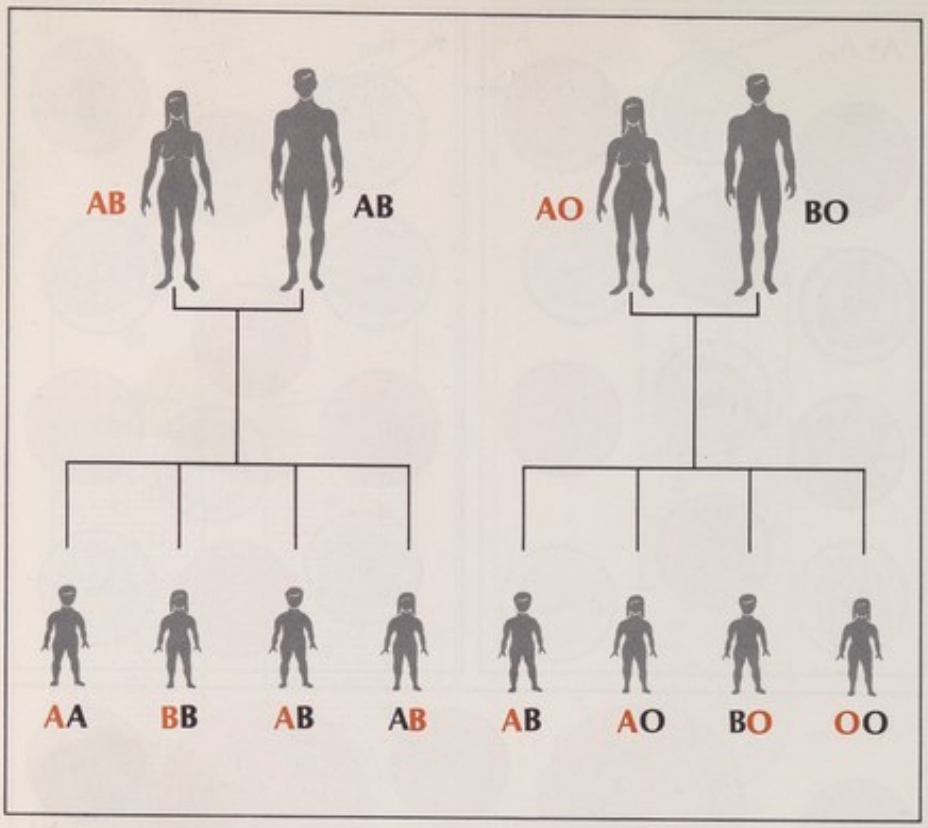
1970 has evidence been growing that within the major histocompatibility complex – between H-2K and H-2D, if you will – are other loci whose phenotypic products, although difficult or perhaps impossible to detect serologically, are highly significant in host-vs-graft and graft-vs-host interactions. Moreover, the recognition of these LD genes, which has shaken up our thinking about the events leading to allograft rejection in vivo, has had, if anything, an even more profound effect upon our thinking with respect to the clinically most widely utilized in vitro reflection of histocompatibility, the mixed leukocyte culture (MLC) technique for evaluating the suitability of potential transplant donors for a particular recipient. To appreciate the importance of these recent developments, we must first examine some of the methods by which histocompatibility has been studied in the human animal.

Man is, of course, much more genetically diverse than inbred strains of mice, nor can he be bred to order. The basic data in human histocompatibility studies, therefore, are exceedingly heterogeneous and correspondingly difficult to analyze; it is only through the collaboration of many laboratories over many years that we are now beginning to reduce them to some sort of order.

Some of the problems can be seen by briefly examining an early, and now superseded, method for matching donor and recipient, the so-called third-man test. In this procedure, an unrelated volunteer is first sensitized to the antigens of the prospective recipient and then grafted with skin from several prospective donors. The more antigens a given donor has in common with the recipient, the more rapidly will the third man reject that donor's graft. However, the fact that a donor and the recipient can be shown to share a large number of antigens says little about the number of antigens they do not share, and in transplantation it is, of course, the difference that matters. There is also the fact that the results of this test will be affected by the number of antigens shared by the recipient and the third man. Obviously, what was needed was some way of defining the individual's complement of antigens more precisely, presumably by react-



When recipient leukocytes (color) of individual A are cultured with mitomycin C-treated leukocytes from an HL-A identical donor B (B_m cells),no increase in "blast" or "transformed" cells is observed over the low number observed in a culture of A cells with A's own mitomycin C-treated cells (A_m). Mixture of A cells with C_m cells results in moderate stimulation in MLC; there is an increase in incorporation of tritiated thymidine (symbolized by black dots) and a corresponding moderate increase in blast cells. Mixture of A cells with D_m cells results in high stimulation in MLC, with marked increase in blast forms. C and D both differ, in varying degree, from A at HL-A locus.



Antigen inheritance is shown in simplified example using ABO blood groupings. In each case, child can inherit only one allele from each parent; the two matings diagrammed (many more combinations are possible) could between them produce all the six possible combinations of alleles. HL-A histocompatibility system is transmitted similarly but instead of only three alleles (A, B, O) involving only two antigens, there are more than 30, each involving a different combination of many different antigens.

ing them against "standardized" antibodies. But the latter proved singularly difficult to obtain.

A good primary source of such antibodies (to define the SD phenotypic products, i.e., antigens) is the serum of a multiparous woman. Such an individual, as is well known, forms antibodies to antigens on the fetal blood cells that were inherited from the father; Rh disease is caused by the mother's formation of such antibodies to erythrocyte Rh antigen. Our concern, however, is with the leukocyte antigens, which have been shown to be implicated in transplant rejection. Clearly, the mother will form antibodies only against such leukocyte antigens as she herself does not already possess, just as only an Rh-negative woman will form antibodies against the Rh-positive fetus. It will be equally clear, moreover, that unless the human leukocyte antibody system is far simpler than that of the mouse – and there is no reason to suppose this – the multiparous mother's serum will be multispecific, containing not one but several antibodies, those formed against several or all the leukocyte antigens that the fetus has inherited from the father but which the mother does not possess.

To obtain a presumed monospecific serum – one containing antibodies to only a single antigen – the multispecific serum can then be reacted against a "panel" of leukocytes from a number of unrelated individuals. Each individual's leukocytes will react with, *and remove from the serum*, any antibodies keyed to their own antigens, but the remaining antibodies will still show activity against appropriate antigens in other leukocytes. If this procedure is repeated enough times with different panel members, all the antibodies that can react with their collective leukocytes will have been removed. What remains will be a serum that cannot react against any member of the panel but will still react with leukocytes of the original immunizer (the fetus), i.e., with an antigen (or antigens) that it, but none of the panel members, possesses.

The reasoning behind this can perhaps be clarified somewhat by a much simplified example. Let us suppose that the fetal leukocytes contain only three antigens that the mother does not possess, called A, B, and C. The mother's serum, then, will contain only three antibodies: a, b, and c. If the first leukocytes in the panel contain, say, antigens B, D, and F, they will remove b from the serum; a second panel member with antigens C, F, and G will extract c. The modified serum will then react with neither member of the panel, but will react with the A antigen in the fetal leukocytes, and will therefore be defined as a monospecific serum.

In practice, of course, the number of antigens would be unknown but presumably much greater than in our example. In addition, there would be no way of assigning particular antigens to particular panel members (had there been, the whole rigmarole would be unnecessary). It will be clear, likewise, that even in our simplified example it would have been perfectly possible to obtain a "monospecific" serum that actually contained antibodies h, k, l, and m along with a – assuming that the fetus possessed the corresponding antigens but none of the other panel members, nor the mother, did. The "monospecific" sera obtained by this method, in other words, are defined in a purely operational sense and, moreover, only in terms of the particular panel employed in a given case. There are other methods by which single antigens can be defined, but these are equally operational and even more complicated.

The initial fruits of these involved operational definitions can be imagined with no great difficulty: confusion. At the Second International Histocompatibility Workshop, held in Leyden in 1965, some scores of antisera prepared in half a dozen laboratories were tested against a panel of 45 individuals. No two sera gave identical results. Happily, however, there was enough similarity in some cases to indicate that different investigators were talking about more or less the same thing. With this encouragement, and through exchange of sera among laboratories, work during subsequent years has clarified the picture to the point where, in one way or another, more than a dozen antigens can

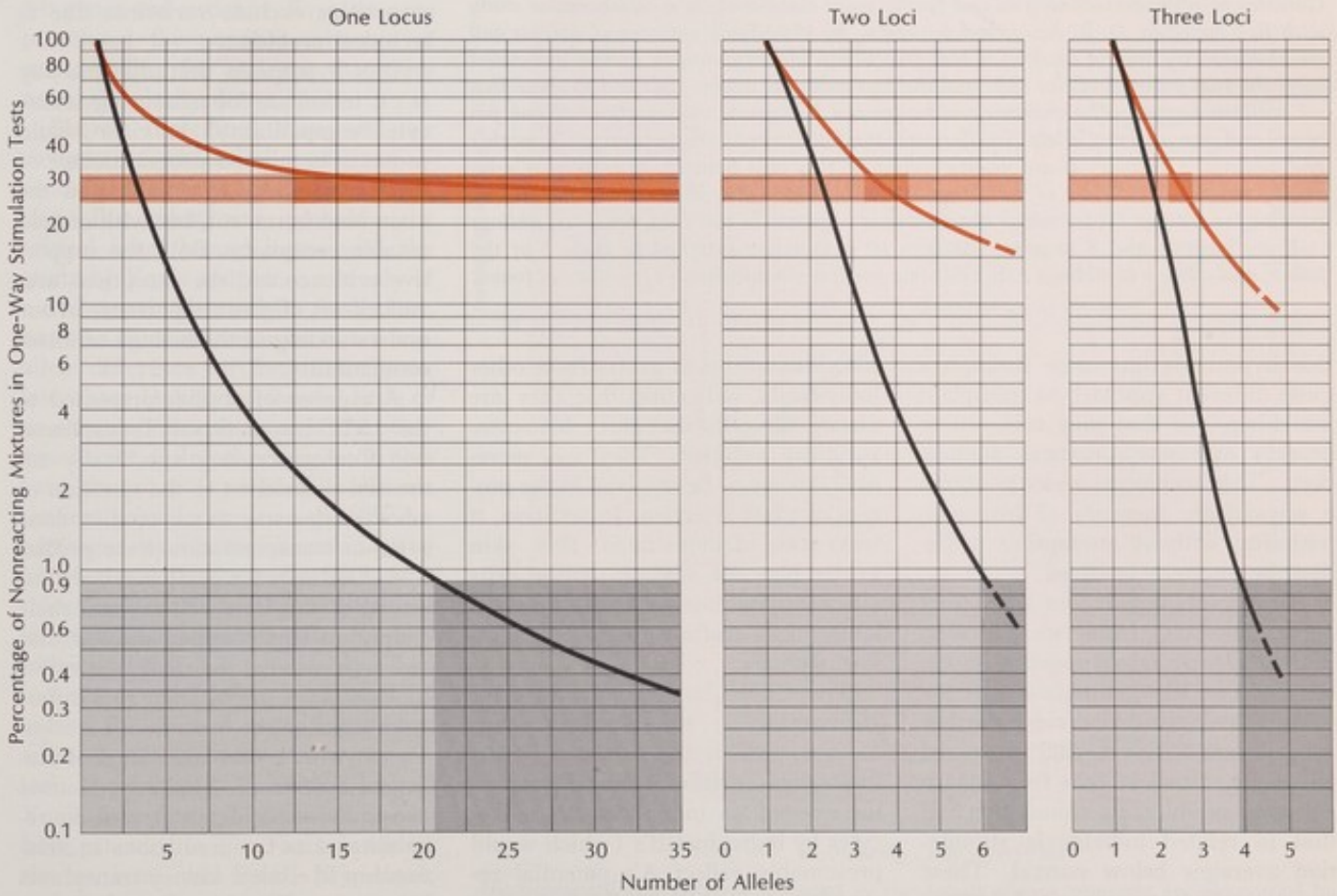
now be detected and identified by sera from a sufficient number of laboratories. These antigens have been given official identifying numbers (HL-A1, HL-A2, etc.).

The identity, however, is still a matter of operational definition. We emphasize this because it is altogether too easy to fall into the trap of equating an antigen with some specific substance, e.g., a particular protein molecule. To some degree this is probably true, but how true remains to be seen. It is a fact that proteins, e.g., bovine serum albumin, possess not one but several antigenic determinants. It is also easy to imagine two different proteins, each capable in theory of generating a distinct antibody, that for one reason or another invariably occur together in leukocytes (or at least in all the leukocytes that have thus far been typed). With present methods, there is no way of producing an antiserum to either of such a pair singly, and therefore no way of distinguishing them; operationally, they constitute a single antigen.

There are other problems. For all the progress that has been made in defining leukocyte antigens, it has now become obvious that we could not have defined them all; to begin with we could only test for the presence of known antigens, and therefore had to be missing the unknown ones. Add to that our new-found "knowledge" that in addition to the serologically defined antigens, there is another whole category of antigens that cannot be detected by currently used serologic methods, the lymphocyte-defined antigens, which are probably important in graft rejection. (Incidentally it should be noted that in using the term "antigens" for phenotypic products that do not elicit a detectable antibody response, we may offend immunologic purists. Nevertheless, it is our feeling that on a functional basis, the nomenclature is justified.)

Nor, for that matter, can we be sure that all leukocyte antigens (either SD or LD) play a role in transplant rejections, or, if they do, whether they play equally important roles. These problems have stimulated further researches, some of which will be recounted later on. Because of the present limitations of leukocyte typing, it has proved of unequivocal practical value largely in matching transplant donors and recipients among siblings. The evidence bearing on typing in relationship to matching between unrelated donor-recipient pairs is less definitive.

More important are the genetic findings. Analysis of antigen patterns in families has established a strong presumption that all the leukocyte antigens are controlled by a single genetic system comparable to the MHC in the mouse. Confirmation of



The number of loci and alleles involved in control of reactions in MLC tests can be predicted by calculating the expected frequency of nonstimulation for siblings (color curves) and unrelated persons (black curves), using different assumed numbers of loci and alleles. Horizontal bars show observed frequencies of nonreaction in MLC tests at 95% confidence levels (darkened areas show where expected and observed frequencies coincide). For each assumed number of loci, the observed data define a possible range in number of alleles but only in the case of a single locus do the ranges for siblings and unrelated persons overlap.

Family	Chromosomes	Antigens 2	3	4	5	6	7	8	9	10	11	12
Father Z	a/b	+	+	−	−	+	−	−	+	+	−	−
Mother Y	c/d	+	−	+	−	+	−	+	−	−	−	−
Siblings E	b/d	+	−	−	−	+	−	+	−	+	−	−
Siblings A	b/c	+	−	+	−	+	−	−	−	+	−	−
Siblings B	b/c	+	−	+	−	+	−	−	−	+	−	−
Siblings C	b/c	+	−	+	−	+	−	−	−	+	−	−
Siblings D	b/c	+	−	+	−	+	−	−	−	+	−	−
Siblings F	a/d	+	+	−	−	+	−	+	+	−	−	−

Alleles												
Father	a	+	+						+			
	b					+				+		
Mother	c	+		+								
	d	+				+		+				

Genetics of histocompatibility in one family were elucidated, in a collaborative study with R. Ceppellini, D. B. Amos, and associates, by identifying antigens of parents and six children. By careful analysis it was possible to identify antigen groups associated with the two paternal alleles (a,b) and the two maternal alleles (c,d) and to show that all patterns found in the children can be formed by various combinations of one paternal and one maternal allele. To illustrate: the father is positive for antigen 3, mother and siblings A through E are negative. Sibling F is also positive. If antigen is arbitrarily assigned to allele a of the father, F has inherited that allele. The father is also positive for antigen 10, for which the mother is negative. In this case, however, siblings A-E are positive and F negative. Antigen 10 is therefore assigned to allele b of the father, and allele b to siblings A-E. Only one possible combination – a/c – was not found.

this hypothesis has come through a quite different approach to transplant matching, one that may turn out to possess extensive practical applications. This technique seeks to obtain a physiologic measure of histocompatibility without attempting to define the antigens involved.

Some years ago it was noted that when leukocytes from two unrelated individuals are mixed together in tissue culture they stimulate one another to enlarge and divide. Further experiments showed that a mixed leukocyte culture of cells from identical twins produces no stimulation and that in related individuals stimulation averages below normal. These and other findings suggested that the MLC could provide a useful measure of histocompatibility in man; this supposition was strengthened by the classical finding that peripheral blood lymphocytes can be used to sensitize individuals to skin grafts from other individuals, indicating that they are among the cells that carry histocompatibility antigens. They are, moreover, known to be involved in the process of graft rejection. In addition, it was soon demonstrated that skin grafts between siblings whose cells did not stimulate in MLC survived longer than grafts between "stimulating" siblings.

There were, however, a number of weaknesses in the original MLC method, notably the fact that it was impossible to differentiate the stimulus exerted on individual A's leukocytes by individual B's (which would presumably reflect A's potential response to a graft from B) from that exerted in the opposite direction (which would prefigure B's response to a graft from A). In 1966, we and our associates at the University of Wisconsin were able to develop a technique for measuring "one-way" stimulation. Cells from a potential "donor" are inactivated with the drug mitomycin C (an experimental antineoplastic compound), which prevents them from responding to stimulation but leaves intact the ability to stimulate other leukocytes from the "recipient;" the latter respond by increasing their synthesis of DNA, which can be measured through the cells' incorporation of isotopically labeled thymidine. As a control, the uptake of labeled material in this system, $A B_m$ (A active, B inactivated), is compared with that produced in the autogeneic $A A_M$ system. In some cases, notably those of siblings whose cells do not stimulate each other, we introduce a further control by stimulating A's cells with inactivated cells from another individual ($A X_m$) and by stimulating these latter with inactivated B cells ($X B_m$). By furnishing control levels for both A's response capacity and B's stimulus capacity, this final step helps exclude variations due to technical problems.

Our hypothesis that the one-way MLC technique for measuring leukocyte "compatibility" in vitro would prove to be a good prognosticator of histocompatibilty in vivo has in the main been borne out, but with certain notable exceptions. Both the supportive evidence and the exceptions are, we believe, of great importance in our understanding of the biology of histocompatibility.

A number of studies presented to the 1972 International Transplantation Conference very specifically addressed themselves to the question of whether one-way MLC studies done prior to transplantation were predictive of subsequent graft survival. For example, van Rood, Koch, and their colleagues from Leyden showed that one could predict the relative survival times of skin grafts from two different donors on the basis of MLC stimulation with a very high level of statistical confidence. Similarly, Kountz of San Francisco demonstrated the reliability of MLC predictions in relationship to clinical kidney transplants between unrelated donor-recipient pairs. Hamburger and his colleagues in Paris have for years presented data consistent with the validity of MLC prognostication.

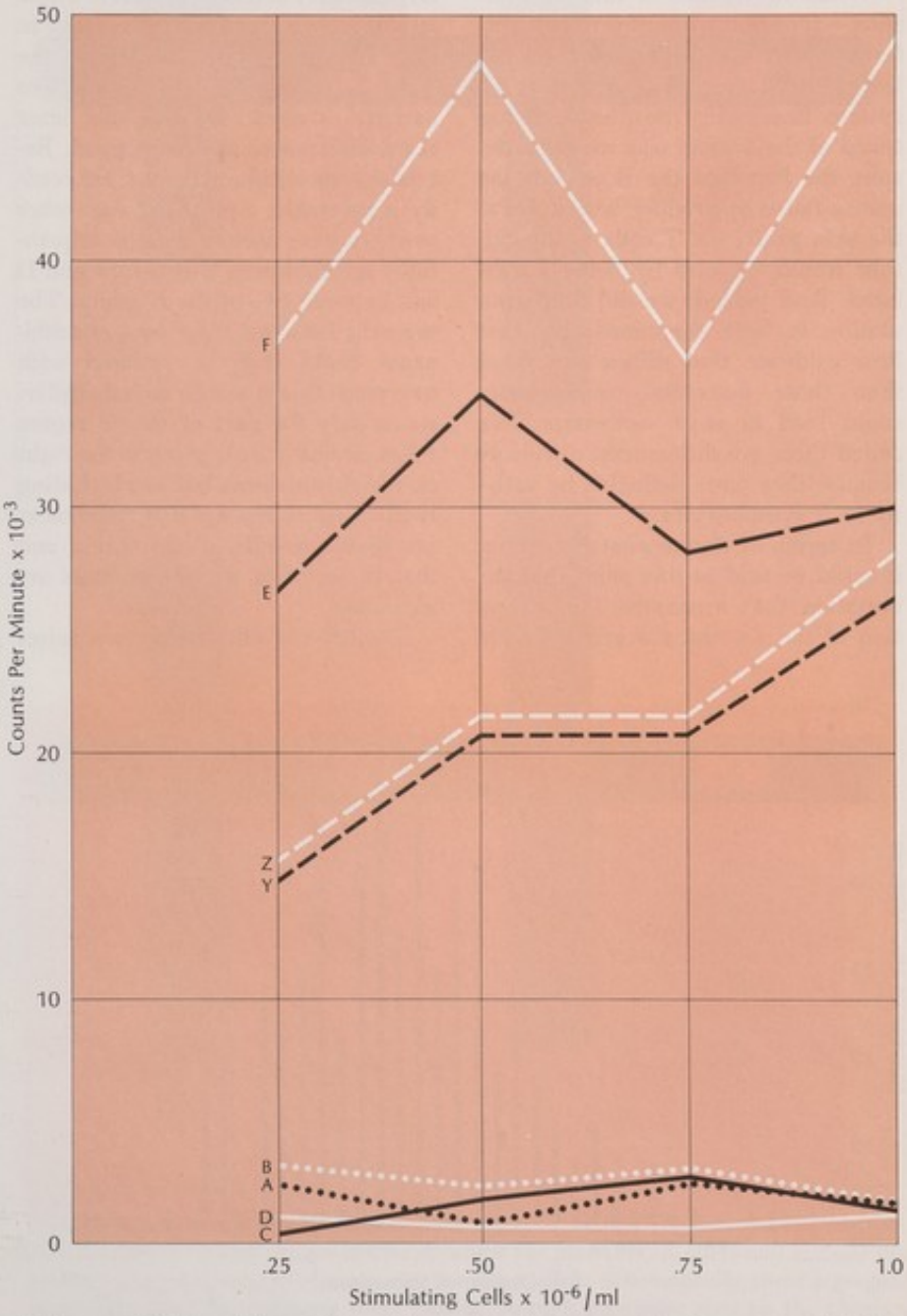
Evidence that the major histocom-

patibility complex system controls both MLC reactivity and the serologically defined leukocyte antigens actually predates considerably our awareness – or even our suspicion – of the existence of the SD-LD configuration that we now call the MHC. In fact, the first hint of the complexity came in the mid-60's. Working with Amos of Duke, we were studying families to determine whether siblings who had inherited the same HL-A (SD) antigens would also manifest identity if their cells were combined in the MLC system. In the vast majority of cases, the answer was affirmative, that is to say brothers and sisters who had inherited the same serologically defined antigens did not demonstrate MLC stimulation. However, we found one exceptional sibling pair. They were identical with respect to their SD antigens, but their cells did stimulate in MLC. In retrospect, this can now be seen as the first suggestion of the existence of what we now talk of as LD genes and antigens. Even then we postulated the possible existence of genes coding for antigens that were not serologically detectable.

Shortly thereafter, Plate and her coworkers found a similar case and essentially suggested the same explanation. But what was probably the critical breakthrough came with the study by Yunis and Amos not only of a sibling pair identical for the SD loci, but of a third sibling who had a one-allele difference for the SD antigens, but whose leukocytes showed little or no stimulation with those of the other two in MLC. In other words, in this case, the HL-A (SD) identity did not prevent stimulation, nor did the SD heterogeneity cause stimulation – a bidirectional demonstration that something besides the serologically defined antigens could govern histocompatibility, and that it is possible that during the course of recombinational events, the SD loci are separable from what we now postulate to be the LD loci. Incidentally, these findings of Yunis and Amos have subsequently been extended in The Netherlands by the work of van Rood, Eijsvoogel, and their colleagues. They have described a number of families in which recombinations have occurred with resultant separation of SD and LD components of the MHC.

These findings in man that we have just cited constitute immensely important contributions to our understanding of the MHC. However, because of the obvious advantages of dealing with inbred strains of rapidly and controllably reproducing species, it is work with mice that has given us a more precise and sharply focused view of what goes on within the major histocompatibility complex. Let us review some of that investigation now.

The first clue to us that the H-2 system had a greater-than-realized complexity came in a collaborative study with Bailey of Bar Harbor, Maine. He had produced a congenic mutant strain of the C57 Black 6 mouse in which the mutation involved what we then thought of as the MHC locus; that is to say, the new H(zl) strain was for practical purposes the same as the original strain except for



Histocompatibility patterns deduced in the family shown opposite are supported by one-way MLC tests in which leukocytes from child A were successively reacted against inactivated cells from every other family member. Children B, C, and D, having same genotype as A (b/c), produced zero stimulation in A. Leukocytes from E and from parents, all of whom were deduced to have one allele in common with A and to differ by one allele, produced moderate stimulation. Leukocytes from F, presumed to have no allele in common with A (i.e., a/d vs b/c) produced maximum stimulation.

an MHC difference. Bailey was able to localize the mutation as being on the H-2K side; and in skin grafting experiments, there was mutual rejection of grafts between the congenic animals in a manner consonant with a major transplantation barrier. Working with Chary and Snell, Bailey then tried to produce antisera to the antigens responsible for rejection. But no serologically identifiable antigens could be detected.

When we learned of this phenomenon – rejection without serologic difference – we asked Bailey if he would let us study these mice in the MLC system. Essentially the question being asked of the system was whether, despite the fact that the B cells of the mouse failed to produce antibodies to the skin graft, the T cells in the culture would respond by being stimulated. And indeed, we did find stimulation in both directions, the first firm evidence that differences other than those detectable serologically could lead to MLC activation. We called these LD differences, obviously because they were definable by activity in T lymphocytes.

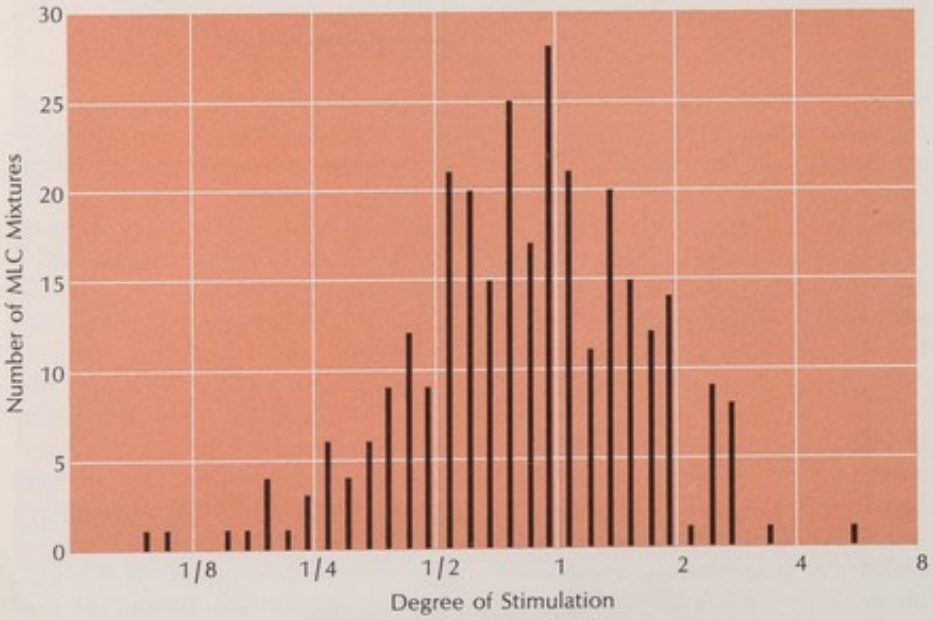
In terms of chromosomal location, it could be said at this point that the mutation that accounted for separation of the SD and LD genes was to the left of the Ss. We could not tell whether it was within the Ir region, between Ir and H-2K, or even to the left of H-2K, that is outside of the H-2 locus. If the last were true, there was a possibility that we were still dealing with differences at the H-2K locus not detectable by our methodology. The opportunity to test this possibility came in experiments with Klein of Ann Arbor. In these investigations, the mouse strains used were B10.A (2R) and B10.A (4R). These strains have a significant advantage in that they are both derived from the same F_1 heterozygote between two parental strains, so that the exact MHC differences can be mapped. Recombinants within the MHC are readily recoverable. Stimpfling and other workers were able to recover sequentially recombinants between Ss and D and between two of the Ir genes. The segment between these two recombinants could then be isolated with assurance that it would include differences only for part of the Ir region (that closest to Ss), going to the right on the chromosome but not including H-2D. It thus appears that differences are in the middle of the MHC, and that in terms of SD the animals are identical.

The 2R and 4R mice did not reject skin grafts from each other. But it was the mixed leukocyte culture studies that produced the remarkable results. In most MLC experiments, whether in unrelated humans or inbred mice, one finds that if cells stimulate in one direction, they will also stimulate in the opposite direction. But here, we found that 4R cells responded to 2R stimulation, but 2R cells essentially were unresponsive to 4R cells. In the light of the mutual tolerance of skin grafts, these MLC results suggested to some that perhaps the mixed leukocyte culture is a very elegant in vitro model with no relevance to the allograft response.

We were, understandably I think, reluctant to accept this proposition. Instead, with Klein, we decided to look at another parameter of the homograft reaction, one which incidentally Simonsen, among others, had always equated with MLC – the graft-vs-host reaction. Klein injected 2R lymphocytes into 4R animals, and then 4R cells into 2R animals. In the first instance, there was no response, but in the second he was able to observe a very strong reaction which was unmistakably graft vs host. In short, there was complete correlation with the MLC observations. There appears to be a region of the MHC between the two serologically defined loci and separable from them by recombination, closely associated with and perhaps identical to the Ir region, the phenotypic products of which lead to MLC activation and to graft-vs-host reactions. These phenotypic products are lymphocytically but not serologically definable.

The one obvious question that arises from this is, if there are LD differences that lead to T cell activation in the form of graft-vs-host reactions, why do they not also lead to allograft rejection via host-vs-graft reactions?

So as not to arouse expectations unduly, let us begin by stating that we are not yet able to provide a definite answer to the question. However, there are some intriguing pieces of evidence that are important in themselves and that provide a springboard for some interesting hypotheses and speculations. First, there has been the most elegant family study in man by Eijsvoogel and his colleagues showing that with cells from siblings one



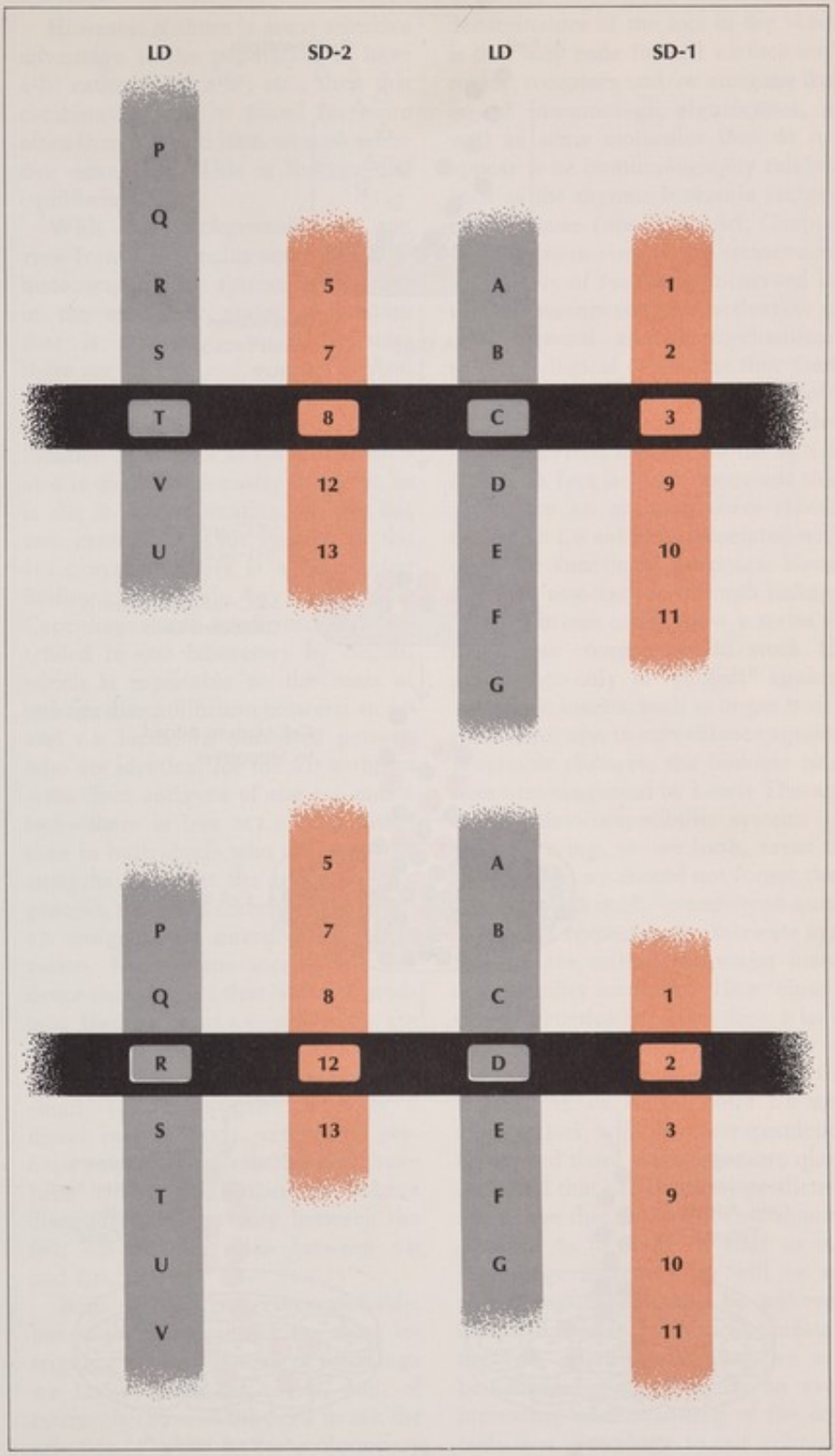
Composite presentation of 17 MLC experiments – in each of which the cells of a potential recipient are tested with stimulating cells from 15 to 20 different potential unrelated donors – indicates that there is a very great variation (more than 60-fold) between the weakest stimulating mixture, i.e., the best recipient-donor combination, and the strongest stimulating mixture. In each experiment the average stimulation of the 15 to 20 different mixtures is equated to 1, and the stimulation in each of the individual mixtures is plotted as a fraction of the average for that experiment.

can obtain MLC activation with an LD difference either in the presence or absence of an SD difference. However, if the cells were then tested in another system, that of cell-mediated lympholysis (CML), in which the critical criterion is the ability of cells from one individual to proliferate and kill target cells from the other individual, one could differentiate between those pairs with only an LD difference and those with both LD and SD differences; specifically, for killing, the SD difference was needed. Schendel and Alter in our laboratory have obtained similar results in some of the mouse strains discussed above.

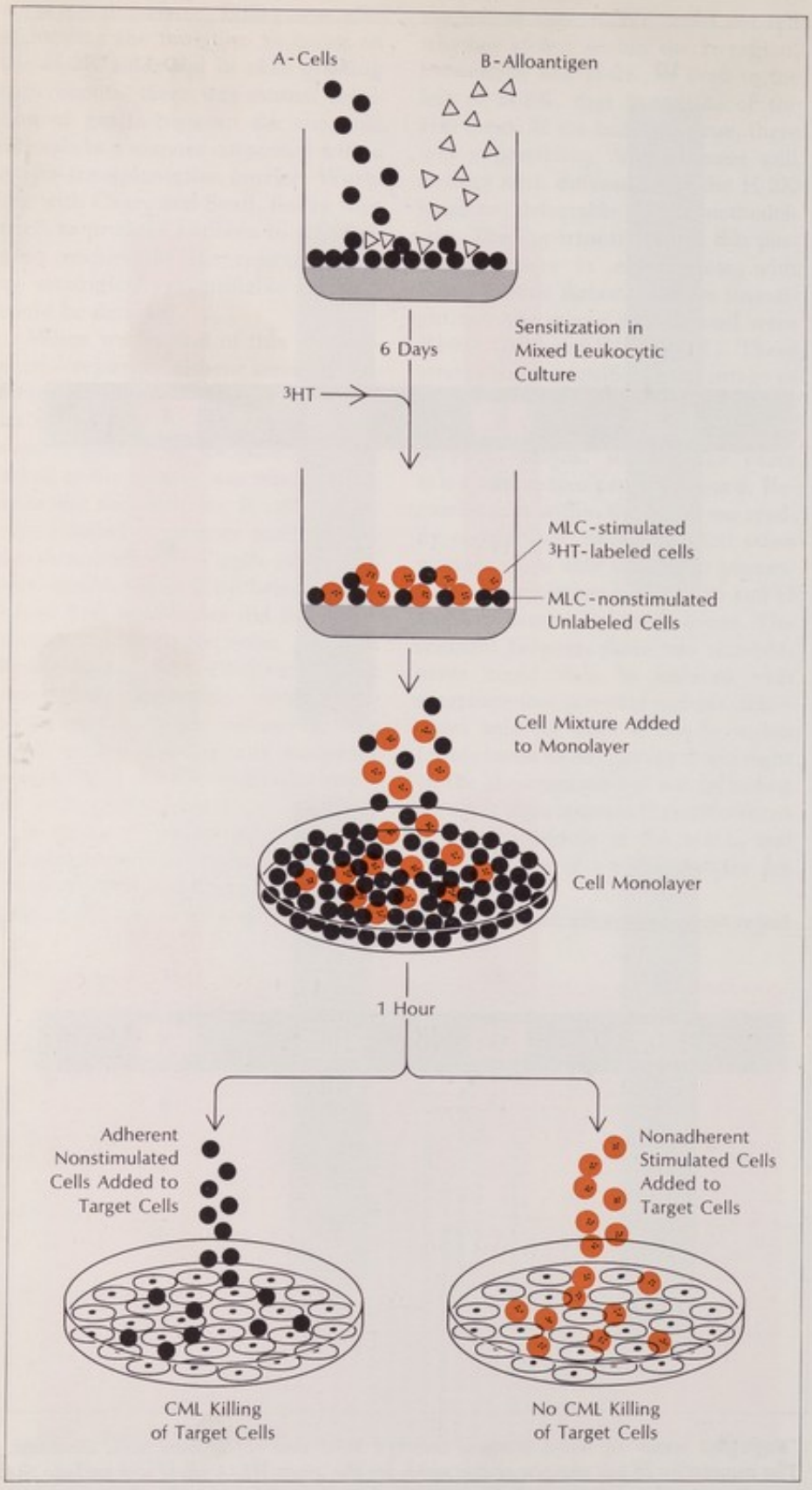
However, Alter has also shown that in some strains, at least, one could get active killing, as measured in the CML system, without a detectable SD difference. A possible explanation for the disparity is that there are two types of LD loci; one results in activation of MLC, the other determines the target for CML. Be that as it may, it seems fair to say that the phenotypic products required to get MLC activation and those that induce cytotoxicity are, at least in some cases, controlled at different genetic loci, and the latter may map with SD.

By extension, this also suggests that in addition to the SD and LD loci, the MHC may also include a third component, one that provides the genetic control for cytotoxicity in the cell-mediated cytolysis system. This may or may not be separable from SD and could be a model, although not likely an exclusive one, for the effector arm of allograft rejection. (And one might add that at some future date the genes functioning at such a locus might provide the ideal target for immunosuppression.)

If one starts from the premise that cytotoxicity is representative of the effector arm of graft rejection and that its genetic controls can by recombination be separated from leukocyte activation controlled at LD, one can visualize this possibility: LD differences within the major histocompatibility complex lead to proliferation of one population of T lymphocytes. At times, as in the C57 Black 6 × H(zl) model, this population can by itself induce graft rejection. More often, these cells act as helpers required to activate a second T cell population or a B cell population (a relationship



"Segregant series" of HL-A antigens (marked SD-1 and SD-2) have been identified. The association of any antigen of one series for the given HL-A allele will exclude the gene for any other antigen of that series from that allele. Schematically, these segregant series are shown as imaginary slides passing through imaginary slots on the chromosome; slots can be filled by genetic material specifying only one antigen from each group. Individual inheriting a gene for a different antigen in each series from his father and mother has genes 2, 3, 8 and 12. There is also at least one LD locus (presumably to the left of SD-2), alleles of which are represented by letters P, Q, R, S, etc. There may be a second LD locus between SD-1 and SD-2. Alleles of this locus are represented by letters A, B, etc. If both LD loci exist, individual would have four antigens associated with the two SD loci, and four "antigens" for the LD loci.



The experimental demonstration that mixed leukocyte culture stimulation can be separated from cell-mediated lympholysis (presumably an expression of the effector arm of allograft rejection) is depicted. If A cells are incubated with B alloantigens in an MLC, the MLC-reactive cells will incorporate radioactive tritiated thymidine, and when placed on a monolayer will not adhere to the monolayer cells; most of these nonadherent cells will not be active in cell-mediated lympholysis. On the other hand, one can identify a nonstimulated, adherent cell population which possesses significant killing activity when added to target cells in the CML system.

perhaps analogous to the T-B cell interactions described by McDevitt in the preceding chapter). This second T cell population, controlled by another MHC difference, possibly at SD, would then act to affect cell killing and allograft rejection.

We have recently obtained some evidence consistent with such a two-cell system, paralleling the dichotomy of the genetic control over MLC and CML. Several investigators have demonstrated that it is possible to adsorb cytotoxicity-mediating cells to monolayers of specific target cells (taken from the individual initially used as donor of the sensitizing cells). We have used this principle by sensitizing cells in MLC, subsequently adsorbing the MLC-reactive cells to a target monolayer. The results indicate that whereas the CML-active cells adsorb to the monolayer, most of the cells dividing in MLC (those incorporating radioactive thymidine) do not adsorb and are found in the nonadherent cell population. This work has been done by Zier, Sondel, and Alter here in Madison.

This concept presupposes the importance of both LD and SD differences to achieve transplant rejection. It also suggests that perhaps this second T cell population might recognize SD antigens and that it might use some surface-bound immunoglobulin as its receptor, a possibility that in the light of current debate about the nature of T cell receptors is certainly controversial. To carry on with the speculation, it also can be considered possible that the gene controlling the T cell receptor is actually an Ir gene. It is clear that LD genes, or at least some of them, map very closely with Ir genes, or at least some of *them*. If LD is Ir, and Ir genes control molecules that are T cell receptors, then it may be that what these molecules recognize are the Ir products.

Probably one of the most attractive things about this speculation is that it postulates a system that would beautifully exemplify evolutionarily derived biologic efficiency. Another is that it could readily explain the predictive value of MLC with respect to transplant survival. If one needs the LD difference to initiate the sequence or to affect the degree of intensity of cytotoxic reactions leading to graft rejections, and if these

reactions are directed against SD antigens, then the severity of the reaction would, given some SD difference, be modulated largely by the degree of activation in the mixed culture – the stronger the reaction, the greater the helper cell effect upon the second T cell population. However, the correlation would not be absolute since a second parameter must be taken into account. Nor can one exclude, in the light of the fact that one can sometimes get skin graft rejections without SD differences, that what we have been talking about is only one of several pathways that arrive at the allograft reaction.

In the preceding discussion, we postulated an evolutionary advantage for the scheme we were proposing for certain interrelationships of MHC loci. Let us switch now to another concept that also has as its teleologic rationale selective advantage. To do this, however, it is necessary first to outline a genetics phenomenon called linkage disequilibrium.

Let us assume that on a given chromosome there are two loci, A and B, and that in a random population each of these will be expressed by four different forms of a trait, or four different alleles. At locus A, the alleles are a^1, a^2, a^3, and a^4. Similarly at locus B. Let us further assume that these alleles are present in the population with equal (25%) frequency. Since the two loci are linked, on any one given chromosome they will segregate together 99% of the time. Therefore, if one parent is a^1b^1 on one chromosome and a^2b^2 on the other, the offspring will receive one of the two combinations intact 99 times out of 100. In the random, unrelated population, these alleles will, given sufficient time, come into equilibrium purely on the basis of allelic frequencies. If we choose one allele of one locus, say a^1, and ask how frequently it will be found with one allele of the second locus, b^1, in linkage equilibrium, the answer will be 25% of the time.

However, if there is some selective advantage to the population to have a^1b^1 rather than a^1b^2, etc., then this combination will be found far more often than it would without such selective advantage. This is linkage disequilibrium.

With this background, one can view from a particular perspective the histocompatibility system as it exists in the randomly mated population that is humanity – HL-A. In man there are two SD loci, LA and 4. And we know that linkage disequilibrium for these loci is common. The combination of allele 1 at LA and allele 8 at 4 is disproportionately frequent, as is the 2-12 combination, to cite but two examples. With respect to the MLC system, there is a fascinating finding, first made by Sorensen in Copenhagen and confirmed and extended in our laboratory by Segall, which is explicable on the basis of linkage disequilibrium between an SD and LD locus. In unrelated persons who are identical for the SD antigens – the four antigens of the LA and 4 loci – there is less MLC stimulation than in individuals who differ for the antigens coded at the HL-A loci. In general, there is a correlation between SD antigens and mixed culture activation. Now if one accepts the evidence that the loci that code for products leading to MLC activation are separable from the SD loci, then there are two explanations for this phenomenon: 1) SD antigens can play a direct role in MCL activation, perhaps only after LD-reactive cells have been activated; 2) there is linkage disequilibrium not only between the two SD loci but also between SD and LD.

Both explanations are reasonable, but to us the second is the more intriguing because of some of the things we know about the MHC, and, of course, because of the need to ask the question, "What is the evolutionary advantage?"

In this context, one of the common denominators of the loci in the MHC is that they code for cell surface molecules, receptors and/or antigens that are of immunologic significance, as well as some molecules that do not appear to be immunologically related, such as the thymus leukemia antigen in the mouse (see McDevitt, Chapter 17). Now, in view of the tremendous complexity of functions subserved by the cell membrane, the activation of cells, general control mechanisms, etc., it is logical to assume that there is a nearly perfect fit among the molecules on the membrane that are coded for at different loci within the MHC. It may in fact be very important that given two SD antigens there should be certain LD antigens associated with them for functional purposes. Keeping such associations through linkage disequilibrium could allow a series of membrane components to work together not only to "protect" against iatrogenic insults, such as organ transplants, but also in surveillance against neoplastic changes, the biologic purpose first suggested by Lewis Thomas for the histocompatibility system.

In entering, as we have, areas of speculation, we should not forget that there are still many unanswered questions with respect to the intricate system we are calling the major histocompatibility complex. How closely can we genetically isolate the LD loci? Are they restricted to the region of the Ir loci? Are they identical with the Ir loci? As we isolate more LD and more Ir loci, will the correspondence hold? And there are many more questions. All that can be safely predicted, other than that there is a vast amount of work to be done, is that as our knowledge increases we will be enriched immunologically by an ever-increasing ability to foresee accurately the fate of allografts, and we will be enriched genetically by an ever-increasing understanding of the controls that give shape to our existence from molecule to whole organism.

Section Five

Multifactorial Genetic Disease

19

Multifactorial Genetic Disease

C. O. CARTER
Medical Research Council, London

Readers of this text will have noted a salient fact about the disorders thus far described: with few exceptions, they are either rare or very rare, with incidences ranging downward from one per thousand to one in several hundred thousand. Collectively, these inborn errors of metabolism and chromosomal defects constitute a public health problem of no small magnitude; individually, they are so uncommon that the average physician may practice for years, even his entire professional life, without seeing a case.

Yet it is becoming increasingly clear that the pathologic and epidemiologic significance of genetic disorders is far greater than might be inferred from the above. Every physician knows of a host of conditions that, in the old phrase, "run in families." These range from congenital defects such as cleft lip, pyloric stenosis, talipes equinovarus, and spina bifida (with incidences of from one to three or four per thousand) to diabetes, ischemic heart disease, schizophrenia, and rheumatoid arthritis (each of which has a morbidity at least an order of magnitude greater and collectively affects a sizable fraction of the population).

The precise nature of the familial element in these conditions is hard to decipher. Pedigree studies show clearly that none of them is inherited in the simple Mendelian fashion of single-gene disorders, whether "dominant" or "recessive," nor has evidence turned up of any simple, underlying metabolic defect such as we find in the enzymatic and similar pathologies. Nor, for that matter, are any of them associated with detectable chromosomal anomalies. Nonetheless, mathematical analysis of familial data (some details of which I shall recount here) has begotten some better-than-plausible hypotheses on the etiology of these medically very important diseases. These can be summed up as follows:

- The familial disorders result from the interaction of both heredity and environment (the latter term including, of course, the intrauterine conditions encountered by the fetus). The relative weight of the hereditary element varies, however, from one condition to another.
- They are multifactorial – meaning that each of them involves a sizable number of different hereditary and environmental influences.
- Their hereditary component is polygenic, reflecting the activity of many genes rather than one, resulting in a continuum of genetic predisposition to the disease in question.
- Their actual expression requires a strong genetic predisposition that pushes the individual beyond the "threshold of risk," at which point environmental influences will determine whether and, in some cases, to what extent the individual is affected.

Let us now examine these hypotheses in detail:

The existence of both genetic and environmental factors emerges most clearly from studies of identical twins. When we consider, for example, cases of congenital cleft lip (with or without cleft palate), we find that where one monozygotic twin is affected, the other will also be affected four times out of ten – about 400 times the "chance" expectation, i.e., the incidence in the general population.

Clearly, there are powerful genetic influences at work here, yet it is equally clear that genetics cannot be the whole story – otherwise we would consistently find both twins or neither affected. Evidently there must be differences in the intrauterine environment even of identical twins, though it is admittedly difficult to conceive of what these might be! The possibility that uterine conditions are *entirely* responsible, i.e., that no hereditary component is involved, is disposed of when we note that among fraternal twins both individuals are affected only 4% of the time, a figure that, though much greater than the chance expectation, is still no more than we find in two siblings born at different times. Similar contrasts in incidence between identical and fraternal twins (or siblings) have been found

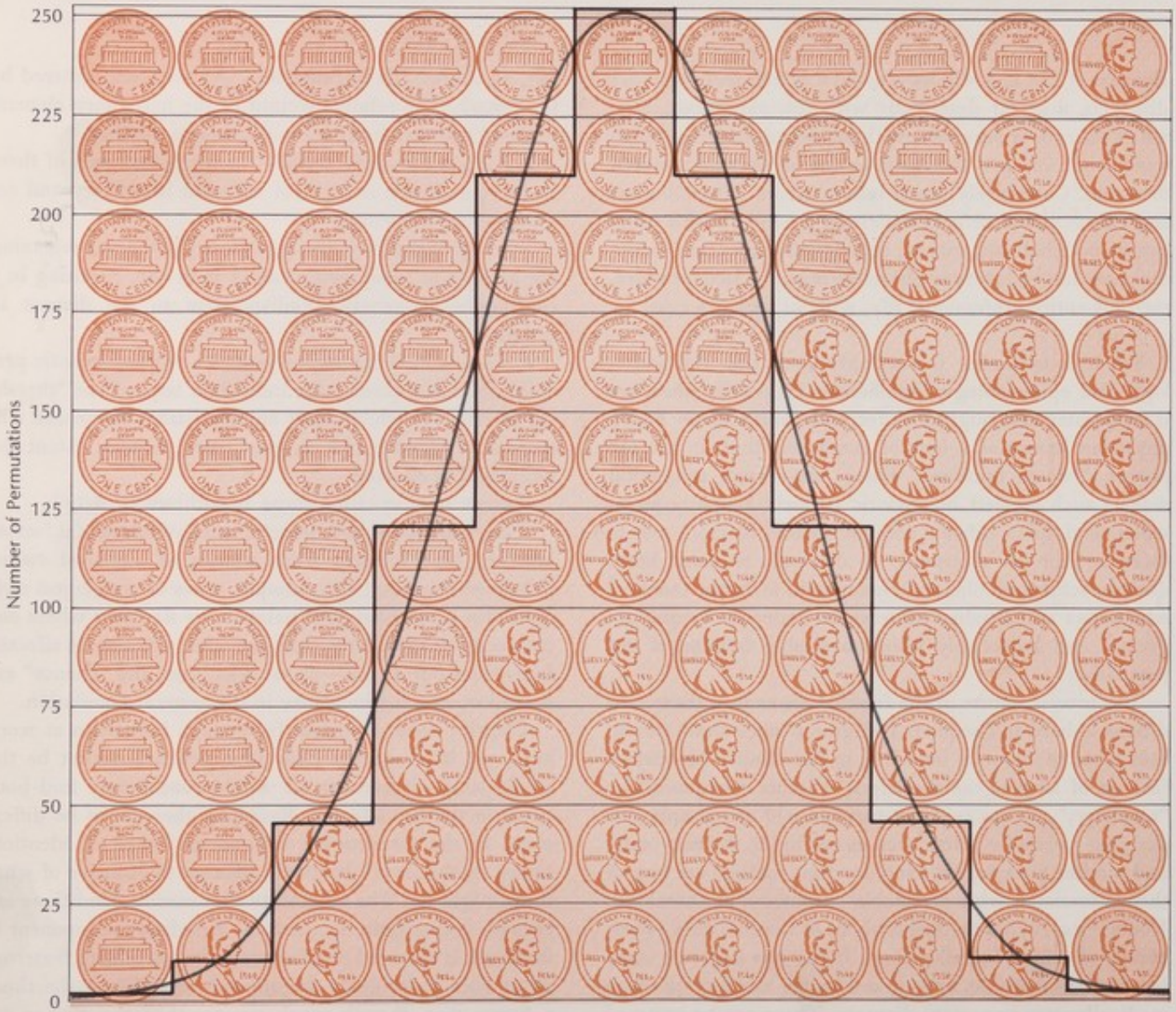
in congenital pyloric stenosis, schizophrenia, and a number of other putatively multifactorial conditions.

To understand the significance of the multifactorial and polygenic aspects of genetic diseases, we must first consider how these phenomena manifest themselves in the normal individual.

Many of the most basic and obvious parameters of the human animal are multifactorially determined. These include height, body build, basal metabolism, blood pressure, and the like. These traits are alike in having a continuous variation across a fairly wide spectrum and in manifesting a distribution described by the "normal" or Gaussian curve of variation, with the bulk of the population concentrated in the midrange and only a few individuals at the extremes. Indeed, the existence of a continuous, normal distribution in respect to any characteristic is the strongest sort of presumptive evidence of multifactorial causation. Why this is so can be seen by considering how the normal curve can be derived. The flipping of coins provides a good example and analogy.

A single flipped coin will obviously come up either heads or tails (we may ignore the remote probability that it will land on edge). This is equivalent to a unifactorial situation: only one variable is operative, and it is either wholly present or wholly absent. Genetically it is analogous to a condition determined wholly by a single gene, i.e., if environment plays no part in the gene's expression. As we know, there are few if any genetic traits that are quite so rigorously determined; perhaps the best examples are the common blood groupings. An individual will be A positive or negative; there is no intermediate condition.

Consider, now, the situation where we flip not one coin but six. The most probable result is obviously three heads and three tails (3H3T). There are, in fact, 20 different arrangements which will produce this result:



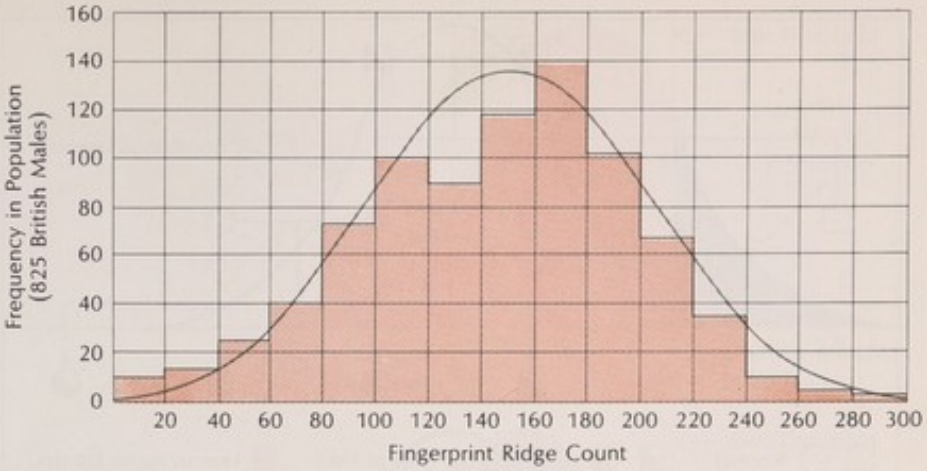
Coin-flipping probabilities provide a model of multifactorial causation. Bar graph showing likelihood of various heads-tails combinations when 10 coins are flipped approximates the bell-shaped "normal" probability curve; if more coins were flipped, the fit would be even closer. Put in medical terms, each "head" would correspond to a single, independent pathogenic factor.

HHHTTT, HTHTHT, HHTTHT, TTHTHH, and so on. If the proportion shifts to 4H2T or 2H4T, the number of possibilities drops to 15 and continues to drop as the proportion of heads to tails (or vice versa) increases, until we find only a single possible arrangement that will produce all heads or all tails. If, now, we graph the number of arrangements corresponding to each possibility, ranging from all heads to all tails, i.e., the probability that any one of them will occur, we will obtain a rough approximation of the normal distribution curve. It will, of course, be a "stepped" rather than a smooth curve, but one can easily see that by increasing our number of coins to 50 or 100 or 1,000 we can smooth the curve as much as we choose, obtaining any desired approximation to the "perfect" Gaussian curve.

Our multiple coin situation is, of course, a paradigm of multifactorial causation in nature. In relation to height, for instance, one can visualize a set of 50 or 100 coins, for each of which heads represents a factor tending to increase height while tails is a factor tending to decrease it, and so for blood pressure, intelligence, etc.

One should note that the term multifactorial does not in itself imply anything about the genetic or environmental balance of the factors. In theory, one can conceive of a multifactorial trait determined entirely by the environment, though no example occurs to me. Nor (again in theory) need a multifactorial trait be polygenic, though the two terms are sometimes used synonymously; its appearance might reflect the actions of one or two genes plus a dozen or two environmental factors. It is also possible to imagine a multifactorial trait that is *purely* polygenic, i.e., produced by the additive action of many genes, but without the intervention of environmental factors. Here, the theoretical possibility can be fleshed out with some actual cases.

The first of these was discovered by no less a personage than Gregor Mendel. In one of his experiments, a cross between purple-flowered and white-flowered beans produced mauve-flowering plants in the first generation and a range of colors from purple to white in the second, results that Mendel suggested must be due to the action of more than one gene pair. Subsequent work has turned up other examples, in both plants and animals.



Near-normal distribution of fingerprint ridge counts in unselected population provides powerful presumptive evidence that multifactorial causation is at work.

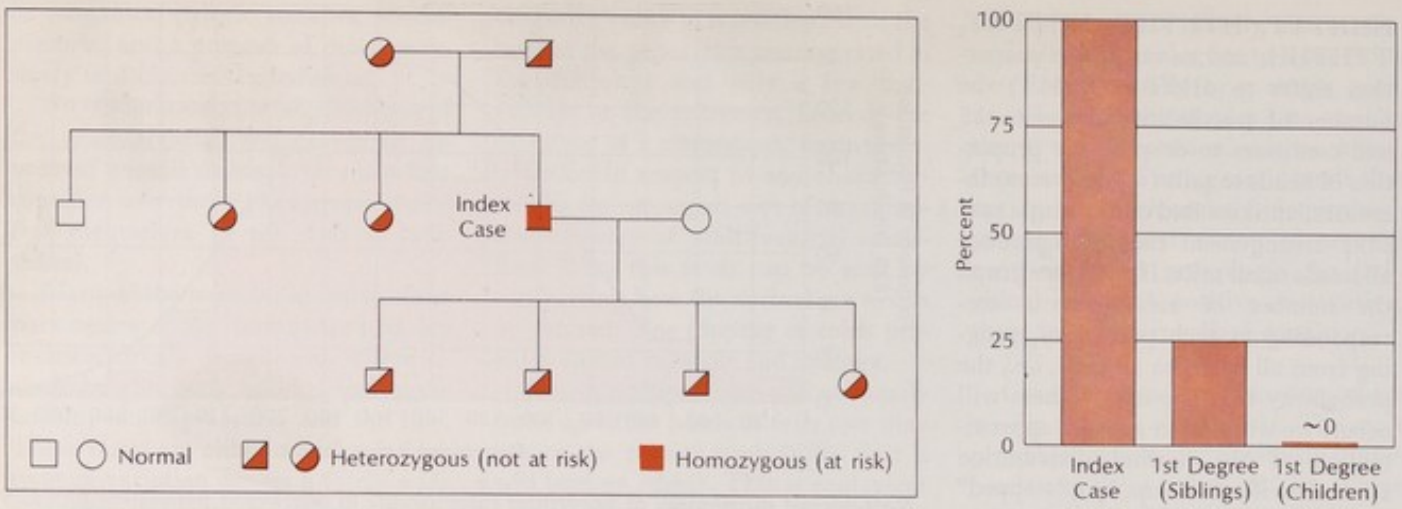
In man, perhaps the "purest" case of polygenic inheritance is that of fingerprint ridges. The variable here is the total number of ridges for the 10 fingers when counted according to certain specified rules — a quantity that is finally determined by the fourth month of fetal life, so that the influence of the environment is minimal. Its polygenic character is shown, first, by its normal distribution in the population and, second, by the fact that the degree of resemblance among relatives with different degrees of kinship corresponds very closely to the number of genes they have in common. Thus in identical twins, with precisely the same gene component, one would expect a correlation of 1.0; the actual figure (for 80 twin pairs) is 0.95 ± 0.01 (even here, it appears, environment plays a small role). For parent and child, the number of genes in common is, of course, precisely 50%; for fraternal twins and non-twin siblings, the average, though not necessarily the individual, resemblance is also 50%. The actual correlations, shown in the table on this page, are very nearly perfect within the mathematical limits of the data. Numerically these correspondences are also what one would expect in the case of a single dominant gene, but the existence of a continuous, normal distribution rules out that possibility.

Having filled in the background of multifactorial and polygenic causation, how do we apply these principles to the genesis of disease? As our example, let us revert to our initial case in point: cleft lip, with or without cleft palate. For reasons already cited, we assume that this condition results from a genetic predisposition com-

Genetics and Fingerprint Ridges

Relatives	*Observed Correlation*	*Theoretical Correlation*
Husband-wife	0.05 ± 0.07	0.00
Monozygotic twins	0.95 ± 0.07	1.00
Dizygotic twins	0.49 ± 0.08	0.50
Sib-sib	0.50 ± 0.04	0.50
Parent-child	0.48 ± 0.03	0.50

Kinship patterns of ridge counts indicate that the causative factors are almost completely genetic. Observed correlations closely approximate the theoretical correlations that can be calculated from the number of genes shared by each pair of individuals.

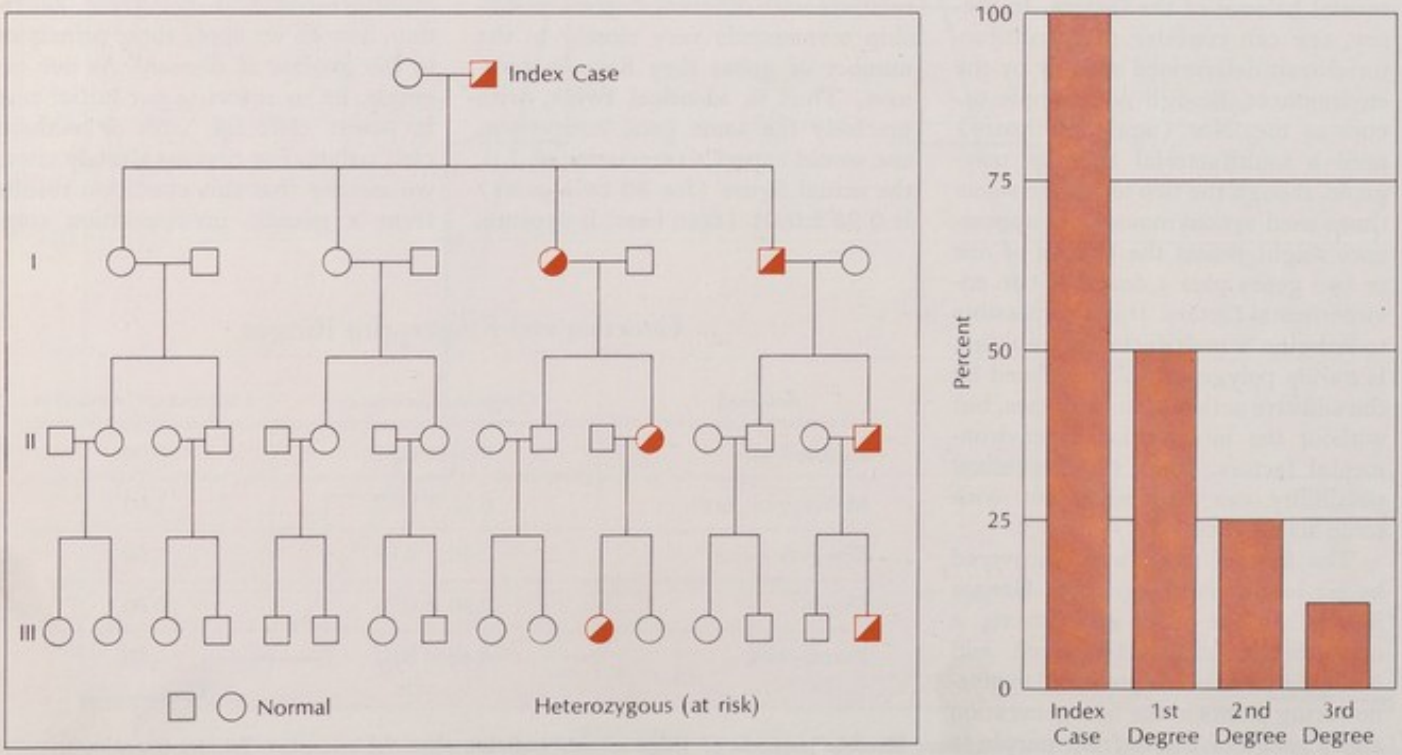


Single recessive-gene model fails to predict incidence patterns in conditions such as cleft lip. Though both siblings and children of an index case are first-degree relatives, sharing the same number of genes with him, the theoretical incidence among the children would approximate zero. In fact, the disease is almost equally common in both groups of first-degree relatives.

bined with unfavorable environmental (in this case, intrauterine) factors. In principle, the predisposition could be monogenic, resulting from the operation of a single dominant or recessive gene, rather than polygenic. However, when we examine the incidence patterns of cleft lip, the case for monogenesis seems improbable.

If we assume that the genetic predisposition is furnished by a single dominant gene, we would expect that the risk, and therefore the incidence, of cleft lip in a particular group would vary directly with the number of genes which that group shares with an affected individual. Thus first-degree relatives (parents, children, siblings), who average 50% common genes with the affected individual, should show half the incidence of monozygotic twins, with 100% common genes; second-degree relatives (aunts, uncles, nephews, nieces) would have one quarter the twin incidence, and so on. In fact, we find that first-degree relatives have only one tenth the incidence of identical twins (4% as against 40%), while second-degree relatives have less than one fiftieth the twin incidence (0.7%).

If we assume a single recessive gene, we are no better off. Here, the affected homozygous individual must inherit the "risk" gene from both heterozygous parents, a combination that occurs with a probability of 25%, and the same 25%, of course, would apply to the siblings of the affected



Single dominant-gene model is equally nonpredictive. Here, in theory, the incidence should be directly proportional to the number of genes shared with the index case; in fact, all degrees of kinship show a far smaller incidence than the model predicts.

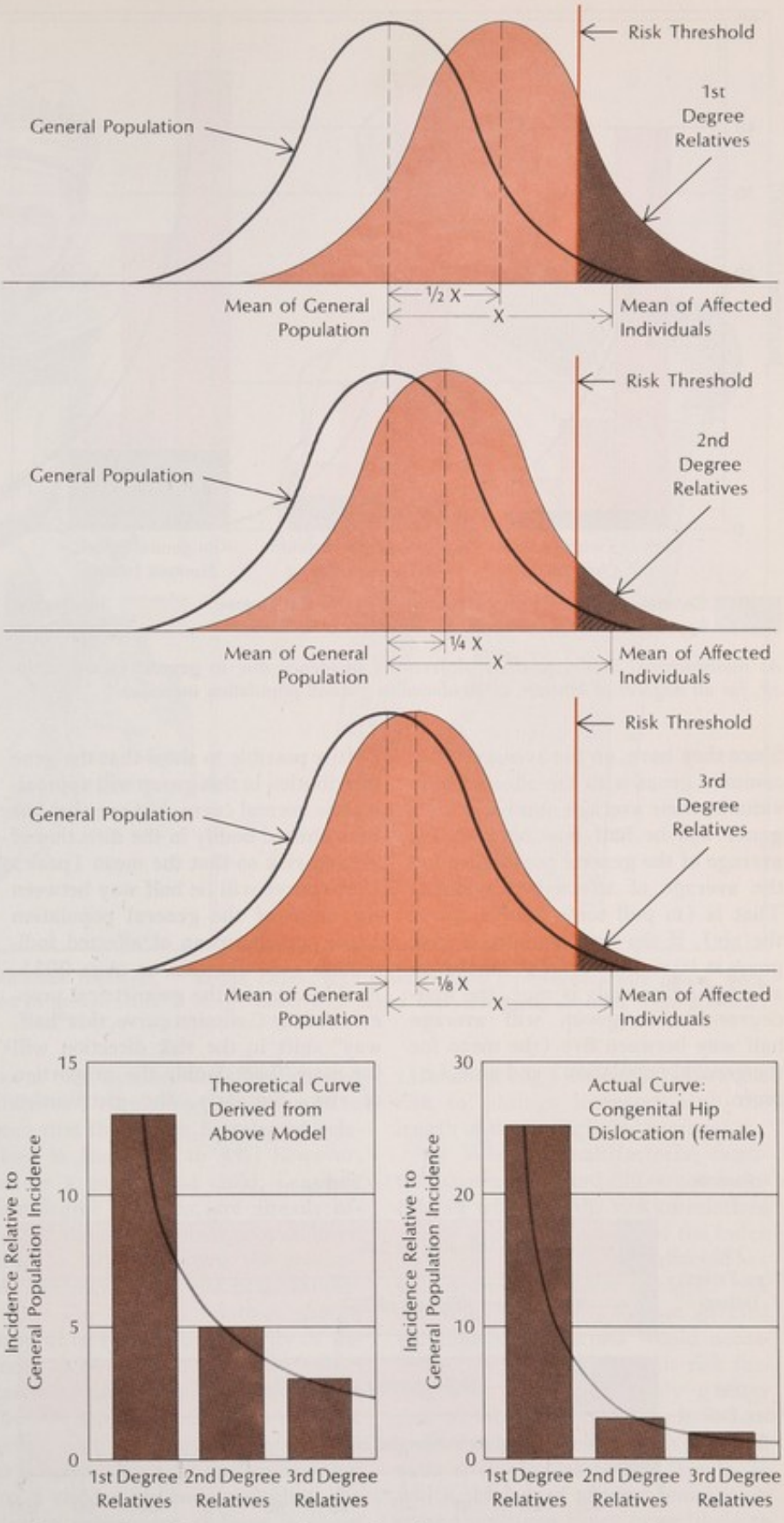
individual: one in four will be at risk.

When we consider children of an affected individual, however, we have quite a different case. They will certainly inherit the defective gene from the affected parent, but barring consanguinity between mother and father, will have very little chance of getting a second defective gene — and with it the risk of cleft lip. It would follow, then, that the risk and incidence of the condition in children of affected individuals should be very much smaller than in siblings of those same individuals. Actually, it is almost the same: in an American study, 4.3% as against 4.6%; in a London group, 3.0% vs 3.2%.

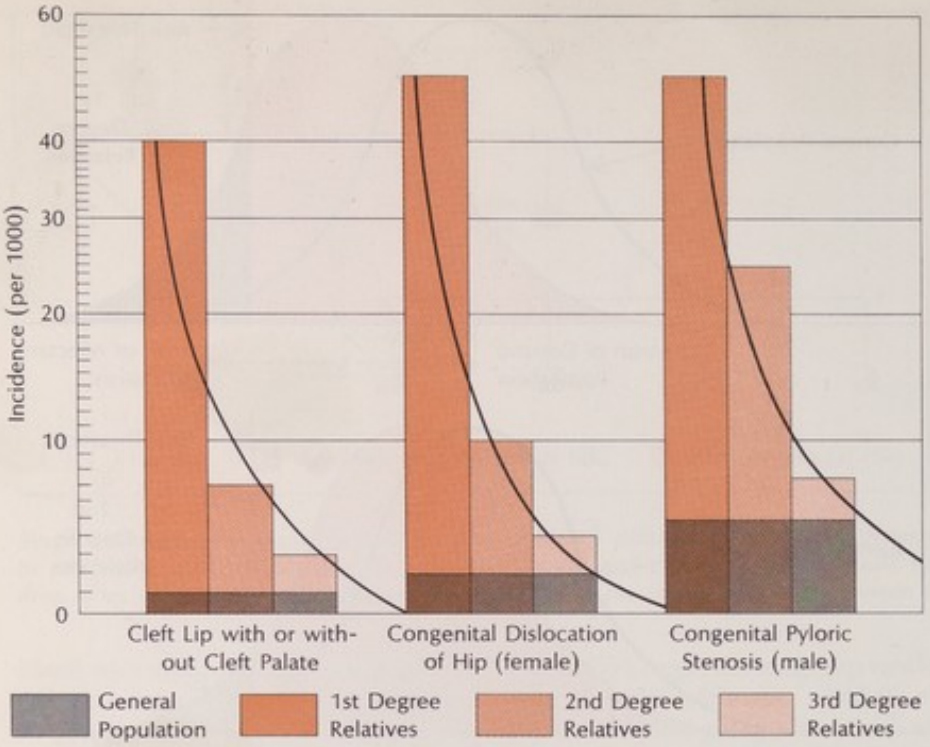
Since the consequences of neither single-gene model seem to square even moderately well with the facts, let us see what the polygenic model will do. Here we assume a sizable number of different "risk" genes randomly distributed throughout the population. We need not assume that any single such gene is "defective," and indeed there are reasons (whose discussion is beyond the scope of this chapter) for suspecting that many of them are not. In combination, however, they generate a risk in the individual whose gene complement happens to include all, or nearly all, of them (the "threshold" effect).

Since we have assumed a good number of independent risk genes, their distribution in the population will form some sort of approximation to the normal curve: most people will possess a moderate number of the genes (say, from one third to two thirds of the total), a few will have almost none of them, and an equal few will have almost all of them — enough to push them over the risk threshold. How many of these actually develop cleft lip will depend, first, on how many risk genes they have and, second, on how favorable or unfavorable is the intrauterine environment. Only some, perhaps a minority, of the individuals with a high proportion of risk genes will develop cleft lip, but all the individuals with cleft lip will possess a high proportion of the genes.

Let us now consider the first-degree relatives of the latter group. On the average, they will have more risk genes than the general population, though less than the affected group.



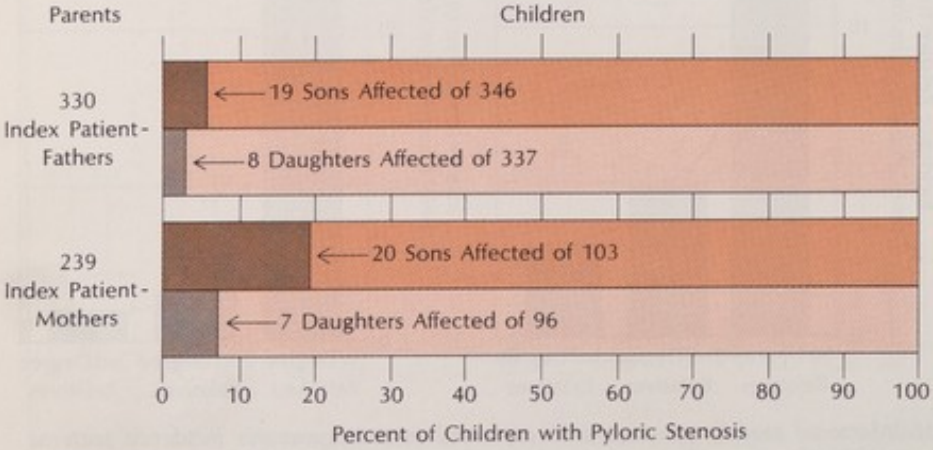
Multifactorial model, constructed as explained in text, generates incidence patterns resembling those actually found in congenital hip dislocation and many other putatively multifactorial conditions. "Normal" curves show size of risk group in relatives of differing consanguinity (dark color areas) compared with that in general population (hatched areas); actual incidence patterns should, and do, reflect these proportions.



As multifactorial model predicts, differential incidence due to genetic factors drops off, for all degrees of kinship, as incidence in general population increases.

Since they have, on the average, 50% common genes with the affected individuals, their average number of risk genes will be half way between the average of the general population and the average of affected individuals. That is (to pull some figures out of the air), if the total number of risk genes is 10 and the average number in affected individuals is nine, the first-degree-relative group will average half way between five (the mean for the general population) and nine, i.e., seven.

It is possible to show that the gene distribution in this group will approximate a normal curve, but one that has been shifted bodily in the direction of greater risk so that the mean (peak) of the curve will lie half way between the mean of the general population curve and the mean of affected individuals (see diagram on page 203). And because of the geometrical properties of the Gaussian curve, this "halfway" shift in the risk direction will far more than double the proportion at risk. Similarly, the distribution curve of second-degree relatives will be shifted only a quarter of the way toward the affected mean, third-degree relatives only an eighth, and so on. And when we employ these curves to estimate the number of individuals at risk in each group (which is, of course, proportional to the incidence in each group) we find precisely the sort of figures that occur in actual studies of cleft lip: a very sharp drop in incidence between first- and second-degree relatives and a much less precipitous drop between second and third. By the time we reach fourth-degree relatives, the difference from the general population approaches marginal significance. Similar patterns of inheritance have been found in talipes equinovarus, congenital dislocation of the hip, and congenital pyloric stenosis.



Multifactorial model postulates that where one sex is less frequently affected (here, females), its affected individuals will transmit a higher risk to their offspring. The incidence figures in congenital pyloric stenosis conform to this pattern.

All this highly involved ratiocination may seem a rather shaky foundation on which to erect so elaborate an etiological hypothesis. It happens, however, that the hypothesis also generates several predictions of incidence that can be independently tested.

1) The influence of relationship on incidence should vary inversely with the overall (the general population) incidence of the disease. That is, the higher the overall incidence, the lower the genetic threshold can be presumed to be, resulting in a smaller span between the genetic mean of the general population and the mean of affected individuals. And the smaller the span, the less the average genetic difference between relatives of affected individuals and the general population, meaning that the increase in risk with increasingly close relationship to an index case will be relatively less, though absolutely greater. And, in fact, we find that in cleft lip, with an overall incidence of .001, the risks for third-degree, second-degree, and first-degree relatives are respectively 3, 7, and 40 times that of the general population, while in congenital dislocation of the hip (in females) the overall incidence is .002 and the comparable risk factors for relatives are 2, 3, and 25 times the general incidence. Similarly, in congenital pyloric stenosis (males only), with overall incidence of .005, the risk factors are 1½, 5, and 10 times.

Postpartum treatment of infant is an environmental factor in congenital hip dislocation. Condition is common in groups such as the Lapps, who swaddle their infants, very rare when "back sling" holds legs flexed and abducted, as with some Chinese.

2) In contrast to the usual situation with single-gene diseases, the risk of polygenic conditions should vary from one affected family to another, since it is not a matter of the presence or absence of a single pathogenic gene but of what one might call the family's average component of risk genes. And the more individuals in the family who are affected, the higher, presumptively, is the risk component. And we find, for example, that in anencephaly and spina bifida, where two children in a family are already affected, the incidence in subsequent children is twice that in families where only one previous child was affected. (We may contrast this with a single-gene condition such as cystic fibrosis, where the risk to a child remains one in four, regardless of whether the mother has previously borne one, two, or even three affected children.) Again, in cleft lip we have seen that the risk in first-degree relatives is from 3% to 4%; however, where a parent and child, i.e., *two* first-degree relatives, are already affected, the risk to subsequent children rises to 10%, reflecting the presumably greater genetic risk in the family.

3) The risk to relatives should vary directly with the severity of the index case; that is, the more severe cases can be presumed to possess, on the average, a higher number of risk genes, with a consequent greater risk to relatives. For cleft lip, we find that for a child with unilateral cleft, the risk to subsequent siblings is about 2.5%, while with bilateral cleft lip and palate, the sibling risk rises to 6%. Similar findings have been reported for aganglionic megacolon (Hirschsprung's disease), where the risk to siblings increases with the length of the aganglionic segment.

4) Where a multifactorial condition shows a marked difference in incidence with sex, the risk to relatives should depend on the sex of the index case: specifically, the less-affected sex will presumably have a higher risk threshold (as in 1), so that affected individuals of that sex would necessarily have a greater genetic risk factor, with a proportionately greater risk to relatives. An example is congenital pyloric stenosis, with a normal ratio of five males to one female, in which children of affected females are more than three times more likely to be born with the malformation than children of affected males.

Thus, for the cited congenital malformations, at least, the predictions

generated by the polygenic-multifactorial theory are consistently confirmed by the facts, and the case for the theory seems better than plausible. For the common diseases of adult life, matters are rather less clear-cut. In contrast with the congenital malformations, which are clearly defined and unambiguously diagnosable, the diagnosis, and in some cases even the definition, of such diseases as schizophrenia, ischemic heart disease, and diabetes mellitus is considerably less than certain. Moreover, again in contrast with the congenital malformations, these adult diseases manifest or fail to manifest themselves in a given population over a span of many years, so that incidence figures must be corrected for age, i.e., for the "unexpired risk" of developing the disease.

Nonetheless, such figures as are available seem to be consistent with the polygenic-multifactorial hypothesis — and not merely its general features, but some of the postulated predictions that I have already cited. In particular, the influence of consanguinity on incidence consistently varies inversely with the overall incidence of the disease in question.

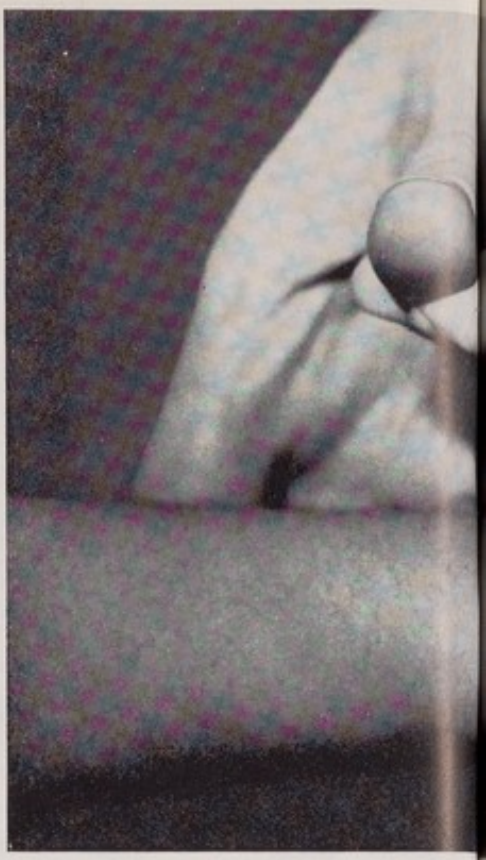
Ischemic heart disease is, of course, a very common condition indeed, and several studies have indicated that in first-degree relatives of such heart patients the risk of death from cardiac ischemia or its complications (e.g., coronary thrombosis) is only about 2.5 times that of the general population. However, when we consider only cases of early-onset ischemia (before age 55 for men, age 60 for women), we find a moderate, general incidence (.015 for men, .010 for women) and a sixfold increase in deaths among first-degree relatives.

Rheumatoid arthritis, with somewhat greater incidence (.020 to .030), has a fourfold increased risk in first-degree relatives; the much rarer ankylosing spondylitis (incidence .002 in men, .0004 in women) shows a greatly increased risk of 35, 10, and 3 times in male first-, second-, and third-degree relatives respectively. Significantly, these latter figures are close to those we find in female congenital dislocation of the hip, which has about the same general incidence.

The existence of a genetic factor in diabetes mellitus has been known for years, but its nature has been debated, with a number of investigators proposing a single-gene mechanism, either dominant or recessive, much modified by environment. However, both these possibilities can be all but eliminated on much the same grounds as we ruled them out for cleft lip. Thus, for example, the incidence in parents of index cases is as high as in their siblings, when the latter are corrected for age, a pattern not consistent with recessive inheritance. We also find consistency with the polygenic postulate of differing risk among families, measured by the number of individuals affected. Thus, the risk to siblings of a diabetic is doubled when one parent is also affected and increases even further if both parents are affected. It is also significant, perhaps, that early-onset diabetes (before age 30), which may be a somewhat different disease and is certainly a rarer one (incidence about .002), shows a much higher (24 times) increased risk to siblings.

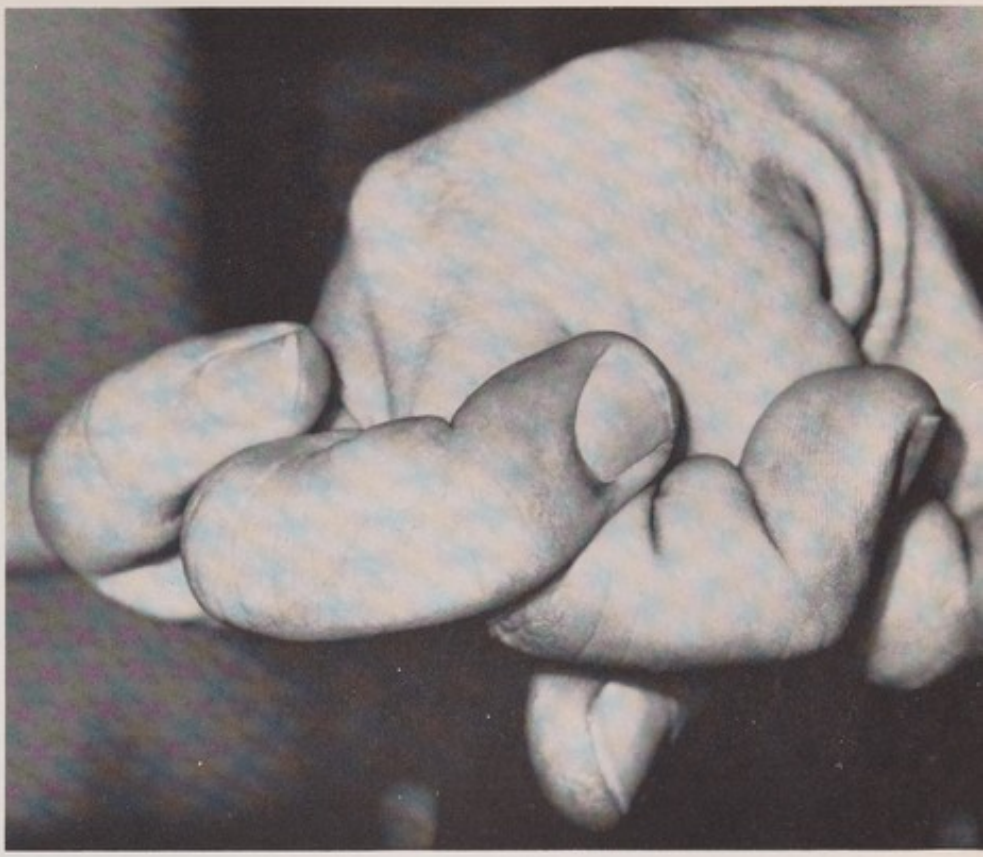
In the case of schizophrenia, not merely the nature but the very existence of a genetic factor has been a matter of bitter dispute. The presence of some sort of familial element has been noted since the 19th century, and a number of studies (notably, those of F. J. Kallmann) long ago revealed some of the incidence patterns that I have already cited. The risk of developing the disease varies directly with the genetic relationship to an index case – close to 50% for monozygotic twins, 10% to 15% for first-degree relatives, perhaps 4% for second-degree relatives. Significantly, too, the risk to dizygotic twins was found to be no greater than that for other first-degree relatives. Nonetheless, the existence of a genetic predisposition was attacked by many psychiatrists, notably those of the analytic persuasion (in some cases so heatedly that one is tempted to postulate some form of Freudian resistance!). They insisted that the incidence figures could be explained equally well by purely environmental causation, i.e., that the closer the social contact with a schizophrenic, the greater the risk of developing the disease.

Whatever the merits of this view, it was persuasive enough to block wide-spread acceptance of the importance of a largely genetic predisposition for a generation. Recent studies, however, have disposed of the environmentalists' objections by showing that in children of schizophrenics the risk is only slightly diminished when they are separated from their parents at birth and reared by normal adoptive parents. That adoption per se is not responsible is shown by control studies of adopted children without schizophrenic heredity. Another of our postulates is satisfied by the finding that where two first-degree relatives are affected, the risk about doubles in both cases.



Familial joint laxity, perhaps determined by a single gene, has been identified as

Similar findings have turned up in the other major form of severe mental illness, manic-depressive psychosis, though the genes involved appear to be quite different from those predisposing to schizophrenia, i.e., an increased risk of one disease does not imply increased risk of the other. There are indications, however, that the genetic risk of schizophrenia is closely related to that for the rather looser diagnostic category of schizoid illness, giving further weight to the belief that both these conditions are



one of at least two genetic factors in the pathogenesis of congenital hip dislocation; it was found in a third of affected females and nearly three quarters of the males. Another genetic factor, acetabular dysplasia, may be polygenic.

different manifestations of the same underlying pathology.

Persuasive though all these statistics may be, one would still like to know more about the genetic – and environmental – factors in putatively multifactorial diseases before embracing this hypothesis to the exclusion of all others. In most instances, we can as yet do little more than speculate. In the case of ischemic heart disease, to be sure, we know a great deal about predisposing environmental factors – cigarette smoking, diet, hypertension, obesity, and the like. About the genetic factors, however, we are almost totally in the dark, though one might hazard a guess that these have to do with differences in fatty-acid metabolism and perhaps also in the structure and/or metabolism of the vascular intima. Similarly, in schizophrenia the environmental component presumably involves certain kinds and degrees of psychic stresses, while the genetic component might consist of imbalances and/or abnormalities in the neurohumors, but the precise nature of either remains obscure.

It happens, however, that in one putatively multifactorial abnormality – congenital dislocation of the hip – we can now identify with confidence at least some of the genetic and environmental components concerned and even reproduce their action in experimental animals.

One of the environmental factors concerns the treatment of the infant postpartum. An early clue came from studies of Lapps and certain American Indians. In both these racially unrelated groups the incidence of congenital hip dislocation is markedly higher than in Europeans, and in both the young infant is swaddled "papoose style" – bound tightly to a cradle board with the hips extended and adducted. Conversely, hip dislocation is reported to be low in tropical populations where the infant's movements are not restricted, and especially in populations where the infant is carried in a back sling with hips flexed and abducted – this being precisely the hip position used in therapeutic correction of congenital dislocation.

A second environmental factor is breech birth. This, too, has been known for some time. A 1953 study found that 16% of congenital dislocation patients were breech births, as against 2% to 3% breech presentations in the general population. A more recent study, which I carried out with John A. Wilkinson, obtained the same figure and also established that another 9% had undergone therapeutic version close to term. These two groups add up to a quarter of the affected group, and it is likely that a sizable proportion of the remainder spent some time in the breech position in utero but underwent spontaneous version before birth. (The length of time spent in the breech position by either group would, of course, provide an additional variable.)

As to the genetic factors, one of them appears to be acetabular dysplasia, with incomplete covering of the head of the femur. The depth of the acetabulum can be estimated quite easily and accurately in children of a certain age by measuring Hilgenreiner's angle: the greater the angle, the shallower the acetabulum. Measurement of the angle in normal populations reveals a normal distribution,

strongly suggesting that it is itself multifactorially determined and probably, though not certainly, polygenic as well. The genetics of acetabular dysplasia are hard to investigate, since Hilgenreiner's angle cannot be measured in adults. Recently, however, Dr. R. Wynne-Davies in Edinburgh, using a measure of acetabular dysplasia more suited to adults, has shown that children with acetabular dysplasia are likely to have one or both parents with the same condition.

Whatever the genetics of acetabular dysplasia, however, it is clearly associated with congenital dislocation. In 25% of our unilateral cases, Hilgenreiner's angle on the unaffected side was more than two standard deviations above the norm, while in another 20% it was between one and two standard deviations.

A second and perhaps more pervasive genetic factor is familial generalized joint laxity. Hyperlaxity of the joints is present in certain syndromes, such as Ehlers-Danlos, Marfan's, and osteogenesis imperfecta. More often, it occurs alone and behaves in families as a dominant trait. In extreme form (possibly corresponding to the homozygous state) it is incapacitating, causing not only congenital dislocation of the hip but also recurrent dislocation of other joints such as the shoulder, clavicle, and patella. In milder (heterozygous?) form it is fairly common, being found in about 6% of normal schoolchildren. In our congenital dislocation group, however, it was present in nearly three quarters of the boys (24/33) and about a third of the girls (18/55). This finding has recently been confirmed by the Edinburgh workers.

The lower incidence of this "predisposing factor" in girls creates a seeming problem, since, as already mentioned, congenital hip dislocation is much more common in girls than in boys, the actual ratio being at least five to one. There must be some additional, and important, factor in females.

One possibility would be a difference in the development of the acetabulum, and in fact studies have shown that this socket is somewhat shallower in female infants. This difference, however, is much too small to account for the radical difference in the incidence of congenital dislocation. A more plausible factor is a temporary, generalized joint laxity in females in the fetal and early neonatal periods.

It has long been known that such laxity develops in women (and other female mammals) during the later stages of pregnancy; the process affects the pelvic ligaments, those of the anatomically adjacent hip joints, and perhaps others as well. Its evolutionary significance is obvious: it imparts an extra elasticity to the pelvis in preparation for the physical demands of childbirth.

In part, the laxity is hormonal; as early as 1926, pelvic relaxation was induced in both mature and immature female guinea pigs by injecting them with estrogen followed by progesterone. In addition, however, it depends on a specific, "sex-limited" property of female ligaments, for male animals do not respond in the same way to the hormone injections. It seems likely, therefore, that the special female propensity to congenital hip dislocation is due to the special properties of female ligamentous tissue stimulated by the endocrines of the female fetus. (The latter, under the stimulus of maternal gonadotropins, secrete both estrone and progesterone during the last trimester of pregnancy.)

In babies, the significance of this sex difference is difficult to demonstrate clinically; however, my associate, Mr. Wilkinson, has reproduced it experimentally in rabbits. If the leg of a 6-to-8-week-old rabbit is splinted in a position of hip flexion, knee extension, and lateral rotation of the leg (the rabbit equivalent of the human frank breech malposition) no dislocation occurs in either sex. If, however, the animals are first treated with estrone, followed by progesterone, the females – but not the males – develop generalized joint laxity and an atraumatic dislocation of the hip very similar to human congenital dislocation.

Thus in congenital dislocation of the hip, we have what seems to be a convincing paradigm of multifactorial disease: two genetic factors – one of them (acetabular dysgenesis) apparently itself multifactorial and probably polygenic, the other (familial joint laxity) possibly monogenic – modified by two environmental factors – one of them prenatal and, again, multifactorial (the existence, timing, and duration of breech positioning), the other postnatal (the "swaddling factor"). In the latter connection, it is significant that in two European populations studied, the incidence of congenital dislocation *at birth* was something like four times that in the general population. Presumably three quarters of these infants recover spontaneously – but would not do so in a culture where swaddling was the rule. If, then, multifactorial disease is in most cases still a hypothesis, in congenital dislocation of the hip it seems to be a demonstrable fact.

Doubtless a great deal of research will be needed before multifactorial causation can be demonstrated in such detail for ischemic heart disease, diabetes, schizophrenia, and the like. Such research, however, can be expected to provide very important dividends, practical as well as theoretical. A clearer understanding of the genetic risk factors would be of obvious importance in genetic counseling. Already, for example, we can estimate that if a man and a woman are both early-onset diabetics, there is about one chance in eight that they would produce a similarly affected child – not an overwhelming risk, certainly, but one worth taking into consideration when deciding whether or not to marry or to have a child.

Long-term studies of individuals presumed to be at high genetic risk might well enable us to identify at least some of the environmental factors that turn risk into disease and, to the extent that these are controllable, to reduce markedly the incidence of multifactorial disorders even in the genetically susceptible.

Finally, one might expect, or at least hope, that ultimately the presumed genetic abnormalities can be associated with biochemical abnormalities (as has already been done in many simpler genetic disorders) – at first in affected individuals, then in their unaffected but genetically high-risk kin, and eventually in the population at large. Developments along this line would do much to achieve what must always be a major goal of medicine: to single out the one individual requiring special care, treatment, or prophylaxis from the ninety and nine who, possessing what was once called a naturally strong constitution, will – barring the accidents of infection or trauma – thrive on little or no medical attention.

Section Six

Clinical Applications of Genetic Knowledge

20

The Nosology of Genetic Disease

VICTOR A. McKUSICK
Johns Hopkins University

At first glance, the subject of nosology – the classification of diseases – might seem to be the essence of theoretical, academic medicine. To the practicing clinician, the question of whether a particular pathologic constellation is filed in the proper conceptual pigeonhole would appear as little relevant as the subtle taxonomic distinction between *Pinus strobus* and *Pinus sylvestris* would be to a practicing lumberman or carpenter. A moment's thought, however, will show that nosology – in medicine, if not in carpentry – is an eminently practical study. One need only recall, for example, that hardly more than a century ago such phenomena as jaundice, dropsy, and anemia were still often classified, and treated, as "diseases" in their own right, rather than as symptoms consistent with many distinct pathologic entities. Accurate nosology is obviously essential in the area of prognosis and therapy (e.g., pernicious anemia versus irondeficiency anemia) , and no less so in the field of prophylaxis, whether this concerns recurrence in a given individual or among the population at large (e.g., jaundice due to Rh disease, to alcoholic cirrhosis, or to infectious hepatitis).

The same considerations obviously apply to the nosology of genetic disease. Here, however, questions of prognosis in the usual sense and treatment tend to be secondary. In the area of prophylaxis, however – meaning genetic counseling – it can be of key importance.

A brief anecdote will make this point clear. It concerns a couple whose first child suffered from what the family doctor diagnosed as the severe type of genetic dwarfism called achondroplasia. Seeking counsel on whether to have additional children, they wrote to a geneticist quoting this diagnosis and noting that the condition had not hitherto occurred in the family of either parent. Rather injudiciously, I would say, he undertook to advise them without himself seeing the patient; since achondroplasia is inherited as a dominant, he wrote, and since it had not been previously observed in the lineage, the child in question could be presumed to represent a new mutation – so that the chance of a recurrence was close to zero.

Subsequent events, unfortunately, made clear that the child did not have achondroplasia but a similar condition called diastrophic dwarfism, which is inherited not as a dominant but as an autosomal recessive, meaning that the risk of recurrence was not zero but 25%. In fact, it did recur – twice.

Beyond even such practical considerations, however, is the fact that accurate nosology is essential to the understanding of disease. If we are to talk intelligently about a disease (or any other subject) we need to be sure that we are actually talking about the same thing, that we are discussing the same pathologic entity. In genetic terms, this amounts to saying that we are considering a phenotype resulting from a specific and unitary genetic etiology. Certain chromosomal defects provide clear examples: Down's syndrome, the *cri-du-chat* syndrome, and some others are consistently associated with specific chromosomal anomalies (see Hirschhorn, Chapter One) and identified unequivocally by karyotyping. A more difficult problem is posed by the so-called point mutations, whose nosology will be discussed in the remainder of this article. Here, present techniques do not permit us to identify the genetic defect directly, so that its existence must be inferred on various grounds – sometimes inadequate ones – of which I shall have much more to say in a moment.

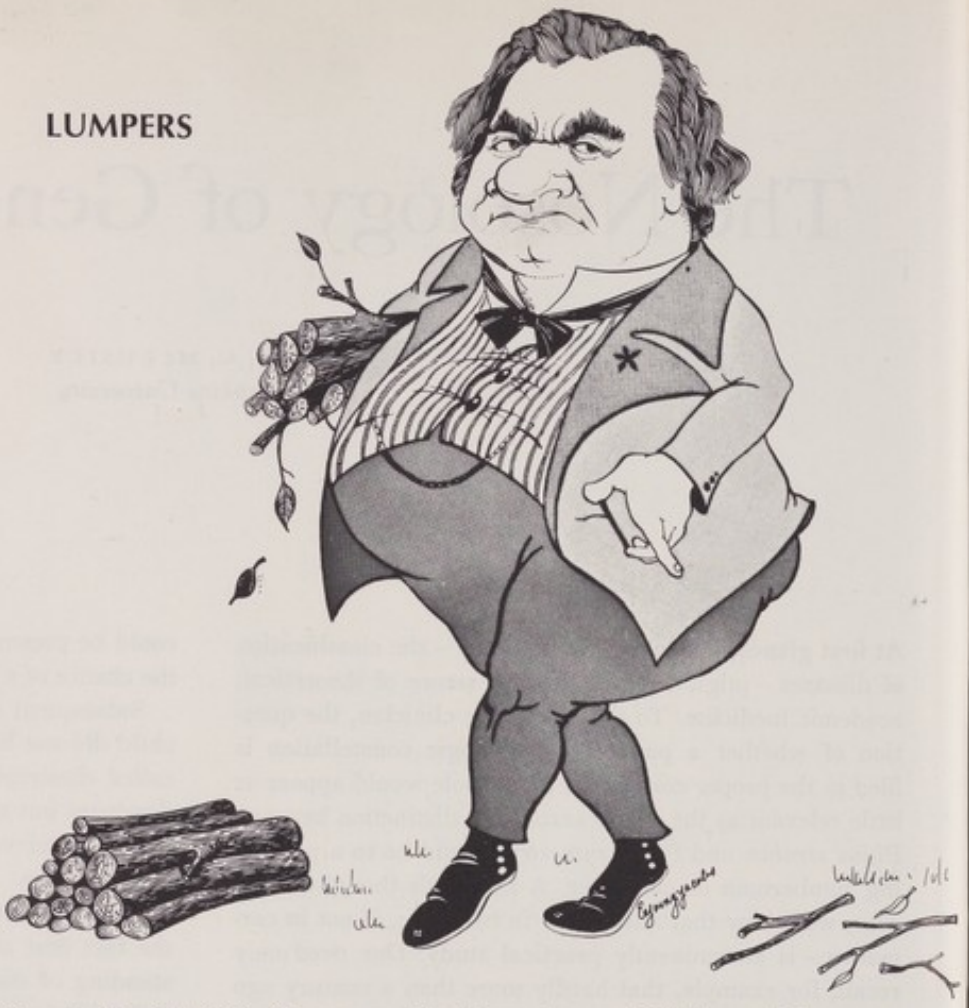
Before going into this, however, I should mention two of the fundamental phenomena that the genetic nosologist must cope with. The first is pleiotropism – meaning the multiple effects of a single mutant gene. The great majority of known gene defects do not affect a single tissue or organ and often not even a single organ system. On the one hand, this is an asset, since it facilitates the separation of similar phenotypes into distinct entities by "the company they keep." But it can also be a source of confusion; since there are more mutations than organs, it follows that different mutations will not infrequently affect

the same combination of organs, and often enough in at least roughly the same way. A further source of confusion arises because a given gene may affect different organs (or systems) to different degrees in different individuals. Thus in the Marfan syndrome, for example, in which major effects are on the eyes, the aorta, and the skeleton, any one of the three may be severely damaged with little or no involvement of the remaining two. Where two similar syndromes show this sort of individual variability, the physician may be hard put to determine whether he is dealing with an "atypical" case of syndrome A or a "typical" case of syndrome B.

The second basic nosologic problem is genetic heterogeneity, which in a sense is the conceptual opposite of pleiotropism; "one from many" as against "many from one." Mutations at quite different loci can give rise to clinically very similar (in some cases, clinically indistinguishable) disorders. In some instances, not merely the genetic locus but the mode of inheritance will differ; for example, spastic paraplegia, retinitis pigmentosa, and Charcot-Marie-Tooth peroneal muscular atrophy are inherited as autosomal dominants in some families, as autosomal recessives in others, and as X-linked recessives in still others. Here, correct nosology is of obvious importance in genetic counseling, but even where the mode of inheritance does not differ, the counseling implications may be important, as we shall see.

The phenomena of pleiotropism and genetic heterogeneity are reflected in two tendencies among nosologists: "lumping" and "splitting." In the early days of medical genetics the splitters tended to dominate, often as an inadvertent result of specialization in medicine. Thus the dermatologic symptoms of pseudoxanthoma elasticum were described as early as 1881, and angioid streaks in the fundus oculi were reported only a few years later. It was not until 1929, however, that these two phenomena were lumped together (appropriately), through collaboration between a dermatologist and an ophthalmologist, as pleiotropic manifestations of a single syndrome.

Even as the lumpers have been correcting much of the "oversplitting" of



"Lumping" and "splitting," here personified by two famous U.S. historical figures, are perennial and complementary tendencies in nosology, as well as in zoological and

earlier years, a new wave of better-founded splitting has greatly expanded the list of discrete genetic entities. This reflects not inadvertence but the development of subtler methods for distinguishing phenotypically similar but genetically heterogeneous conditions; as such, it can be expected to continue.

In seeking to establish an accurate classification of genetic disease, the nosologist relies on three categories of evidence: clinical, genetic, and biochemical.

Clinical evidence – differences in the phenotype – is by all odds the least reliable category of evidence. These differences can too easily lead the nosologist into circular reasoning: having selected a particular group of cases because of certain symptomatic similarities, one assumes that the differences among them represent the "normal" range of variation for the entity in question. In the absence of other evidence, however, we may actually be dealing with an "entity" more specific but hardly less heterogeneous than the old-time physician's "dropsy" or "jaundice." Alternatively, one may assume that the differences observed reflect the existence of two (or more) distinct entities, when in fact they stem merely from normal variability, in particular the influence on the mutant gene of the rest of the individual's genetic complement. That this can be considerable has been demonstrated experimentally in mouse-breeding experiments. In human beings, a given mutation can express itself with noticeable variability even in sibs, where (on the average) 50% of the genetic complement is identical. Thus in two Marfan sibs, one may show ectopia lentis while the other has normal eyes.

However, when other evidence of genetic heterogeneity (or homogeneity) is present, phenotypic differences

botanical taxonomy. The author's selection of the dramatis personae (Stephen A. Douglas and Abraham Lincoln) suggests his own strong bias in favor of splitting.

begin to attain their proper perspective, and, not infrequently, unnoticed differences emerge to serve as truly diagnostic phenotypic features. Thus in gargoylism (mucopolysaccharidosis), for example, evidence of the existence of an X-linked form (Hunter's syndrome) as well as an autosomal recessive form (Hurler's syndrome) led to the finding that in the former corneal clouding was uniformly absent. Subsequently, phenotypic and other evidence (e.g., that of Neufeld, see Chapter 14) has split mucopolysaccharidosis into no less than six distinct entities (often called mucopolysaccharidosis I, II, III, etc.).

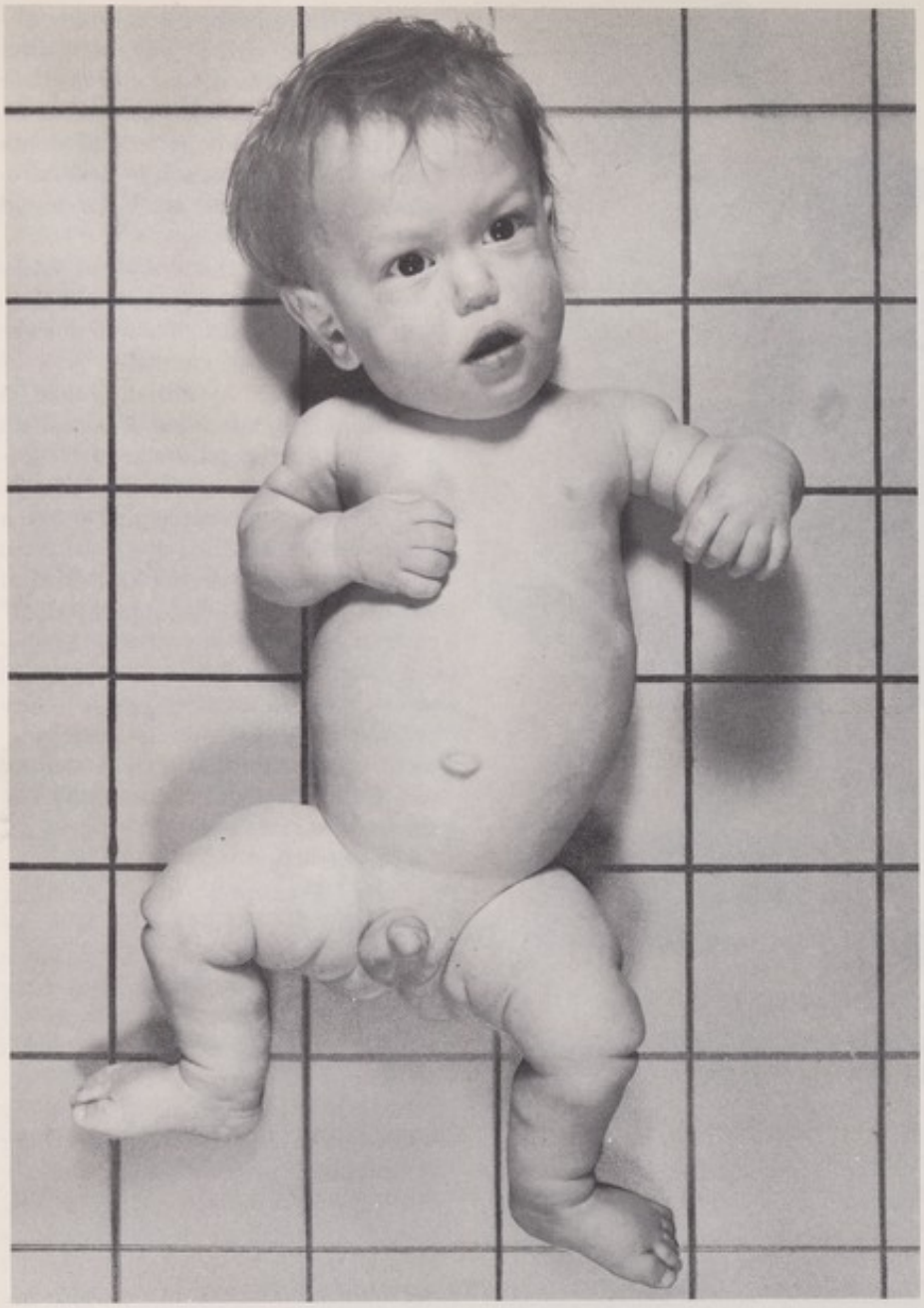
Perhaps the commonest form of genetic evidence comes from studies of large, extended families (kindreds) or "demes," the latter being isolated, inbred populations that are in effect extended kindreds. The demes in particular have played a most important role in genetic studies of all sorts; examples include groups in the more remote cantons of Switzerland, on islands off the Yugoslav coast, and, in North America, endogamous religious communities such as the Hutterites of Canada and the Amish of our own country. Twenty, or even a dozen, cases of a presumed entity occurring within a kindred or in a deme, where precisely the same (ordinarily rare) gene can be presumed to be operating, are far more valuable in defining the phenotype in question than twice the number of sporadic cases. In the latter situation, indeed, even the bare fact of a genetic etiology may be in doubt.

Studies of the Amish, for example, have distinguished two types of dwarfism: the Ellis-van Creveld syndrome (a condition previously described on the basis of occasional cases) and cartilage-hair hypoplasia, a "new" entity that has only more recently been recognized in the non-Amish. Both were formerly lumped with achondroplasia, though they differ markedly from true achondroplasia both in phenotype and in mode of inheritance; they are autosomal recessive rather than autosomal dominant. True, or "classic," achondroplasia, in fact, has not been found among the Amish.

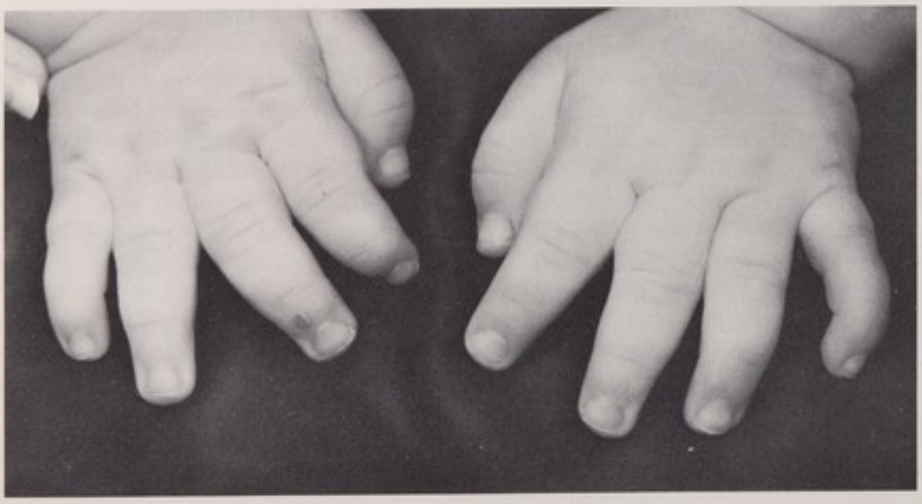
At times, a "clustering" of cases can lead to the identification of a hitherto occult kinship. Thus a former associate of mine studied a type of amyloid neuropathy initially found in 11 seemingly unrelated kindreds in Maryland and nearby states. Genealogic sleuthing then revealed that in fact all 11 were descended from a single couple that had migrated from Germany during the 1700's. Studies of the newly defined "superkindred" could then define the disease phenotype sharply enough to distinguish it clearly from a similar genetic entity previously unearthed in Portugal – while establishing its virtual identity to a form of amyloid neuropathy discovered in an Indiana kindred of Swiss extraction.

In discussing the nosologic relevance of consanguinity, whether established (kindred) or putative (deme), we have of course already moved into our second category of evidence, the genetic. The fact that a phenotypically similar condition recurs among related or presumably related individuals is obviously strong presumptive evidence that we are dealing with a discrete genetic entity, though naturally not by itself conclusive; as we all know, syphilis can occur familially but is in no sense genetic. An even more *outré* simulation of genetic disease was the severe neurologic condition called kuru, described some years ago among certain tribes in New Guinea. This initially appeared to show all the characteristics of a genetic entity, and was classified as an autosomal recessive. Subsequently, however, it was shown to be caused by infection with a slow virus, whose familial transmission was produced by (among other things) religious cannibalism of dead family members!

There are several other kinds of genetic evidence on which the nosologist relies. I have already mentioned differences in mode of inheritance – dominant versus autosomal recessive versus X-linked recessive – as a



"Classic" achondroplasia, inherited as a mendelian dominant, is a common form of dwarfism known for thousands of years. It is easily diagnosed by bulging calvarium, scooped-out bridge of nose, and "trident hand" caused by gaps between fingers.



means of distinguishing phenotypically similar diseases. However, this particular type of evidence is intrinsically a very coarse screen, since there are only three common modes of inheritance (only a very few X-linked dominants have been described), while experience has shown that not a few conditions exist in considerably more than three distinct forms. Differences in mode of inheritance may prove a degree of genetic heterogeneity but the absence of such differences proves nothing whatever. Thus of the mucopolysaccharidoses, only one has a distinctive (X-linked) mode of inheritance; all of the others are transmitted as autosomal recessives, though they can be clearly distinguished on other grounds. I might add that it is perfectly possible in theory (and already some evidence of this is developing) that the X-linked mucopolysaccharidosis might eventually turn out to include two (or more) distinct genetic entities.

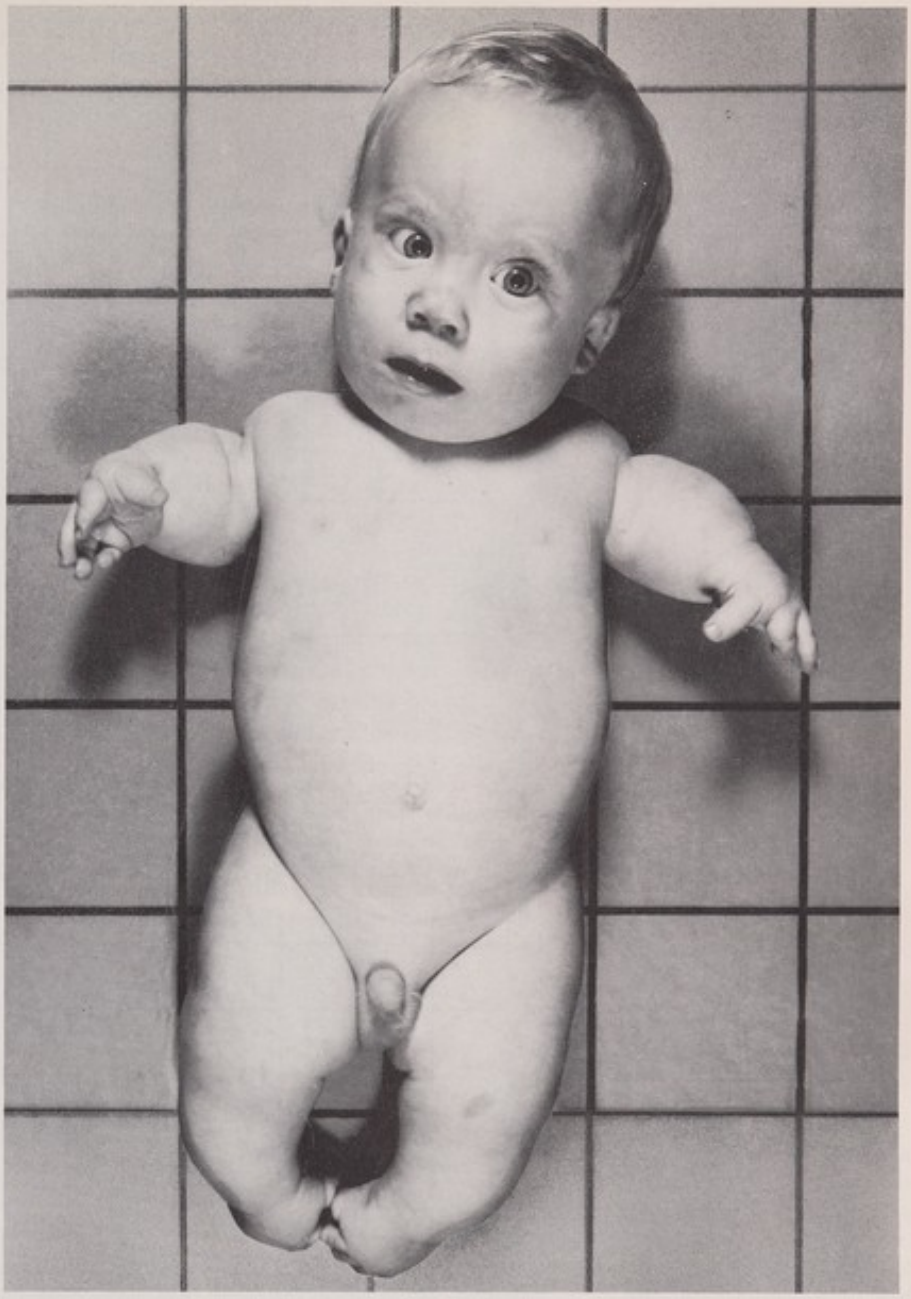
It is perhaps worth noting in passing that where similar conditions are found to be inherited in all three modes (e.g., spastic paraplegia) we generally find that the recessive form is the most severe and the dominant is the least severe, while the X-linked form occupies an intermediate position. This is of course what one would expect. A dominant disorder severe enough to produce early death, early and grave disablement, and/or serious impairment of the reproductive capacity will eliminate itself quite rapidly from the gene pool, while a recessive gene may persist for generations in the harmless heterozygous form. An X-linked gene, which is in effect dominant in men but recessive in women, will fall between the other two.

Two other types of genetic evidence, though relatively uncommon, have proved useful in distinguishing closely similar diseases. The first of these concerns genetic linkage. As is well known, genes located close together on the same chromosome will tend to stay together during the "crossover" of genetic material in meiosis during gametogenesis, meaning that they will tend to be inherited together. By contrast, genes located on different chromosomes, or widely separated on the same chromosome, will be inherited independently, mean-

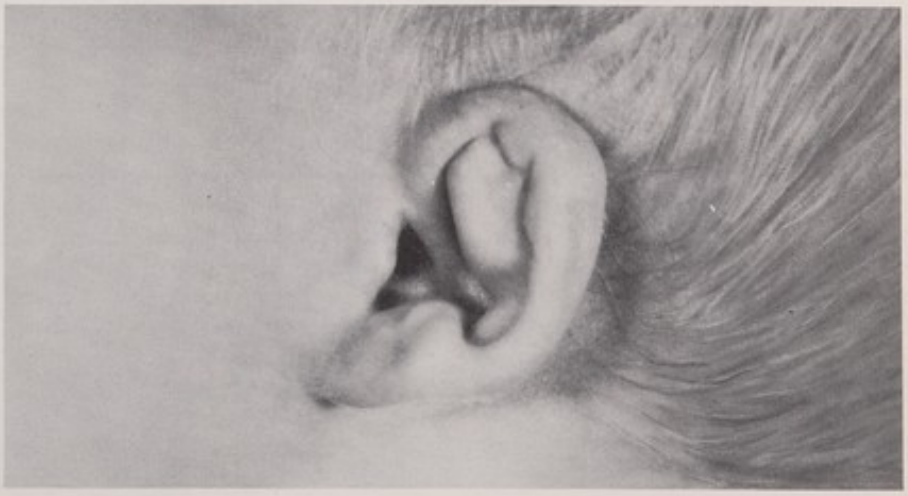
ing that after meiosis they will have only about a 50-50 chance of turning up in the same gamete. Linkage has been demonstrated exhaustively in the geneticist's favorite animal, the drosophila, and to a much lesser extent in human beings. And occasionally it happens that, through intensive studies of pedigrees, a particular disease can be linked to some other known, and clearly distinguishable, genetic disease or trait. Thus one form of elliptocytosis, for example, has been shown to be closely linked with the Rh blood group gene, while another form, which is as yet phenotypically indistinguishable, shows no such proximity. X-linked hemophilia had earlier been shown through other evidence to occur in two forms (A and B); subsequently, it has been found that the A form is closely linked to the genetic locus that determines the enzyme glucose-6-phosphate dehydrogenase and to those that transmit color blindness. The gene for B form hemophilia lies at a distance from these loci.

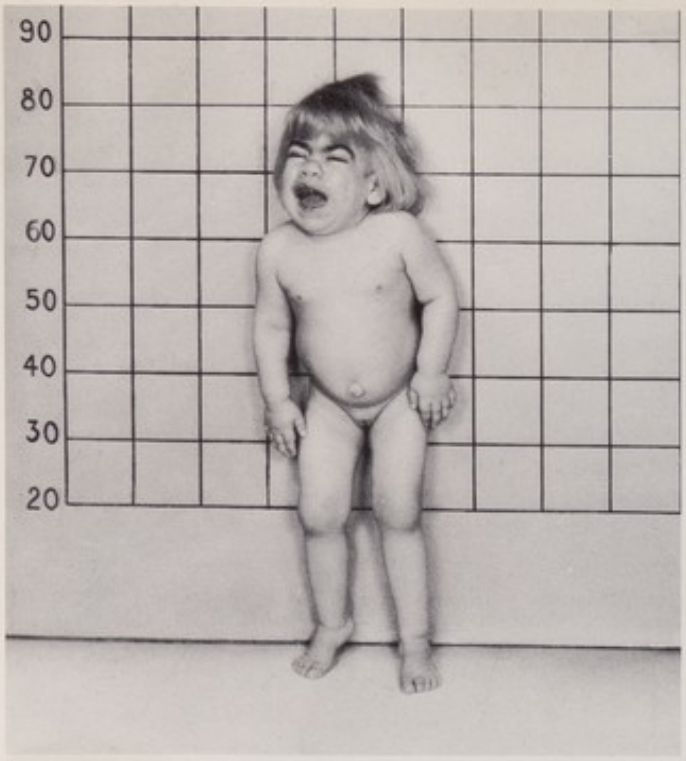
Still another category of genetic evidence comes from matings between individuals suffering from apparently identical autosomal recessive disorders. Such individuals must, of course, be homozygous for the defective gene in question, from which it follows that all their children would be homozygous as well. In fact, however, such matings occasionally produce not merely one normal child (which could conceivably be ascribed to the rather mysterious phenomenon known as reduced gene penetrance) but several normals – and no abnormals. The only explanation seems to be that while both parents are homozygous at a locus determining the disease, it is a different locus in the two parents, meaning that all the children will be heterozygous at both loci ("doubly heterozygous") but normal.

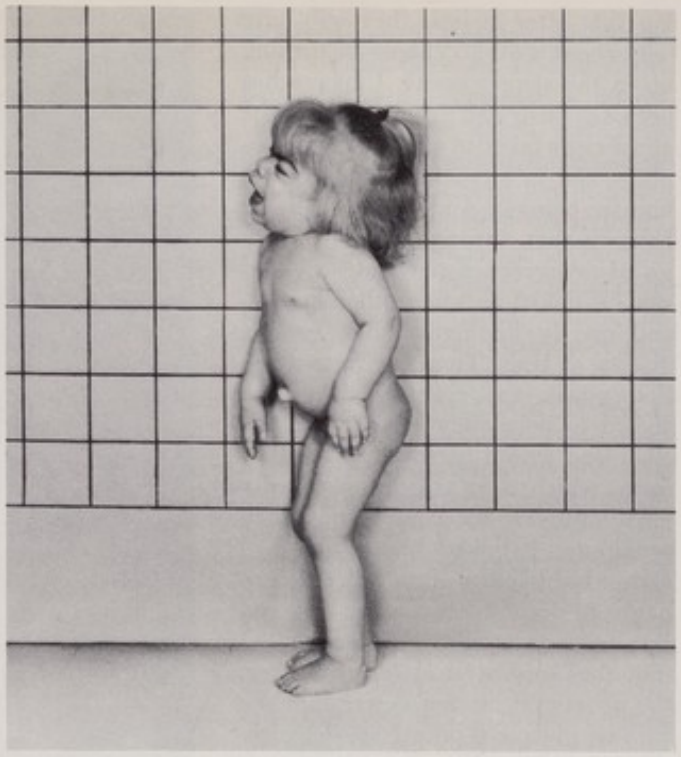
A good example concerns the condition known as deaf-mutism or more often nowadays as congenital deafness (specialists in the field consider "mutism" to suggest an overly pessimistic prognosis). This is a not uncommon condition, and one in which, for obvious reasons, sufferers are likely to be thrown into one another's company and therefore to marry. Yet there are many cases in which such marriages produce nothing but normal children.



Diastrophic dwarfism, inherited as a mendelian recessive, is one of several conditions once lumped with achondroplasia but in fact readily distinguished from it – for example, by clubbed feet, "hitchhiker's thumb," and deformation of the external ear.

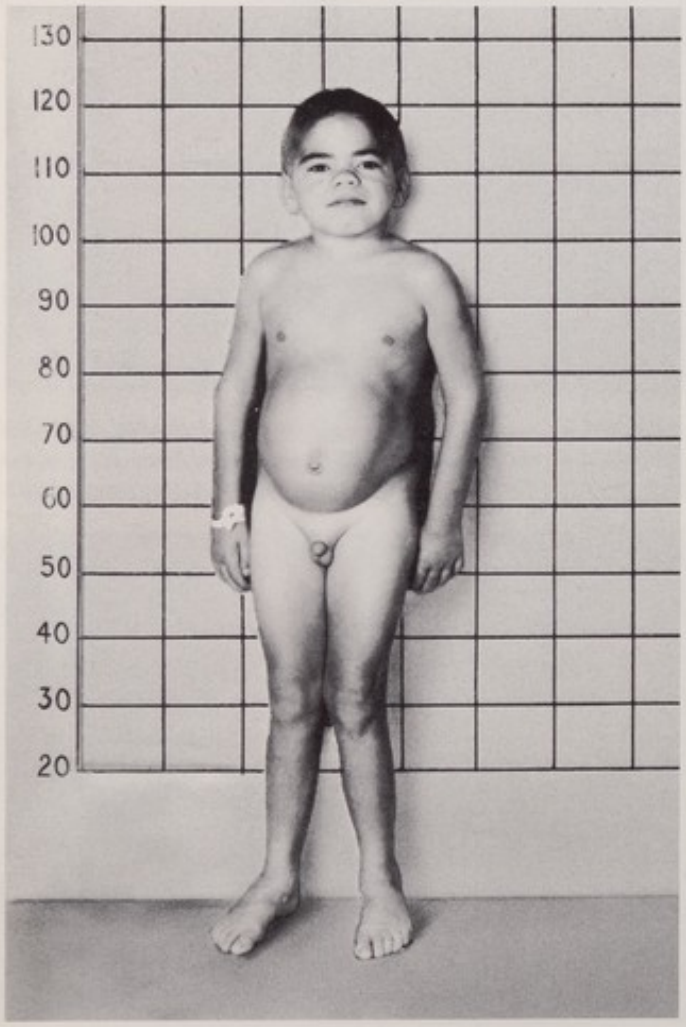


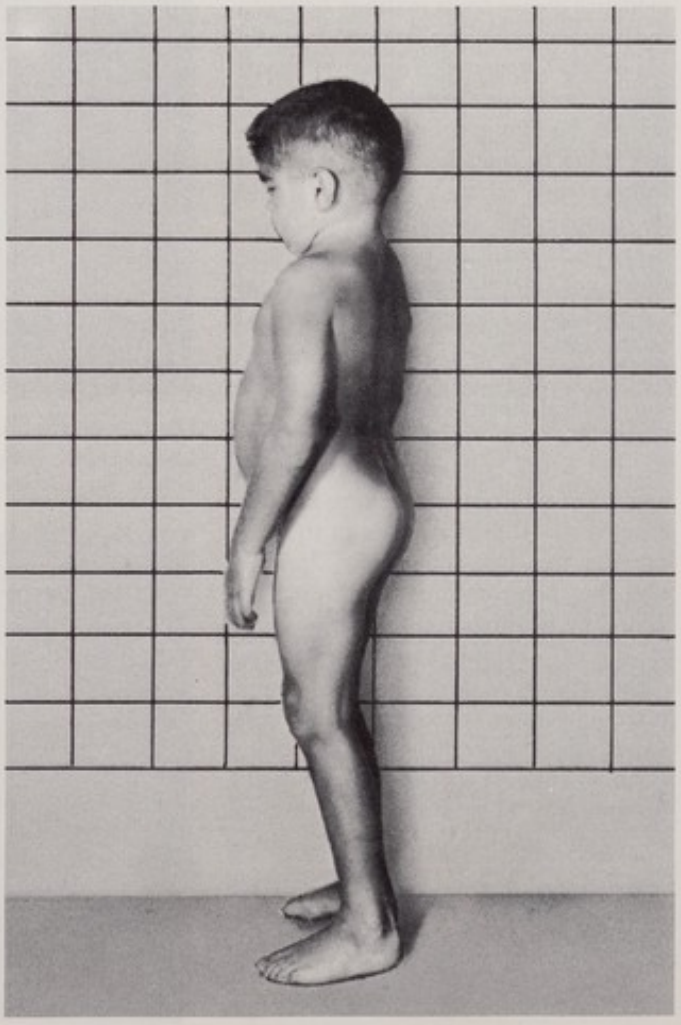




Hurler's syndrome (above) and Hunter's syndrome (below), two of the mucopolysaccharidoses, are not readily distinguishable by appearance, since both show such signs as coarse facies and prominent abdomen (the two individuals shown are not related).

Demonstrated differences in their biochemistry and mode of inheritance (Hurler's is an autosomal recessive, Hunter's an X-linked recessive) ultimately helped single out subtler clinical signs by which accurate differential diagnosis can usually be made.





It now appears that there may be as many as 35 phenotypically identical but genetically distinct forms of congenital deafness, which probably makes this condition the current champion illustration of genetic heterogeneity.

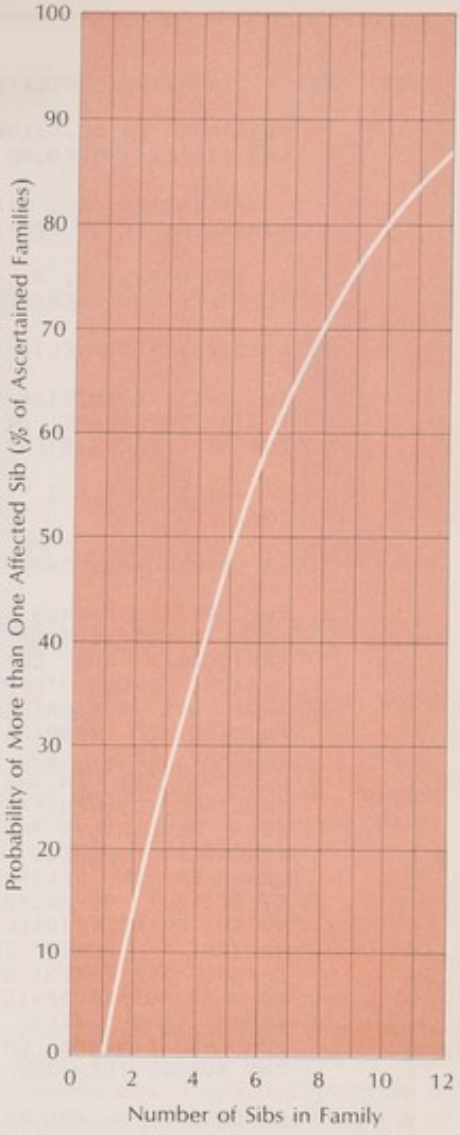
Barring such rare genetic evidence as that provided by linkage or particular matings, by all odds the most definitive means of establishing genetic heterogeneity is through biochemical studies. In general, one would expect to find a distinctive biochemical pattern in each genetic disease entity; in practice, however, the singling out of significant differences among the body's multitudinous chemical processes has too often proved an insurmountable task. This is particularly true when the biochemical changes observed are far removed from the primary effect of the gene mutation. Yet even here biochemistry can sometimes help us to make significant and useful distinctions. Thus qualitative differences in the urinary excretion of mucopolysaccharides (MPS) helped distinguish mucopolysaccharidoses III and IV from types I and II, and from one another – though the same technique provides only an uncertain distinction between I and II. In Chapter 14 Dr. Neufeld carries on the story, showing how "cross-correction" studies and, more recently, enzymatic characterizations have confirmed (but in some cases upset) the earlier nosology. The cross-correction phenomenon in the mucopolysaccharidoses is similar to that which first established the distinction between hemophilias A and B; blood from either type will correct the clotting defect in the other.

In the case of the mucopolysaccharidoses, we seem to be approaching the "ideal" situation in genetic disease: the relation of a specific phenotype to a particular defect in protein synthesis, since it is here that we arrive at the primary action of genes, mutant or normal. A great variety of genetic diseases has now been firmly related to specific defects in hemoglobin synthesis (see Conley and Charache, Chapter Five), to the inactivity or absence of particular enzymes (see Knox, Chapter Six, Hsia and Holtzman, Chapter 23, and Seegmiller, Chapter 10), and, less firmly, to similar deficiencies of the enzyme-like transport proteins (see Scriver, Chapter Eight). Moreover, we can now clearly distinguish the multiple forms of, say, glycogen storage disease or hereditary nonspherocytic hemolytic anemia on the basis of the particular enzyme that is nonfunctional.

Even where a specific enzyme deficiency has been pinpointed as the source of a given disease, there may remain a degree of residual genetic heterogeneity, in the sense that the enzyme in question may be shown, through molecular studies, to possess different structural defects in different lineages. Subtle differences of this sort have been demonstrated in two types of hereditary nonspherocytic hemolytic anemia, respectively due to deficiency of glucose-6-phosphate dehydrogenase and of pyruvate kinase, and in galactosemia; they are suspected or have been tentatively demonstrated in several other conditions.

In many instances, these different abnormal forms of the enzyme are determined by alleles, i.e., genes at one and the same genetic locus. Referred to as allelic disorders, they may have a strikingly different clinical picture as in the Hurler and Scheie syndromes which are thought to be allelic (see Neufeld, Chapter 14). On the other hand, the clinical picture may be identical even though different changes in the same enzyme are present. Mutations of the beta hemoglobin chain (see Conley and Charache, Chapter Five) illustrate the wide diversity of phenotypes that can result from change in the same gene product: cyanosis, polycythemia, or anemia of several different types. Conversely, several different mutations can lead to the same phenotype, e.g., cyanosis. The group of genes leading to change in one polypeptide chain is referred to as an allelic series.

In merely quantitative terms, the progress of genetic nosology can be shown by the steady increase in the number of identified genetic entities. From something over 400 in 1958 it has risen to more than 1,000 in 1973, plus about an equal number of suspected but not proved entities. The increase reflects not merely the discovery of hitherto unknown entities but the splitting of known ones according to the criteria already mentioned. It is worth noting that even these figures somewhat understate the diversity of genetic disease, since they somewhat arbitrarily lump together all defects of a given gene locus, e.g., all beta-chain hemoglobin defects.



Multiple incidence of a disease within a sibship is an important clue to a possible recessive genetic etiology but is unlikely to occur in present-day small families. This fact gives a special importance to studies of such groups as the Amish and Hutterites, in which not only inbreeding but also large families are very common.

The current tabulation of genetic entities contains one interesting anomaly: whether one considers only "established" or also "suspected" entities, there are distinctly more dominant than autosomal recessive conditions. This is in sharp contrast to other species that have been closely studied, such as the laboratory mouse. The reason, of course, is that laboratory mice are closely inbred and, moreover, have larger "families," in which a particular recessive anomaly is

300 22240 DIAPHRAGM, UNILATERAL AGENESIS OF

PASSARGE ET AL. (1968) REPORTED UNILATERAL AGENESIS OF THE DIAPHRAGM IN A BROTHER AND SISTER AND FOUND FOUR REPORTS OF MULTIPLE AFFECTED SIBS IN THE LITERATURE.

PASSARGE, E., HALSEY, H. AND GERMAN, J.* UNILATERAL AGENESIS OF THE DIAPHRAGM. HUMANGENETIK 5* 226-230, 1968.

TEN KATE, L. P. AND ANDERS, G. J. P. A.* UNILATERAL AGENESIS OF THE DIAPHRAGM. (LETTER) HUMANGENETIK 8* 366-367, 1970.

22250 DIASTEMATOMYELIA

IN THIS CONDITION THE SPINAL CORD IS DIVIDED LONGITUDINALLY IN THE ANTERO-POSTERIOR PLANE BY A FIBROUS OR BONY STRUCTURE. THE CASES ARE USUALLY ISOLATED BUT AFFECTED SISTERS WERE REPORTED BY KAPSALAKIS (1964).

KAPSALAKIS, Z.* DIASTEMATOMYELIA IN TWO SISTERS. J. NEUROSURG. 21* 66-67, 1964.

*22260 DIASTROPHIC DWARFISM

THE PATIENTS SHOW SCOLIOSIS, A FORM OF CLUBBED FOOT BILATERALLY, MALFORMED PINNAE WITH CALCIFICATION OF THE CARTILAGE, PREMATURE CALCIFICATION OF THE COSTAL CARTILAGES AND CLEFT PALATE IN SOME CASES. PARTICULARLY CHARACTERISTIC IS THE 'HITCH HIKER' THUMB DUE TO DEFORMITY OF THE FIRST METATARSAL. THE TERM DIASTROPHIC WAS BORROWED BY LAMY AND MAROTEAUX (1960) FROM GEOLOGY* DIASTROPHISM IS THE PROCESS OF BENDING OF THE EARTH'S CRUST BY WHICH MOUNTAINS, CONTINENTS, OCEAN BASINS, ETC., ARE FORMED. CASES HAVE BEEN DESCRIBED UNDER MANY DIFFERENT DESIGNATIONS IN THE PAST. SEE THE CASE DESCRIBED BY MAU (1958) IN HIS SECTION ON 'MULTIPLE CONGENITAL MALFORMATIONS AND CONTRACTURES.' THESE CASES HAVE FREQUENTLY BEEN PLACED IN THE WASTEBASKET OF ARTHROGRYPOSIS MULTIPLEX CONGENITA IN HOSPITAL DIAGNOSTIC FILES. MANY CASES OF SO-CALLED ACHONDROPLASIA WITH CLUBFOOT ARE EXAMPLES OF DIASTROPHIC DWARFISM (E.G. KITE, 1964). THE FOOT DEFORMITY IS RELATIVELY REFRACTORY TO SURGICAL TREATMENT. LANGER (1967) REFERS TO AN ENTITY WHICH PHENOTYPICALLY IS A MILD FORM OF DIASTROPHIC DWARFISM AS 'DIASTROPHIC VARIANT.' BONEY CHANGES ARE QUALITATIVELY SIMILAR BUT LESS SEVERE. SOFT TISSUE CHANGES ARE ABSENT OR MILD AND THE CLUBFOOT IS NOT AS RESISTANT TO TREATMENT AS IN REGULAR DIASTROPHIC DWARFISM. CONSANGUINITY WAS NOTED IN THE REPORTS OF TAYBI (1963) AND JAGER AND REFIOR (1969). KNOWN TO ME ARE TWO AFFECTED WOMEN EACH OF WHOM HAD A NORMAL CHILD DELIVERED BY CAESARIAN AND A 50 YEAR OLD AFFECTED MAN WITH TWO NORMAL TEEN-AGE DAUGHTERS.

R E C E S S I V E

JAGER, M. AND REFIOR, H. J.* DIASTROPHISCHER ZWERGWUCHS. Z. ORTHOP. 106* 830-840, 1969.

KITE, J. H.* THE CLUBFOOT. NEW YORK* GRUNE AND STRATTON, 210-218, 1964.

LAMY, M. AND MAROTEAUX, P.* LA NANISME DIASTROPHIQUE. PRESSE MED. 68* 1977-1980, 1960.

LANGER, L. O., JR.* DIASTROPHIC DWARFISM IN EARLY INFANCY. AM. J. ROENTGEN. 93* 399-404, 1965.

LANGER, L. O., JR.* MINNEAPOLIS, MINN.* PERSONAL COMMUNICATION, 1967.

MAU, H.* WESEN UND BEDEUTUNG DER ENCHONDRALEN DYSOSTOSEN. STUTTGART* GEORG THIEME VERLAG, 1958. P. 108 FF.

MCKUSICK, V. A. AND MILCH, R. A.* THE CLINICAL BEHAVIOR OF GENETIC DISEASE* SELECTED ASPECTS. CLIN. ORTHOP. 33* 22-39, 1964.

TAYBI, H.* DIASTROPHIC DWARFISM. RADIOLOGY 80* 1-10, 1963.

*22270 DIBASICAMINOACIDURIA II

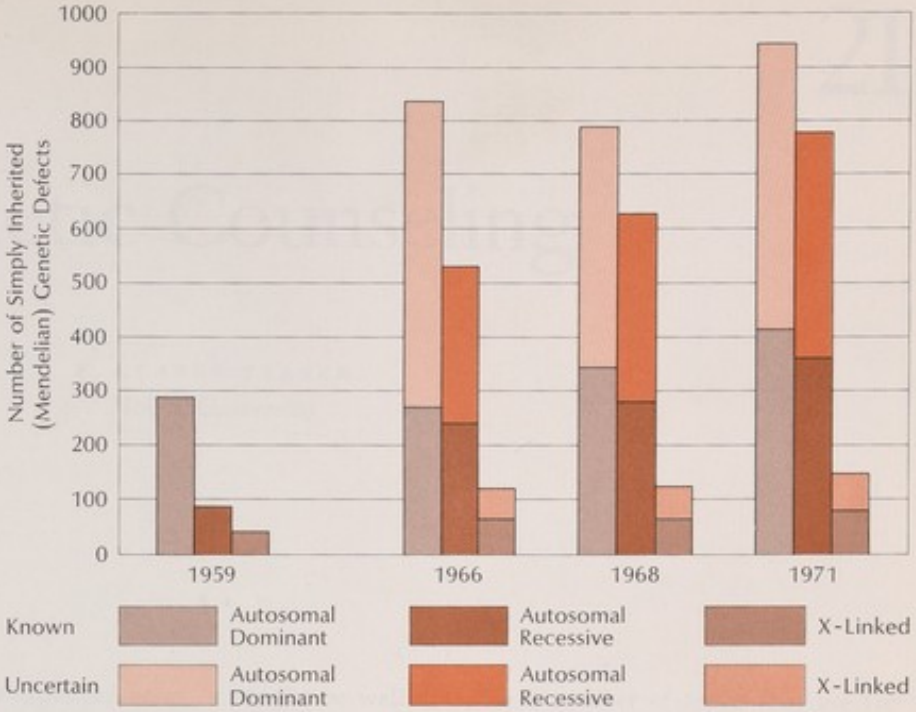
OYANAGI ET AL. (1970) DESCRIBED SEVERE MENTAL RETARDATION, PHYSICAL RETARDATION, MILD INTESTINAL MALABSORPTION SYNDROME, AND INCREASED URINARY EXCRETION OF LYSINE, ORNITHINE AND ARGININE IN TWO JAPANESE SISTERS WITH SECOND COUSIN PARENTS. CYSTINE EXCRETION WAS ALWAYS WITHIN NORMAL LIMITS. THIS DISORDER SEEMS TO BE PARTICULARLY FREQUENT IN FINLAND. KEKOMAKI ET AL. (1967) DESCRIBED AN AFFECTED MALE AGE 23 YEARS AND HIS AFFECTED 15 YEAR OLD SISTER. BOTH REFUSED PROTEIN-RICH FOOD. INSTITUTION OF COW'S MILK AT AGE 1 YEAR RESULTED IN PROLONGED WATERY DIARRHEA AND RETARDATION OF PHYSICAL DEVELOPMENT. WITH INCREASE IN PROTEIN IN HIS TEENS THE MALE GREW BUT MENTAL FUNCTION DETERIORATED AND TYPICAL ATTACKS OF STUPOR

much more likely to turn up in two or more sibs and thereby rouse the suspicion of genetic etiology. In most human populations, by contrast, the carrier of a recessive gene has a much smaller chance of mating with another carrier, and even where such mating occurs, unless the offspring are fairly numerous the homozygous "visible" expression of the mutation may not turn up or, if it does, may turn up as a single, sporadic case that will escape the geneticist's scrutiny.

I have already suggested the special importance of inbred groups such as the Amish in revealing hitherto unsuspected recessive disease entities. As with the laboratory mouse, inbreeding increases the number of homozygotes (despite a common misconception, it does *not* increase the number of "bad" genes in the population, merely the likelihood of their coming together in a given individual). The large families common among such groups increase the likelihood of multiple cases in a single sibship, while further "visibility" is contributed by the spatial concentration of the lineage or deme in question.

However, the genetic isolation of these groups itself limits their value; unless a mutant gene is already present it is unlikely to come in from "outside" (though it might, of course, arise through spontaneous mutation). Delineating discrete genetic entities in less inbred populations is a much tougher job. Joseph Warkany of Cincinnati, for example, has pointed out that among the several hundred thousand inhabitants of American institutions for the mentally retarded, there must be many hitherto unidentified entities whose isolated examples have not yet been "pulled together" into a clinically intelligible or identifiable group. The problem is a matter of what might be called information logistics – in modern parlance, "getting it all together."

It is perhaps appropriate, then, that I close this chapter on a note already sounded several times by other contributors. Continued progress in nosology, as in all other aspects of genetic disease, depends on the close collaboration of investigators and clinicians from all parts of the country – and, indeed, the world.



Rapid increase in number of known genetic diseases is shown by comparison of Verschuer's 1959 listing with later Johns Hopkins' compilations. "Uncertain" category includes mainly conditions in which only the mode of inheritance is not yet certain, plus a few in which the genetic character of the disease is still in doubt.

Johns Hopkins' computerized master catalogue lists all known and suspected mendelian diseases with supporting references; it is periodically published in the form of print-out sheets like that at left. Asterisks show conditions in which the mode of inheritance is definitely known.

21

Genetic Counseling

F. CLARKE FRASER
McGill University

Our understanding of genetic disease has expanded enormously in the past 20 years, but the proportion of cases where this understanding has led to a successful approach to treatment is still disappointingly small. Though we now diagnose genetic anomalies more surely, and comprehend their genesis and mechanics more deeply, in relatively few cases can we intervene effectively to prevent, reverse, or ameliorate their manifestations.

This means that if we are to aim at reducing the amount of ill health caused by deleterious genes we must aim at avoiding their transmission rather than reversing their effects. The expanding list of known genetic anomalies, more accurate knowledge of their modes of inheritance, and the employment of new investigative techniques, together enable us to predict with increasing accuracy the probability that a particular couple will produce an abnormal infant, and thereby reduce the likelihood that such an infant will be born or even conceived. Prophylaxis of this sort is the province of the genetic counselor.

Genetic counseling is not a recognized medical specialty, and may never be, since its practice, strictly speaking, does not require full medical training. Nonetheless, considered in its full range, it requires an impressive amount of specialized knowledge – far too much to be conveyed within the bounds of a single article. My intention here is therefore not to turn the reader into a full-fledged genetic counselor, but rather to describe the potential role of the community physician in this area of preventive medicine. In a sizable number of what might be called uncomplicated genetic conditions, even the nonspecialist can give his patients useful advice, and in other, less clear-cut, situations he can still facilitate the work of the professional counselor and help ensure that both the patient and the community derive maximum benefit from the latter's recommendations.

Before discussing these points in detail, however, it might be well to define a number of terms that are frequently confused by laymen and occasionally by physicians. These are *congenital*, *genetic*, *hereditary*, and *familial*.

When we say that a particular disease or other somatic abnormality is *congenital* we are saying no more than that it was present at birth. Such conditions may be environmentally caused, i.e., they result from injury during the birth process or from more or less fortuitous events in utero, such as a maternal rubella infection. Others are *genetic*, i.e., caused by an abnormality of a single (mutant) gene or by an excess, deficiency, or rearrangement of the chromosomes, which may involve large numbers of genes. Others, of course, result from an interaction of genetic and environmental factors.

Some genetic diseases are congenital, others are not, or not detectably so; the degenerative brain disease known as Huntington's chorea, for example, notoriously manifests itself only when the individual is well into adult life. In the case of phenylketonuria, we have a similar though more subtle distinction: The underlying (genetic) metabolic defect is congenital, while its somatic consequences are not, developing only gradually during the postnatal period – which is why they can be prevented or at least ameliorated by appropriate dietary measures.

A *familial* disease is simply one that "runs in families." It can be familial because it is caused by genes or because it results from some environmental agent that runs in families, such as pinworms or tubercle bacilli (though here there is probably a genetic susceptibility as well). Geneticists usually use the term *hereditary* as a synonym for genetic, but in the older literature it was often used to refer to diseases that appeared in parent and child through successive generations, or what we now call *dominant inheritance*.

Chromosomal anomalies are genetic, in the sense that they involve changes in the genetic material, but if they

Autosomal Dominant Conditions

Achondroplasia
Alport's syndrome (hereditary nephropathy with deafness)
Aniridia
Acute intermittent porphyria
Hereditary hemorrhagic telangiectasia
Multiple intestinal polyposis
Myotonic dystrophy
Neurofibromatosis
And, of course, any condition showing a Mendelian dominant pattern in the family concerned

Autosomal Recessive Conditions

Adrenogenital syndrome
Albinism
Ataxia-telangiectasia
Cystic fibrosis
Cretinism, familial goitrous
Homocystinuria
Laurence-Moon-Biedl syndrome
Niemann-Pick disease
Phenylketonuria
Riley-Day syndrome (familial dysautonomia)
Sickle cell anemia
Tay-Sachs disease (infantile amaurotic idiocy)
Werdnig-Hoffmann disease (progressive spinal muscular atrophy)

Sex-Linked Recessive Conditions

G-6-PD deficiency
Ectodermal dysplasia, anhidrotic type
Hemophilia A
Hemophilia B
Duchenne muscular dystrophy
Lesch-Nyhan syndrome
Nephrogenic diabetes insipidus
Ocular albinism
X-linked ichthyosis

are severe enough to kill or sterilize the affected individual they are not inherited, i.e., passed from parent to child. The same is true of dominant mutant genes. But it is quite possible for a chromosomal anomaly to be inherited. For instance, though most cases of Down's syndrome (trisomy 21) are sterile, several cases are on record of a female with Down's syndrome passing her extra chromosome on to her child, which of course was similarly affected. And in the case of one very uncommon form of the condition, in which a somatically normal parent possesses the chromosomal disorder in which the two number 21 chromosomes are attached together (a 21-21 translocation), trisomy 21 will occur in all the offspring (there is at least one tragic case of such a couple producing five successive children with Down's syndrome).

The importance of these distinctions will be seen as we examine the situations in which the physician may be asked to provide genetic counseling. By far the most common is the case of the couple who have produced a child who is born with or soon develops some disease not obviously environmental in origin. The couple will of course want to know the cause of the problem, but usually with particular reference to the probability of its recurrence: "Should we have another baby?" Less often, the physician or counselor may be consulted by an adoption agency concerning a child's suitability for adoption in the light of a family history involving mixed "racial" ancestry or putatively hereditary disease, or perhaps by individuals contemplating marriage. Thus, cousins may want to know the risks of consanguineous marriage, or a young woman or (less often) man may be concerned about the family history of a prospective partner.

It is worth noting that in most of the latter cases the physician may find himself dealing with what one might call the folklore of genetics, notably the myth of the "throwback." I recall, for instance, a young man who planned to marry a girl one of whose grandparents had pronounced negroid characteristics. She herself, however, reflected this ancestry only in a slightly darker than average skin. His parents objected violently to the match on the basis that their grandchildren might be black "throwbacks." Again, there was the case of a couple that contemplated adopting a child with some American Indian ancestry but were concerned because they believed that all Indians go berserk at the slightest exposure to alcohol – and that the child might have inherited this mythical genetic trait.

At the risk of telling the reader what he already knows, I should make clear that the "throwback" theory is pernicious nonsense. With parents of different "racial" backgrounds, the children may resemble one parent more than the other in certain characteristics, but are most likely to be intermediate between them. And in a Negro-Caucasian mating, specifically, the children will in no case be darker than the dark parent – or, for that matter, lighter than the light one. Finally, without going into the alleged "average" intellectual inferiority of Negroes one can say simply that a man's intelligence cannot be measured by his skin color and that there is no evidence that Negro ancestry influences a child's IQ – either way.

Abandoning mythical for scientific genetics, one notes that the commonest question of genetic counseling – "Will it happen again?" – is also, from the physician's standpoint, the simplest. He is confronted not with the possibility that something *might* go wrong but with the certainty that something *has* gone wrong – specifically, a deformed or diseased child. Assuming that the damage can be traced to genetic causes, these will fall into one of two categories: *chromosomal aberrations*, which can usually be identified and defined by cytogenetic studies of the child and (if necessary) the parents, or one of the many *point mutations* that produce the inborn errors of metabolism and various other sorts of diseases and defects. Many of the latter can be diagnosed unambiguously by symptomatic and physiologic criteria; the former, by contrasts, are as a class a good deal more variable in their somatic manifestations. For this reason, among others, they are best referred to a specialist in human cytogenetics, who can undertake the necessary cytologic studies and interpret their results in terms of any risk to future children.

Assuming that the problem can be

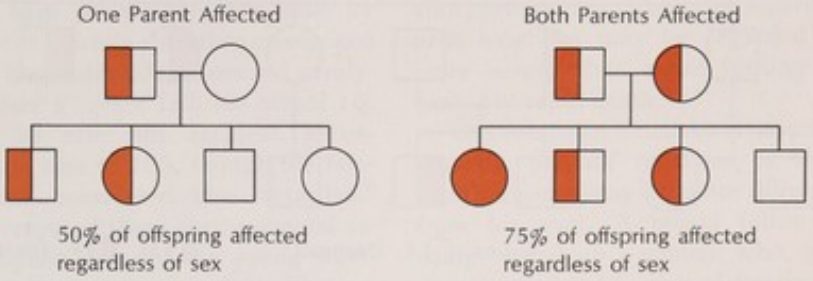
clearly identified as a point mutation, however, referral may well be unnecessary. Most of the genetic errors that occur with reasonable frequency, and so are likely to be seen by the community physician, are inherited in one of only three ways: as autosomal recessives, as sex-linked recessives, and as autosomal dominants. For each mode of inheritance the risk to a subsequent child is known, as the adjacent illustrations show.

The most common among these three forms is autosomal recessive inheritance, found in phenylketonuria, sickle cell anemia, cystic fibrosis, and a host of other conditions. Here, frank disease is found only in individuals homozygous for the mutation in question as the result of having inherited the mutant gene from both parents. Each parent, then, can be presumed to be heterozygous for the mutant trait, assuming that neither of them is affected, so that the chance that a second child will be homozygous is one in four. These odds, of course, apply to any child of such a couple, a fact that sometimes needs emphasis in discussing the problem with the parents, since some people are under the impression that since they have had one affected child the "law of averages" will ensure that the next three will be normal. In fact, it is perfectly possible, though unlikely, for the couple to produce four affected children in a row.

Considerably less common is sex-linked recessive inheritance, notoriously found in hemophilia, the Duchenne type of muscular dystrophy, and more than fifty other conditions. Here the mutant gene is located on the X chromosome (no pathologic mutation is known for the Y chromosome) and expresses itself almost exclusively in males. The male infant with a single X chromosome (received, of course, from his mother) will carry only a single set of "X genes," and if one of these is defective the defect will certainly manifest itself somatically. On the other hand, the female infant who inherits a mutant sex-linked gene from her mother will ordinarily inherit the normal form of the same gene on the X chromosome she receives from her father (assuming he himself does not manifest the disease), so that she will

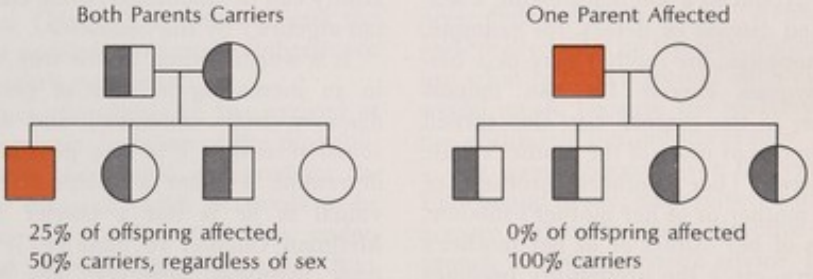
Autosomal Dominant Conditions

Counseling is relatively simple in autosomal dominant disorders, provided the diagnosis is unambiguous, because of the unambiguous relationship between affected parent and affected child. If neither parent is affected, the affected child probably represents a new mutation, and parents can be assured there is little risk to future siblings. However, if the condition is known to skip a generation occasionally, or has a variable age of onset, the counseling must be more cautious.



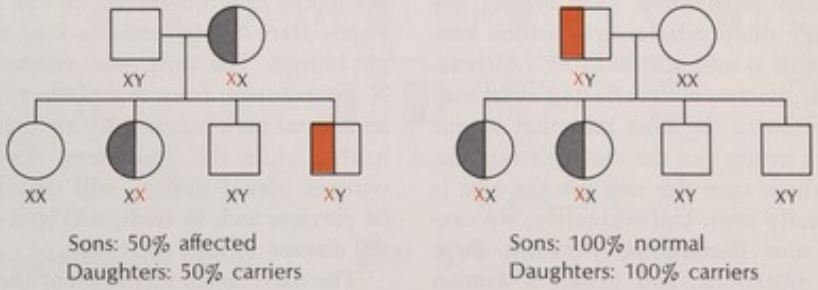
Autosomal Recessive Conditions

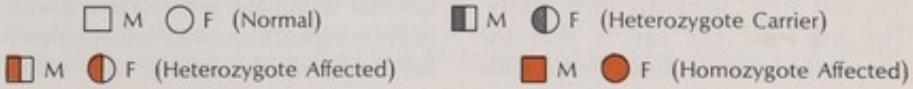
In autosomal recessive conditions, the presumption following the birth of an affected child is that both parents are carriers; they would be advised in counseling that the risk that a second child would be affected is 1:4. Because of the recessive nature of the condition, an affected individual cannot produce an affected child unless the other parent is a carrier, but all his offspring will be carriers.

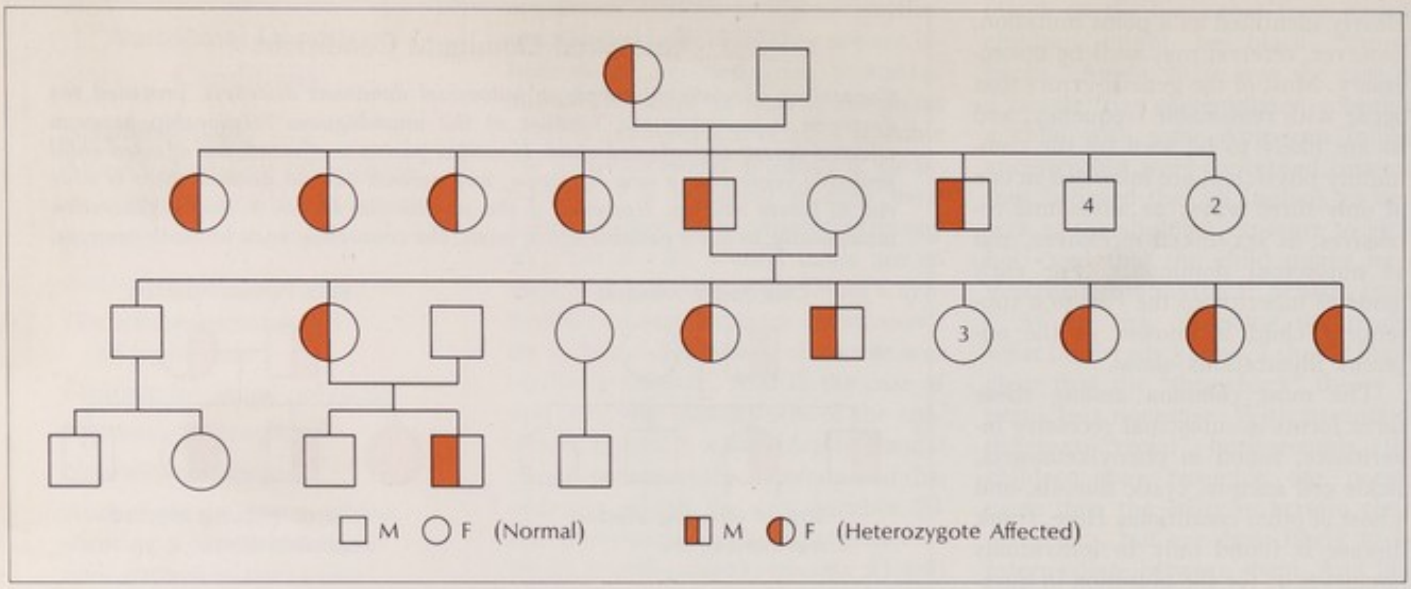


Sex-linked Recessive Conditions

Counseling in X-linked disorders usually turns on whether the mother is a carrier. If she has two or more affected sons, or one affected son and an affected male relative (brother, father, maternal uncle) she must be considered a carrier. A single affected son, however, may represent either a hitherto occult mutation in the mother's ancestors or a new mutation; sometimes special tests of the mother can settle this point. When the mother's carrier status is known and diagnosis is unambiguous, counseling can usually be given by the community physician.







Pedigree of family afflicted with congenital aniridia shows a typical pattern of autosomal dominant inheritance. Theoretically, individuals manifesting the disease should produce offspring 50% of whom are affected, 50% normal. The actual proportions affected here are 50% in the first generation of offspring, 54% in the second, and 50% in the third.

usually be clinically normal, though she can still transmit the disease to her own sons. Thus, where a family has produced a son manifesting a sex-linked disease or defect, for example, hemophilia, the mother may be a heterozygous carrier for the mutant gene; if the disease has also turned up in one or more of the mother's male relatives (her brothers, brothers of her mother or of her mother's mother, sons of her sisters or of her mother's sisters, etc.) the probability becomes a near-certainty. Any future sons she produces will then have a 50% chance of hemophilia; her daughters will be quite safe, but will have the same 50% chance of being carriers.

If the first hemophiliac child is apparently unique in the family tree, however, the genetic situation is a good deal less certain. It may be, of course, that the mother inherited the mutant gene, but that simply by luck it failed to manifest itself among any of her male relatives, in which case the risk would still be 50%. Alternatively, however, the affected child may represent a de novo mutation in one (or a group) of the mother's oocytes, in which case the risk for the sibs is virtually zero. Unfortunately, we cannot now distinguish between these two situations, and to tell a woman that the risk is "somewhere between zero and 50%" is hardly satisfactory. The best one can say is that the greater the number of unaffected male relatives in a hitherto hemophilia-free family, the lower the probability that the mother is a carrier, and this probability can be calculated (using Bayesian algebra) by the counselor.

It is worth noting, by the way, that in an increasing number of genetic diseases, both sex-linked and autosomal recessive, it is now possible to determine whether a "suspect" individual is or is not a carrier (see Mellman, Chapter 16, page 157). The tests involved, however, are sufficiently specialized so that the physician will almost certainly prefer to refer his patient to a counseling center with experience in this work.

Some sex-linked conditions – those that do not lead to early death or incapacitation – can be passed on through the affected father if he begets children despite his infirmity or prior to its development; one such disease is sex-linked choroideremia, an eye disorder. Here the genetics is very simple indeed. The sons, who receive no X chromosome from their father, will be normal both somatically and genetically, while the daughters, though without visual defect, will certainly be carriers and, as such, will transmit the disease to half their sons.

The most easily identified of the inherited pathologies are those with autosomal dominant inheritance, found in Huntington's chorea, "lobster-claw" deformity of the hands, and many other conditions. Here only a single mutant gene is required to produce frank disease, so that the identification of carriers is normally not a problem. An affected individual with one normal parent is presumably heterozygous for the mutant gene, and the chance of its being transmitted to his offspring is one in two. Normal members of the family are likely to be genetically normal once they have passed the age of onset. Some of these conditions, however, occasionally "skip" a generation, meaning that due to "reduced penetrance" they will be transmitted through an apparently normal carrier.

Thus far I have said nothing about the so-called multifactorial diseases, which are produced by a combination of several genetic and environmental factors; they include most cases of cleft lip and palate, diabetes, schizophrenia, and other conditions both common and uncommon. In this area, the basis of inheritance is still poorly understood, though progress is being made (see Carter, Chapter 19, page 199); the probability of their occurrence in a given situation is estimated entirely from empirical data for that particular situation. For these reasons, such problems are best left to the specialist.

The same is true of even the point mutations in cases where the genetics are less clear-cut, e.g., where the problem is not an actual diseased

child of a given couple but of putatively genetic disease elsewhere in the family – for instance, Huntington's chorea in a grandparent or muscular dystrophy in a cousin. Assessment of the probabilities in these and similar cases, e.g., problems of adoption or incest, may involve the compiling of an extensive family pedigree and the drawing from it of inferences that are often both complicated and tenuous. Special tests may also be required. The community physician (unless he has made a scientific hobby of human genetics) will doubtless prefer to turn over the problem to a professional counselor who may be able to reduce some of the uncertainties and is at least accustomed to them.

I should emphasize, however, that even in these cases the contribution of the community physician need not, and should not, be limited to mere referral. Before going into these aspects of the matter, however, I should like to mention some aspects of the "uncomplicated" cases in which the community physician himself undertakes to offer advice.

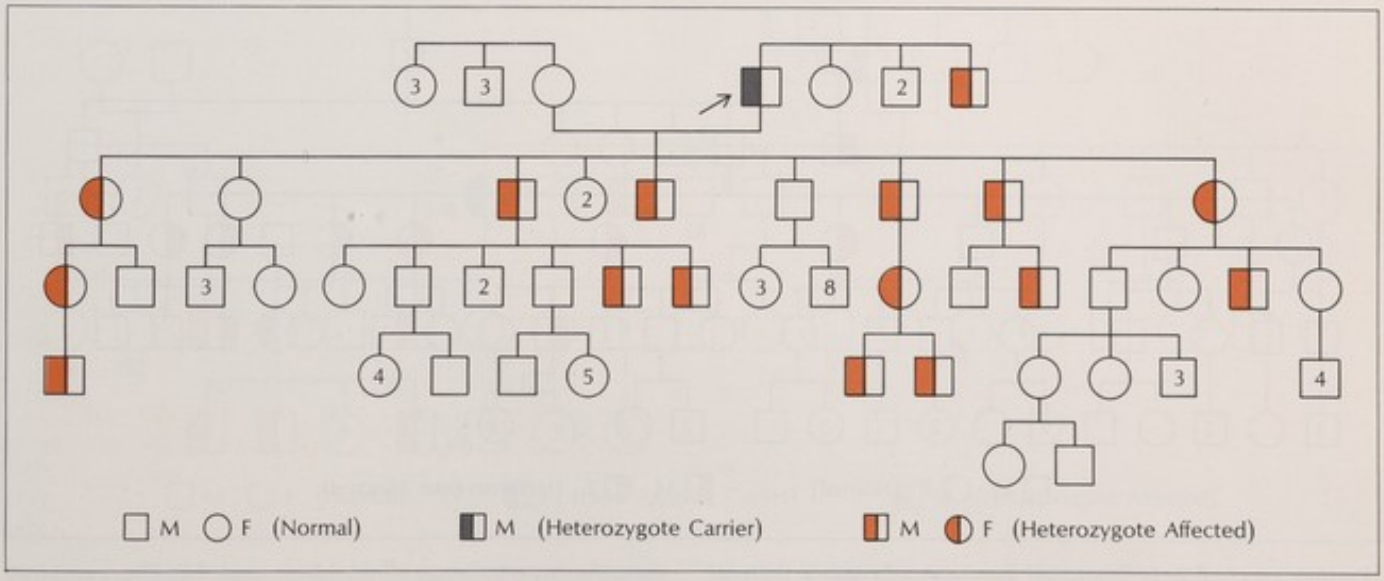
The first point, which should go without saying, is to be sure of the diagnosis – in particular, to make sure that the problem in question is really of genetic rather than environmental origin. This is particularly true with such pathologic catchalls as "mental retardation," which, as we well know, can result from a great variety of causes. I think, for instance, of a young woman patient justifiably concerned about having children because a brother and two first cousins were apparently mentally retarded. However, investigation revealed that the brother and one of the cousins both had shown signs of cerebral damage from hypoxia at birth (the former had been born prematurely, the latter by cesarian), while the other cousin had been hospitalized for coma and paralysis after a severe fall; his actual IQ, only 85 after the accident, subsequently rose to 108, though the family still considered him "retarded." Awareness of these environmental catastrophes changed the young woman's genetic prognosis from discouraging to reassuring, and her children were in fact quite normal.

Second, there is the problem of "genetic heterogeneity." Clinically similar diseases may show different patterns of inheritance – for instance, there are two types of mucopolysaccharidosis that resemble each other quite closely, but one shows autosomal (Hurler's syndrome) and the other sex-linked recessive inheritance (Hunter's syndrome).

Once the physician has arrived at a prognosis, in terms of the probability that the couple will produce another diseased child, he needs to be as sure as possible that they understand just what the risk means. Obviously, different couples will react in different ways to the same risk, and not always in the ways one expects. An obvious variable is the severity of the disease involved; generally speaking, people are willing to take a relatively high risk if the possible disease is mild, while being unwilling to take even a small risk if it is severe. And subjectively, at least, diseases causing slow, progressive debility or a permanent handicap may be regarded as more severe than those leading to early but rapid death.

The counselor must keep in mind that the parents' reactions to their child's disease may be quite different from his own. A recent follow-up study turned up a woman who, told that she had a 5% risk of bearing a second cleft-lip child, avoided pregnancy for years and underwent an illegal abortion when she accidentally became pregnant, a course that most of us would consider overreaction. In contrast was another woman who, having taken a one-in-four chance of bearing a second albino child and lost, expressed the wish that I had presented the facts more forcefully and "really laid it on the line" concerning what having a second such child would mean, because she understood the arithmetic but not the reality.

My own tendency over the years, I am bound to say, has evolved toward a more rather than less directive approach in giving genetic advice. I do not think the physician should go so



In "lobster-claw" deformity, as in some other autosomal dominant diseases, inheritance patterns may be confused by "reduced penetrance" of mutant gene, whereby a few individuals possessing it are somatically normal. Arrow shows man in first generation who, though apparently normal, must have possessed the gene, since both his brother and six of his children were affected.

far as to say "Thou shalt not," but where the risk is serious and the possible consequences severe, it seems to me perfectly proper to say the equivalent of "Now be quite sure you have really thought about what another defective child would mean, in terms of the specific problems that would be encountered." And if the parents ask me "What would you do?" I am prepared to say that putting myself in their shoes as best I can, I would or would not, as the case may be, take the chance.

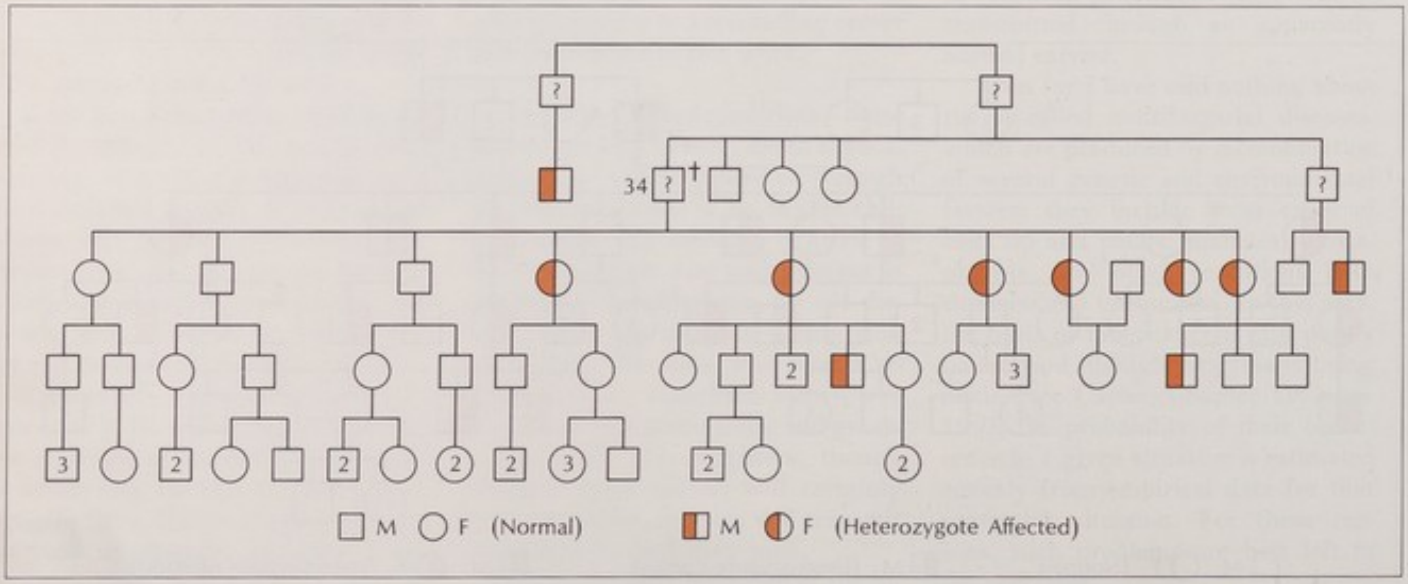
Finally, we should be aware that both the demand for genetic counseling and the options available to the counselor and his patient are almost certain to increase as the result of two developments: changing attitudes toward early interruption of pregnancy and new techniques for detecting genetic disease, or the possibility of it, in utero (see Dancis, Chapter 24). Increasingly, the physician will be asked not merely "Should I have another child?" but "Should I have this child?" In some cases information obtained through amniocentesis can help provide an answer. Suspected chromosomal disorders have been detected through examination of fetal cells, and certain enzymatic defects also seem to be detectable, although sometimes more elaborate techniques are required. In the case of sex-linked diseases, determination of the sex of the embryo (which can be reliably done by cytologic techniques) will at least narrow the possibilities, even when the disease itself cannot be identified in utero. A woman carrying the hemophilia gene, for example, might well choose abortion of a male fetus, which has a 50% chance of being diseased, while carrying to term a female, who at worst would be a carrier.

Where termination of pregnancy is possible and legal, the physician would do well to acquaint himself with attitudes and procedures at the hospitals where the procedure might be undertaken, lest he raise false hopes in his patient. Even in places where abortions are legal, or legal under certain circumstances, some hospitals, for one reason or another, drag their feet on approving them; in others, mere bureaucratic paperwork may have a similar effect. A case in point (though amniocentesis was not involved) is that of a woman who had borne one hemophiliac son and, despite attempts to employ birth control, a second one, as well as a normal daughter. In desperation, she applied to a hospital for a tubal ligation, but by the time they had approved it she had been forced by another pregnancy to change her application to one for abortion. The hospital committee refused the application on the grounds that the admitted 25% risk of yet another hemophiliac child was "not sufficient" to justify the procedure; the woman is now the mother of three hemophiliacs.

The medical profession, I believe, must face up to the fact that in the area of genetic disease, at least, preventive medicine in some cases will mean the prevention of birth and in others the prevention of conception by surgical means. It seems to me highly deplorable that in so many places laws or professional attitudes still make it difficult to obtain abortion or sterilization when parents, on reasonable grounds, decide they do not want to take the risk of bringing a deformed or diseased child into the world.

Let us turn now to cases in which the physician has referred his patient to a professional counselor. There are three areas in which his assistance can be valuable. First, there is the simple matter of making sure that the counselor gets a complete record of the referring physician's observations, along with the results of any tests that have been made – this in the interest of saving everyone's time and the patient's money.

The second area lies more in the psychologic than in the strictly medical field. Giving genetic advice, as can be imagined, is always a delicate business, and doubly so in the case of the professional counselor, who will



Further uncertainty in autosomal dominant inheritance patterns may occur when the disease manifests itself relatively late in life, as with Huntington's chorea (pedigree above). Two individuals in first and second generations are not known to have suffered from the disease, though all must have carried it. The reason may have been either reduced penetrance of the gene or their death before symptoms became manifest (one is known to have died at age 34). Numbers indicate additional unaffected individuals.

probably have seen the patient only once or twice before. His efforts can often be more diplomatic if there is information on the family background. For instance, the presence of, or the possibility of, a diseased child in the family is a fertile source of marital discord; occasionally the result is an actual breakup. Often grandparental attitudes are important and may influence the parents' attitudes, especially if the condition is such that it can be "blamed" on one or the other side of the family. To the extent that the referring physician can alert the counselor to sensitive areas of this sort, the latter's dealings with the patient can be more tactful and therefore more effective.

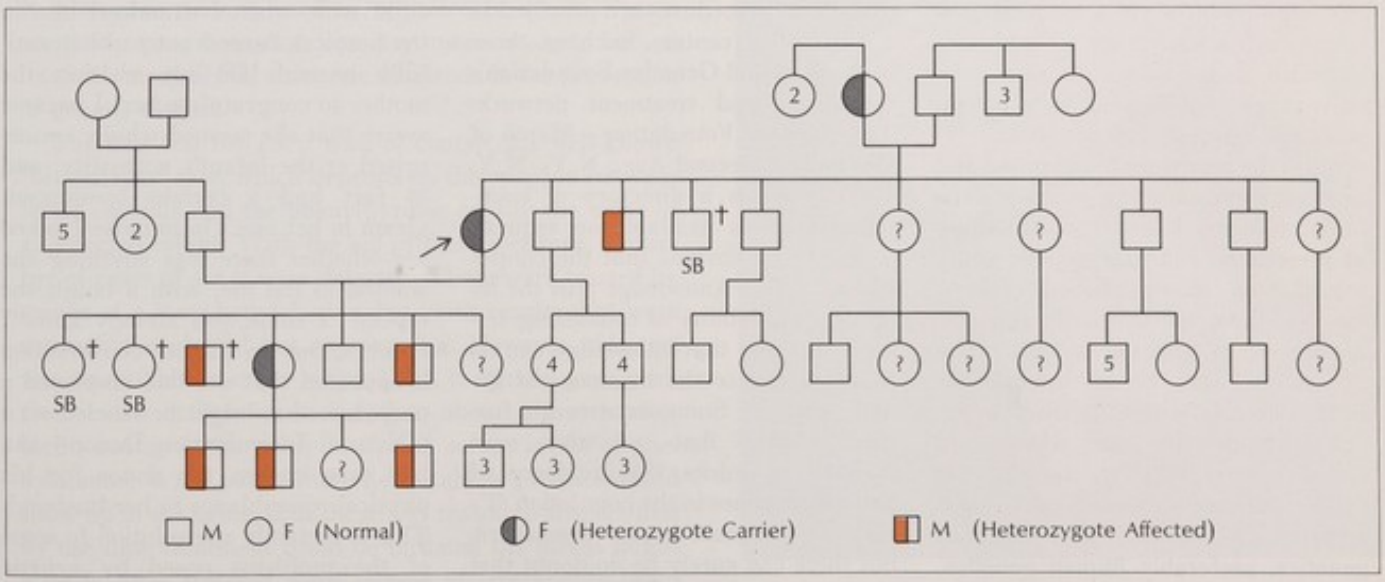
The third area is that of follow-up, both of the patient and (on occasion) of the relatives. So far as the patient is concerned, this should initially be done by the counselor himself, who can thereby check on whether his advice has been understood and also, by ascertaining what people actually do about his advice, improve his professional art. If, however, a second interview by the counselor is impracticable, e.g., for geographical reasons, follow-up by the referring physician in a week or two is desirable, if only to check tactfully on his patient's comprehension (he should, by this time, have received a copy of the counselor's advice). I have known cases where parents have somehow converted a counselor's opinion of three-to-one odds *for* a normal child to three-to-one *against*. And subsequent follow-up, with support and reinforcement, is often a valuable resource to the family that their physician can supply much better than the counselor.

Certain kinds of genetic findings – notably those in the areas of sex-linked recessive and autosomal dominant disease – will imply a risk not merely to the patient being counseled but also to certain of his or her relatives. On occasion, the mere fact of an adverse genetic finding in one member of a kinship will stimulate other members to seek counseling themselves, which is of course all to the good. Often, however, these "secondary cases" for one reason or another fail to show up, which poses a serious dilemma for both the counselor and the referring physician. On the one hand, where the findings identify an individual who is unknowingly at serious risk of bearing a diseased child, someone surely has the professional responsibility of warning her. On the other hand, anyone attempting to exercise this responsibility, be it the counselor, the referring physician, or some third party such as a public health nurse, stands a good chance of reaping the usual reward for the bearer of unsolicited and unwelcome advice: resentment.

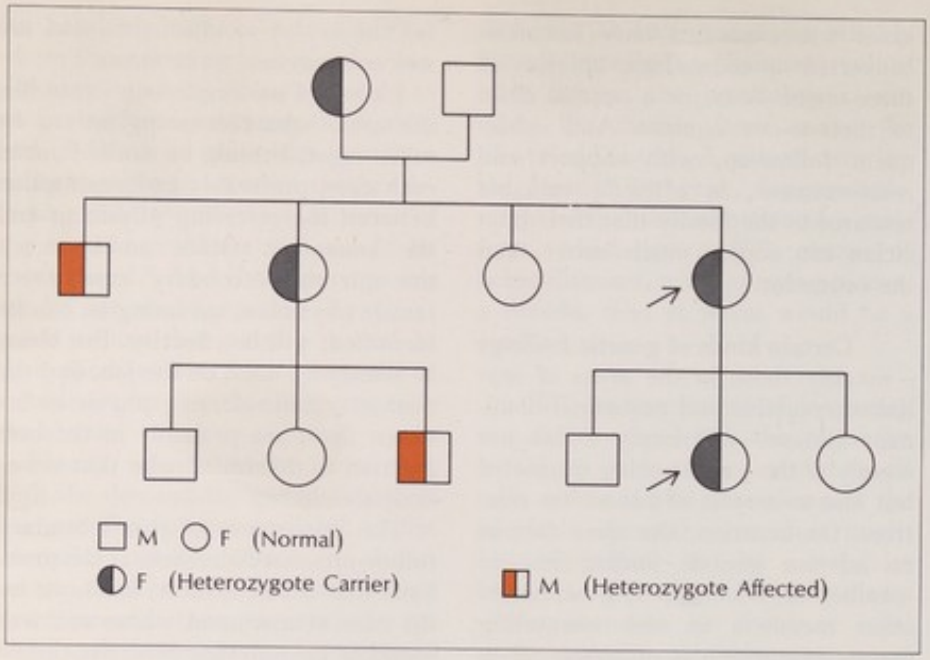
I know of no simple solution to this dilemma; whatever steps are to be taken must, I think, be worked out in each case, preferably by consultation between the referring physician and the counselor. Often an approach through the secondary case's own family physician, assuming he can be identified, will be effective. But clearly somebody must do the job, and the counselor and referring physician between them are probably in the best position to determine who that somebody should be.

The importance of this secondary follow-up, as well as some of the problems that it can involve, is shown by the case of a woman whose son was found to be suffering from sex-linked choroideremia, which produces a progressive loss of vision. Investigation revealed that she had inherited the trait from her father, who was totally blind as a result of the disease, yet no one had warned her that she was certainly a carrier. Her sister, too, is certainly a carrier but still does not know it – and the two are not on speaking terms.

Given the probable increased demand for genetic counseling, the community physician, particularly if he practices outside a large metropolitan area, may well wonder whether it would be worth while to set up coun-



In many sex-linked diseases, the carrier status of a given woman can be ascertained only by inference from pedigree patterns. In this "classic" hemophilia pedigree, one woman (arrow) must be a carrier, since her brother and two sons are affected; the same is true of one daughter and (by a different line of inference) of her mother. Her other surviving daughter, her granddaughter, her sisters, and their daughters may or may not be carriers; there is no way of being certain one way or the other.



In some sex-linked diseases, carriers can be identified biochemically as well as by pedigree patterns. This pedigree for Duchenne muscular dystrophy reveals two women as obligatory carriers. Of five other possible carriers, whose genetic status was left uncertain by the pedigree, two (arrows) showed abnormally high levels of the enzyme creatine phosphokinase and are therefore deemed to be carriers; those with normal enzyme levels are likely, but not certainly, noncarriers.

seling services in one or another of the local hospitals. In fact, the requirements of such a service are not very elaborate.

A basic requirement – we are now discussing counseling beyond the "uncomplicated" level discussed above – is a laboratory that can undertake cytogenetic studies, both sex chromatin and chromosomal. If a hospital possesses such a laboratory, or wishes to establish it for other purposes, the only other requirement is a person with the interest and competence to acquire the necessary background and to engage in counseling on a part-time basis. (To my knowledge, there are at present no full-time genetic counselors even in established centers; most of them are primarily engaged in research or teaching or both.) The counselor might be an internist or pediatrician interested in medical genetics, preferably one who could spend several months, at least, at some existing counseling center. Alternatively, he might be a Ph.D. in genetics, preferably human genetics, resident in the community (such individuals are not uncommon in academic towns), though here I would consider a period of experience in an existing center virtually obligatory. I know of one such arrangement in one of our smaller Canadian towns, where the wife of a local doctor, with a degree in human genetics, does counseling in conjunction with the local hospital. Facilities for blood-grouping and biochemical screening tests should also be available. More special diagnostic procedures are provided in a number of centers, such as those in the National Genetics Foundation's counseling and treatment network. The National Foundation – March of Dimes, 800 Second Ave., N.Y., N.Y. 10017, makes a directory of counseling services available on request.

It will be obvious that the continued growth of knowledge plus the increased availability of counseling services to apply that knowledge can do much to reduce the incidence of genetic disease. Some experts go further, feeling that we may even manage to reduce the incidence of pathologic genes in the population. To me, this forecast seems premature, but there can surely be no doubt that genetic counseling can obviate a great deal of physical and mental suffering.

In conclusion, I might mention a couple of instances demonstrating that even the serious, and occasionally somber, business of genetic counseling has its lighter side. A colleague of mine, for example, cites the case of a young albino man who wanted to know the likelihood of his children being albino. Since his fiancée was neither albino nor related to him, he was told that the chance of her carrying the gene for this autosomal recessive condition was about 1%, and the chance of a child being albino was about one in 200. Curiously, this apparently reassuring advice seemed unwelcome to him. It was then suggested that an examination of the fiancée (by transillumination of the iris) might help determine if she was in fact a heterozygote, but he strenuously resisted this proposal. After further discussion, he admitted that not he but the fiancée was pressing the case for marriage and that he had hoped to use the counselor's opinion as an argument against the match!

A second case concerns a young woman patient of mine whose first baby showed an unusual combination of malformations. Since her husband's brother (who had married a cousin) had produced two similarly defective children, she was naturally concerned. Though I could not (and indeed still cannot) identify the disease with any known genetic disorder, there was obviously at least a presumption of appreciable genetic risk.

Subsequently, she had a second child who, when I examined him in the hospital, turned out to be beautifully normal. On my visiting the mother to congratulate her, I became aware that she seemed wholly unsurprised at the infant's normality, and in fact had a certain complacent gleam in her eye. On impulse I asked her whether there was anything she wanted to tell me; with a laugh she replied "I think you already know!" I didn't, but on further conversation it appeared that she had employed a procedure that might be labeled NID – Natural Insemination Donor – and had even chosen the donor for his physical resemblance to her husband. This is certainly one solution to some of the problems posed by genetic disease, though hardly to be recommended by even the most sympathetic counselor or physician!

22

Mass Screening for Genetic Disease

ROBERT GUTHRIE
State University of New York, Buffalo

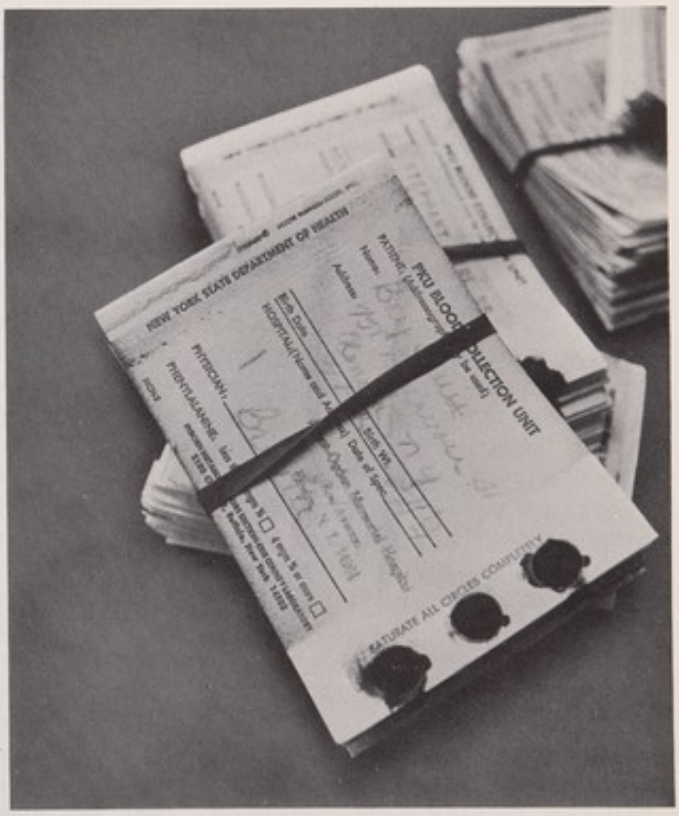
It is now some five years since mass screening of neonates for phenylketonuria became routine throughout the United States. The tests have been made legally mandatory in 43 states and are carried out voluntarily in the other seven, thereby covering an estimated 90% of the 3.5 million infants born annually in this country; similar measures are under way in at least 15 other nations. All told, these screening programs add up to perhaps the most extensive exercise in preventive medicine since the development of polio vaccine.

Given a public health effort of this magnitude, a review of the results would be appropriate in any case. But such a review seems especially desirable now, in the light of some current efforts to discontinue screening programs under the guise of a misconceived "economy." In fact, as I shall demonstrate, PKU screening has proved itself not merely a medical but also an economic success, figured in the most hard-nosed budgetary terms: the ounce of prevention that saves an expensive pound of amelioration. It is also, as we are learning, a prototype of other mass screening programs that can pay more modest, but no less real, medical and economic dividends in the prevention of genetic disease.

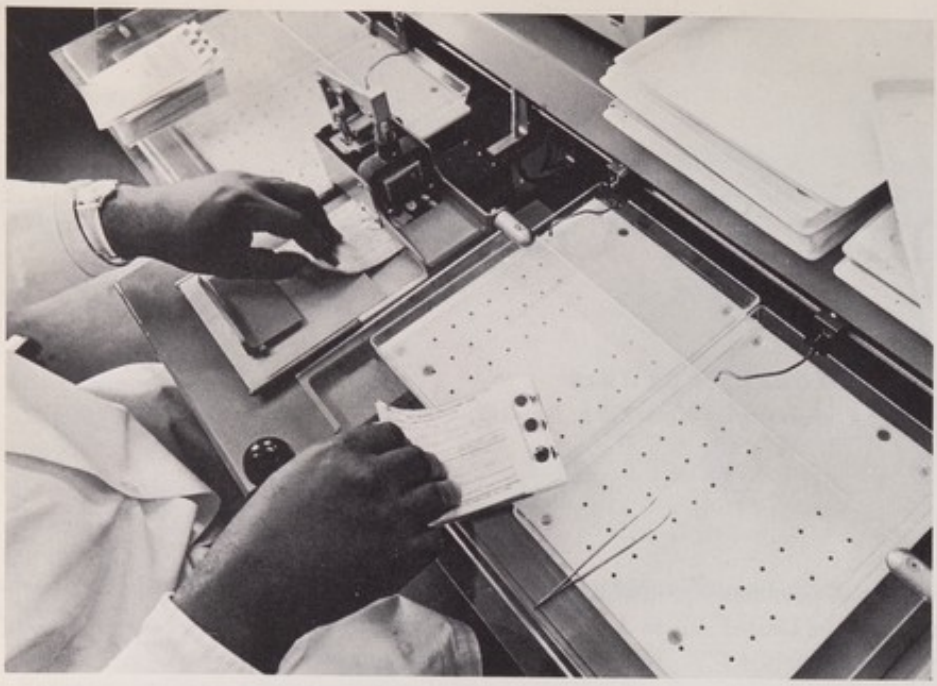
The first test for PKU was, of course, the well-known "blue-diaper" test, which depends on the reaction between ferric chloride and the phenylpyruvic acid in the urine of an affected infant. With the aid of this procedure, a number of cases of PKU were detected. These were treated by means of a low-phenylalanine diet, with results that were somewhat equivocal but encouraging. However, this test was unsuitable in many ways. First, it usually gave positive results at best no earlier than about a month after birth, owing to the delay in rise of serum levels of phenylalanine to the point where its metabolic product would show up in the urine; thus there was reason to suspect that by the time treatment could be initiated the infant might already have suffered some degree of irreversible brain damage. Then, its success as a mass screening device depended heavily on the cooperation of untrained or inexperienced individuals – parents, public health nurses, etc. – who would have only the most general notion of what to look for, since their chances of having seen another case of this rare condition were remote.

In the late 1950's, my associates and I were asked to run tests on serum phenylalanine levels to monitor the dietary treatment of two PKU patients. Dissatisfied with existing procedures, which were cumbersome and costly, we devised the bacterial inhibition test, which is neither.

This procedure employs cultures of *Bacillus subtilis* in an agar medium. Normally, this organism is capable of



Automated screening for PKU and other inborn metabolic disorders begins with the arrival by mail of blood samples like those shown above; on a recent Thursday a total of 137 arrived at the New York State health department's laboratory in Buffalo.



Filter papers containing the blood samples are placed in machine for punching (left); one dried blood spot yields four discs (as shown at far right), each of which is positioned by a suction device in a separate dimpled tray (center) for testing. Another spot is punched for trays containing agar medium (trays of this type can be seen at the technician's right). Twelve tests are made

synthesizing phenylalanine (which it requires for growth), but if the medium contains the compound β-2-thienylalanine, an analogue of phenylalanine, some essential metabolic process is blocked, so that the bacterium cannot reproduce. When a filter-paper disc impregnated with serum or blood from an affected individual is placed on the medium, however, the exogenous phenylalanine diffuses into the medium, producing a zone of growth surrounding the disc. By appropriately adjusting the concentration of the inhibitor, the test can be made sensitive over virtually the entire range of expected phenylalanine concentrations, the width of the zone of growth being proportional to the amount of phenylalanine present.

Since (as was soon determined) the serum or blood did not need to be fresh, the test could be applied on a mass basis, with samples processed in batches at central locations. There was a consequent gain in both reliability and economy, an advantage already suggested by trials with urine-impregnated filter papers. Further modifications involved using standardized discs, mechanically punched from the impregnated paper, and spores of *B. subtilis*, rather than the organism itself, for "seeding" the medium. This "instant" culture technique obviated the sterile procedures previously needed to maintain the culture and perform the test.

It is this test that has become the standard method used in nearly all PKU screening programs. It has proved itself accurate, quantitative, and reliable, as any test designed to screen for a rare disease must be. As might be expected, its widespread use has taught us a good deal about PKU itself. In most European and American populations, PKU incidence turns out to be about one case per 10,000 births, implying a gene frequency (allowing for consanguinity) of something less than one per hundred of population. Though this incidence is low, it is – significantly, I think – a good deal higher than previous estimates based on urine testing. One might note, incidentally, that both Ashkenazic Jews and (especially) Afro-Americans appear to have an unusually low incidence of PKU.

The fact that the test is carried out in the immediate postnatal period (3 to 5 days in normal infants, 5 to 8 days in prematures), in addition to making possible more timely treatment, has sharpened our awareness of the potential significance of particular phenylalanine levels. Originally, a level of 20 mg% or more was considered diagnostic, but the discriminatory level has steadily been lowered to the point where now the test is so adjusted as to pick out all infants with levels exceeding 4 mg%. A test this sensitive can be expected to turn up many "positives," and in fact we find that we can expect to retest 30 to 50 suspicious cases to turn up one actual case of PKU. Nonetheless, the extra sensitivity is justified, since it has become apparent that there is no sharp correlation between inherited lack of the enzyme phenylalanine hydroxylase (the fundamental biochemical defect in PKU) and actual serum levels of the amino acid. Of all infants with the enzyme defect, as determined by phenylalanine loading tests, some 25% maintain phenylalanine serum levels of between 8 mg% and 15 mg% on a normal diet – levels that would once have been considered benign. Some investigators believe that infants with these low levels do have a genetically variant form and are not at risk for retardation (see Hsia and Holtzman, Chapter 23).

Of course we cannot be certain that these infants would grow up retarded, since the only conservative (and ethical) course is to limit their phenylalanine intake as if they were "typical" PKU cases. But it is significant that studies in Germany, Austria, Denmark, Scotland, and the U.S. of men-

possible by the three samples. Four are routine: for phenylalanine, leucine, methionine, and tyrosine.

tally retarded populations have turned up a significant number (25% of persons with elevated phenylalanine levels) of individuals with serum levels in this low range – levels, incidentally, at which phenylketone does *not* show up in the urine, suggesting that a more accurate term for the disease would be phenylalaninemia. We don't know how high the levels were in these individuals in infancy. There is evidence that a drop in phenylalanine levels occurs in phenylketonurics as they get older.

My own feeling is that infants in whom phenylalanine levels of 6 mg% or greater persist for two weeks after the initial test, and in whom the finding of phenylalaninemia is confirmed by more precise techniques such as chromatography, should be considered at least provisional PKU patients and placed immediately on a low-phenylalanine diet. This conclusion is based not merely on the general desirability of giving any disease prompt treatment but on what we know about the nature of PKU specifically. It is generally believed that most if not all the pathologic effects of phenylalaninemia derive from inhibition, by the abnormal concentrations of this amino acid, of cellular uptake of other equally essential amino acids, with resultant growth retardation Because of its extraordinarily rapid expansion in the postnatal period, the brain is the main and perhaps the only "victim" of this growth limitation. It follows, then, that any delay in treatment beyond the minimum period necessary for a confirmed diagnosis risks brain damage; it is during the first few months of life that the brain grows most rapidly.

It is possible that in some of these marginal cases the patients may not be true PKU sufferers but may be manifesting a different variant, an idiosyncratic phenylalaninemia that seems not uncommon among young infants. These cases, however, can be eliminated by periodic monitoring of blood phenylalanine levels, which on a low-phenylalanine diet could be expected to drop below normal (roughly 2 mg%) as the transient disturbance subsides. And monitoring is of great importance in any event, to check on the effects of the dietary regimen. Here, as elsewhere, we are learning that overtreatment, resulting in hypophenylalaninemia, may be quite as damaging as undertreatment. Moreover, clinical experience has demonstrated that individual requirements for phenylalanine among PKU patients vary greatly; one child may require three times as much dietary phenylalanine as another to maintain the same blood level. The difference, apparently, is in the rate at which the amino acid is metabolized and excreted by alternative metabolic routes.

Monitoring is no more difficult than the initial screening. Indeed, it does not even require the infant's being brought to a clinic or physician's office, since the mother can quickly be taught how to take a blood specimen in a standard filter-paper blank that can be mailed to the testing center.

Given prompt, individualized, and carefully monitored treatment along these lines, there is every reason to hope for a favorable outcome, in the sense of producing a normal, healthy child and adult. There has been much debate among pediatricians and medical geneticists over what constitutes a rational prognosis in PKU, based on the undoubted fact that treatment has seemed much more effective in some children than in others. My own belief is that much or all of this variability can now be accounted for, first by variation in the time when treatment was instituted, and, second, by the variations just cited in the "minimum daily requirement" of phenylalanine.

Whether early treatment, adjusted to the individual's metabolic needs, will prove successful in *all* cases remains to be proved; these sophisticated modalities have been employed for only a few years, so that long-term studies of such patients do not yet exist. It can be said, however, that several such studies have now passed the five-year mark, and that in all cases the IQ of the children falls into a normal distribution. Moreover, the incidence of behavioral disturbances such as hyperactivity, which are a conspicuous feature in some (though not all) cases of untreated PKU, seems no greater than normal.

The prospect that most or all PKU patients will grow up into essentially normal individuals, however, raises another health problem that will require attention: prenatal or "maternal" phenylalaninemia. Physicians have already accumulated a limited amount of experience with this condition, owing to the fact that a small but significant fraction of *untreated* PKU patients are of normal or low-normal intelligence – i.e., not so retarded as to require institutionalization and therefore (if female) not unlikely to become pregnant.

The majority of infants born to such mothers show severe retardation even though their (postnatal) serum phenylalanine levels are normal. Not infrequently we find evidence of more serious growth disturbance, such as microcephaly, which is seldom found in "ordinary" PKU. The obvious presumption is that they have been damaged by high phenylalanine levels in the mother's blood. The amino acid crosses the placenta, and studies of a pregnant monkey in which phenylalaninemia was induced by means of a special diet have shown that levels of the compound were even higher in the fetus. The equally obvious conclusion is that during pregnancy the mother should be put on a low-phenylalanine diet, and this has been done in several cases (see illustration on page 235)

Evaluation is complicated by the fact that a minor fraction of infants born to PKU mothers show no intellectual impairment. One source of this apparent anomaly may well be the previously cited "marginal" cases in

Newborn Screening Tests[1] for Inherited Abnormalities

Disease	*Test Substance*	*Test*	*Automated*	*Treatable*
Phenylketonuria	Phenylalanine	BIA[2]	✓	✓
Maple Syrup Urine Disease	Leucine	BIA	✓	✓
Tyrosinemia (transient and permanent)	Tyrosine	BIA	✓	✓
Homocystinuria	Methionine	BIA	✓	✓
Histidinemia	Histidine	BIA	✓	✓
Valinemia	Valine	MBIA[3]	✓	✓
Galactosemia Transferase or Kinase Deficiency	Galactose	MBIA or Coliphage	✓	✓
Transferase Deficiency only	Galactose Uridyltransferase	Beutler	Partly	✓
Argininosuccinic Aciduria	Argininosuccinic Acid Lyase	EA[4]	✓	✓
Orotic Aciduria	Orotidine-1′-phosphate Decarboxylase	EA	✓	✓
Hereditary Angioneurotic Edema	Cl Esterase Inhibitor	FST[5]	Partly	No
Emphysema (adult) Liver Disease (infant)	α-Trypsin Inhibitor	FST	Partly	No
Sickle Cell Anemia	Hemoglobin	Electrophoresis	Partly	Palliative only

1 *Using dried blood spot filter paper specimens* 2 *Bacterial Inhibition Assay*
3 *Metabolite Bacterial Inhibition Assay* 4 *Enzyme-Auxotroph Bacterial Assay*
5 *Fluorescent Spot Test*

which the mother's serum phenylalanine approaches normal levels; certainly in one case where all the children were of above-normal intelligence, the mother's serum phenylal anine was found to be only 10 mg%. It is also not unlikely that variations in fetal enzymes are involved here. We now know, contrary to earlier beliefs, that the enzyme phenylalanine hydroxylase is manufactured prenatally in the fetal liver, but there is doubtless appreciable variation in the time when that enzyme becomes manifest. Accordingly, there must be variation in the degree to which it can protect the fetus by metabolizing the exogenous (maternal) phenylalanine.

Be all that as it may, it is clear that with some hundreds of PKU cases being detected and treated annually, the population at risk of prenatal phenylalaninemia can be expected to rise rapidly as the "first generation" of screened and treated girls reaches childbearing age. The low-phenylalanine diet will probably be discontinued at puberty or (usually) earlier, both for practical reasons and because it may be found that phenylalanine has only minor effects, in most cases, after brain development is completed.

In principle, prophylactic dietary treatment of pregnant PKU patients should not be particularly difficult or expensive – *provided the patients are known*. This points up the importance of long-term follow-up of these individuals – many of whom may be totally unaware that they suffer from what would have been, if untreated, an incapacitating metabolic defect. What we have here is a problem in the logistics of medical information: how does the physician, knowing that his patient is at risk of producing an abnormal child 10 or 20 years from now (by which time she may well have moved a thousand miles away), ensure that she and/or her then physician have the information needed to avert the risk? The answer is certainly not a simple one – but it is relevant not merely to PKU but also to other genetic diseases, to the extent that they too are correctible and liable to affect the patient's offspring.

Considering only the United States, well over 3,000,000 neonates are now being screened for PKU annually; this, at a generous estimate, represents an outlay of something like $10 million a year. Compared to the tens of billions consumed annually by the Defense Department, or even the billions expended on space programs and superhighways, this is pretty small potatoes, but it is still big enough to draw the eyes of state legislators and budgeters seeking "frills" whose elimination can reduce the admittedly severe current demands on their funds. The first attempt to discontinue PKU screening occurred in Illinois two years ago, and though it was beaten back by the efforts of parents and physicians, similar attempts can be expected as the budgetary pinch on state governments grows tighter – which it seems likely to do over the next few years at least.

It is therefore necessary to say flatly that – and I choose my words carefully – discontinuance of PKU screening would represent not merely medical but also fiscal irresponsibility of the grossest sort. The detection and treatment of one case of PKU represents an outlay (assuming it is the outcome of 10,000 screening tests, each costing 50¢ to $4.00 in the U.S.) of up to $50,000; but failure to detect that case means a child that must almost certainly be institutionalized for the rest of its life, representing an outlay of at least $250,000 (this assumes an average life span of 50 years and an annual expenditure for custodial care conservatively estimatable at $5,000). The $250,000 figure includes no allowance for the future earnings of the treated case, or of the tax income from such earnings. It does not allow for the suffering experienced by the parents of a permanently retarded child – which is not, in any case, quantifiable in terms that are likely to make sense to a budgeter. We are not talking here about human values but purely economic ones: whether it is better to spend $50,000 now or five times that sum later.

There is, for that matter, good reason to expect that the $50,000 figure can itself be somewhat reduced. One approach to this is through the use of

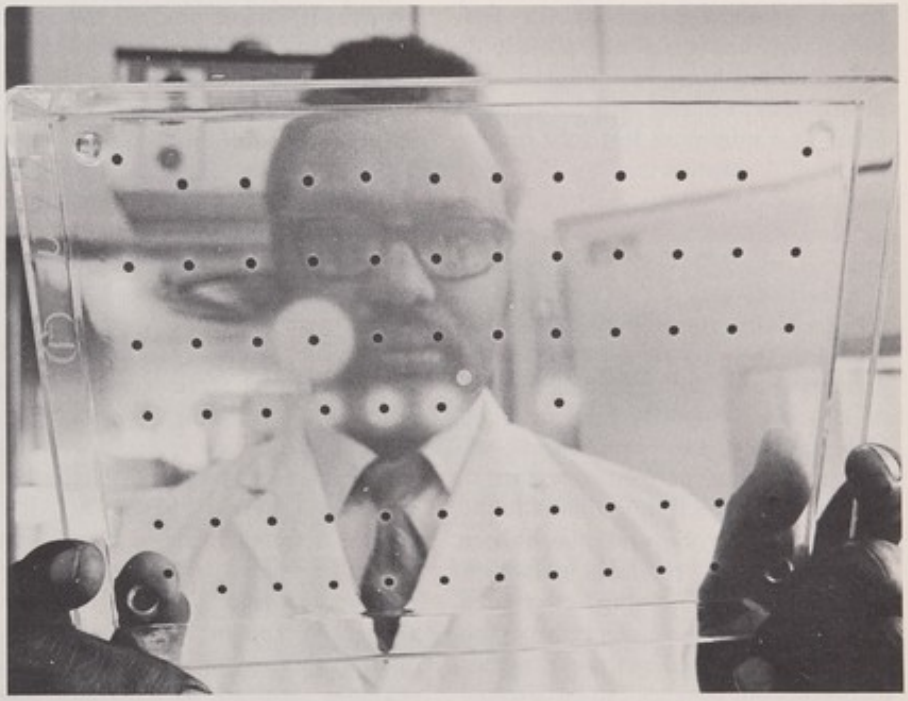
automated test equipment, of which I shall have more to say in a moment. However, to get full value from this apparatus, testing must be done on a large scale: a minimum of 25,000 births a year. This constitutes a powerful argument for regionalization of screening programs; given the validity of tests based on dried blood spots, which can be mailed in to the testing center, there is no good reason why any hospital should find it necessary to carry on its own, small-scale, screening program. The centralized, consolidated programs also possess advantages from the standpoint of staff morale. A program covering only a few thousand births per year can easily run for several years without turning up a single true case of PKU, so that the personnel involved may acquire the feeling they are engaged in a pointless exercise. When the screening load reaches the level of 50,000 or 100,000 per year, however, the expectable number of cases will provide tangible evidence that the program is performing a valuable prophylactic function.

It would be misleading, however, to suggest that the cost of screening is likely to be markedly lowered by even the most rational organization of these programs. The main cost factor is, and will remain, the actual gathering of specimens and the necessary follow-up on positive tests; the actual processing is a relatively minor item. Authorities in New Zealand, for example, find that with automated equipment the entire testing load – some 50,000 New Zealand births per year, plus several thousand others mailed in from various Pacific islands – can be easily handled by one person, whose work includes preparing and "seeding" the culture trays, placing the test discs on them, and reading the results. The economics of mass screening will continue to rest, as they do now, on the demonstrated fact that prevention is cheaper than nonprevention.

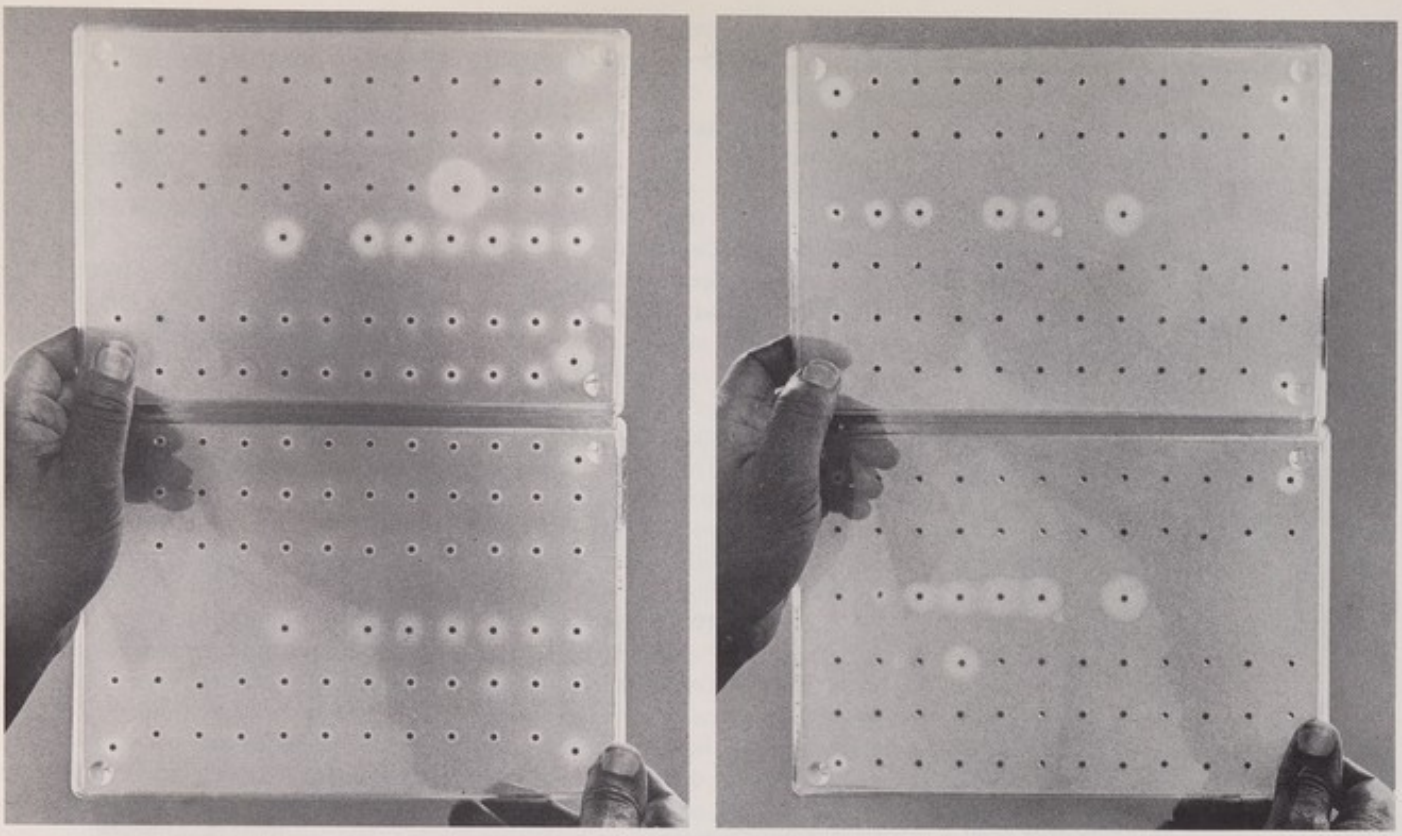
The automated equipment grew out of the finding that the inhibition assay procedure can be adapted to the detection of several other hereditary aminoacidemias simply by varying the inhibitor dissolved in the culture medium. These conditions include valinemia, tyrosinemia, histidinemia, homocystinuria (in which we test for methioninemia), and maple syrup urine disease (leucinemia). With automation, all these tests plus the phenylalanine test can be carried out at no more cost than the latter alone, unautomated.

The machine, called a punch-indexer, handles test specimens in batches of 56. It automatically punches out four ⅛-inch discs from each specimen blood spot and transfers them to predetermined locations on four different culture trays, each containing a different inhibitor – in each case, of course, the same spatial location for a given specimen. The testing array is completed with four control discs, one in each corner, as a check on uniformity of the medium, and a row of up to 12 control discs across the center of the tray, which contain blood (taken from the "expired" bottles of a blood bank) loaded with several standard concentrations of the amino acid in question. In PKU testing, for example, we use discs with normal, 2, 4, 6, 8, 10, 12, and 20 mg%; since each of these produces a growth zone in the culture of differing width, any positives in the test specimens can be approximately quantified immediately by visual comparison with the range of control discs.

Much the same mechanized procedure can be used for large-scale testing to unearth other metabolic defects, though the biologic and chemical rationales differ somewhat. Galactosemia, for instance, can be detected by the metabolite inhibition assay test, here employing a mutant strain of *E. coli*. (The principle for this test was suggested by Dr. Kenneth Paigen, Roswell Park Memorial Institute, Buffalo, N.Y.) This organism has the same metabolic defect as human galactosemics, so that galactose (or galactose-6-phosphate, in another form of the disease) accumulates in the cells, inhibiting growth. The diameter of the "zone of inhibition" (as opposed to the "zone of growth" in the tests described earlier) is proportional to the amount of galactose in the specimen disc. The same test can detect valinemia, since *E. coli* growth is inhibited by valine also. Galactosemia of one type can also be diagnosed by the Beutler test, which is chemical rather than biologic and measures enzyme activity directly rather than the accumulation of an unmetabolized substrate. The Beutler test involves a reaction between the enzyme galactose uridyltransferase,



Chief technician Emerson Barr examines agar tray that, after incubation overnight, shows one positive PKU result requiring follow-up of the infant. This can be seen in row 3 from top; "halo" surrounding disc indicates that phenylalanine in blood sample prevented the usual inhibition of growth of B. subtilis *by the β-2-thienylalanine present in medium. Row 4 contains control discs with known concentrations of phenylalanine.*



Shown above are results of additional bacterial inhibition assays. Trays at left were tested with phenylalanine (top) and tyrosine (below; positive can be seen in row 5, below control discs; each tray corner also has control disc). Trays at right show (top) tests of leucine (for maple syrup urine disease), with no positives, and (below) methionine (homocystinuria, one positive in row 4).

present in normal erythrocytes, and a synthetic substrate, the product of which is fluorescent. For the test, blood-specimen discs are placed, by the same machine, in "dimples" in a plastic tray, each of which serves in effect as a miniature test tube for the reaction. A solution of the substrate is added and, after incubation, a failure to fluoresce points to enzyme-deficient cells.

Direct assays of enzyme activity are generally considered superior to assays of the enzyme substrate, since they measure the metabolic defect itself rather than its consequences. Unfortunately, however, enzyme assays cannot be performed in PKU, since the enzyme in question is found only in the liver. Several other inborn errors, however, resemble galactosemia in that the "missing" enzyme is a normal constituent of erythrocytes, and is not destroyed by drying, so that its presence or absence can be detected in blood-spot samples.

Two of these tests, developed by Dr. William Murphey in our laboratories, are of the Enzyme-Auxotroph Assay type, employing mutant strains of *B. subtilis*. One strain, for example, requires exogenous arginine for growth; it cannot produce the amino acid from a medium containing the substrate argininosuccinic acid. Normal erythrocytes, however, contain the enzyme argininosuccinic-acid lyase, which liberates arginine from that substrate, permitting the organisms to grow. If the organisms fail to grow, the enzyme is absent from the specimen, which is to say that the individual suffers from argininosuccinic aciduria. A similar test with a different bacterial strain can detect absence of orotidine-1′-phosphate decarboxylase, which is the underlying defect in orotic aciduria. Altogether, some 20 different enzyme activities can be detected in dried blood from normal neonates by one or another means. For most of these there are as yet no known instances of congenital absence or inactivity of the enzymes involved, so there has been no attempt to adapt the tests to mass screening.

In addition to clinical conditions produced by enzyme deficiencies, there are also conditions resulting from a deficiency in an enzyme inhibitor. Recently, Dr. Murphey has developed two simple and accurate fluorescent spot tests, applicable to the dried spots of blood, for inherited deficiencies of enzyme inhibitors. One of these, inherited as a Mendelian autosomal dominant trait, is hereditary angioneurotic edema, associated with a deficiency in C1 esterase inhibitor. Although several hundred cases have been listed in the literature, mass newborn screening reveals it to be a very rare condition. The second condition, α-1-antitrypsin deficiency, inherited as a Mendelian recessive trait, appears to be more frequent, since it can be associated with at least 1% or 2% of all cases of emphysema (see Sharp, Chapter 13), and, moreover, with infantile liver disease. Its true frequency remains to be determined, since routine screening of newborns for this condition has only just begun.

It is likely that other enzyme deficiencies exist but are sufficiently uncommon to have thus far escaped detection. Alternatively, such defects could be not uncommon but associated with little or no impairment of physio-

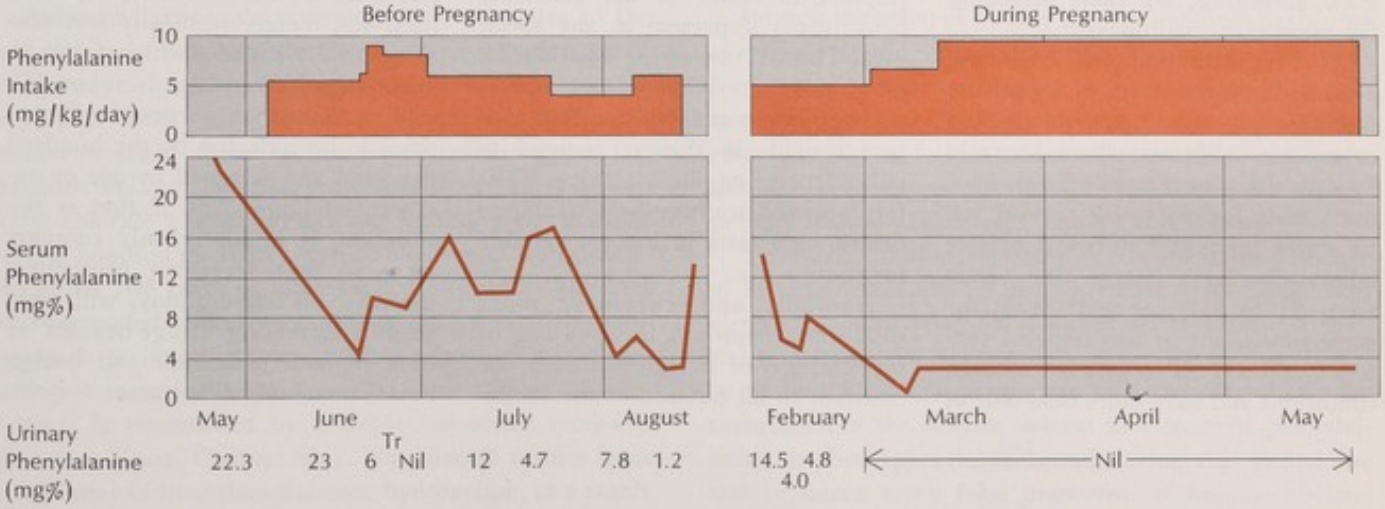
logic function. This appears to be the case, for example, in many or most types of glycinuria, several cases of which have been discovered accidentally in the course of other, unrelated investigations. Because clinical and research attention has, for obvious reasons, been focused on pathologic errors of metabolism, we have doubtless overlooked many anomalies whose functional effects are obscure or benign. Yet I believe that with increasing knowledge it will become clear, as Garrod suspected more than 60 years ago, that pathologic inborn errors are in fact only exaggerations of the innate chemical differences present in all of us.

Recently, a practical method of using the newborn dried blood spot (or specimen) in mass screening for sickle cell hemoglobin and other hemoglobin variants has been developed by Dr. Michael Garrick in our laboratory. This method uses discs punched and placed by the punch-index machine in the dimples in the plastic dimple-tray. These discs are then eluted in the dimples, after which the eluates are used for electrophoretic separation of hemoglobin bands in cellulose acetate strips. A follow-up procedure, using agar electrophoresis, can differentiate the S A heterozygote from the S S homozygote. Using this procedure, filter paper specimens for the PKU test can be screened by a laboratory already performing multiple tests on the specimens at a material cost of $0.03 per specimen. The capital investment for equipment is less than $1,000 to acquire the capability for screening up to 100,000 specimens per year.

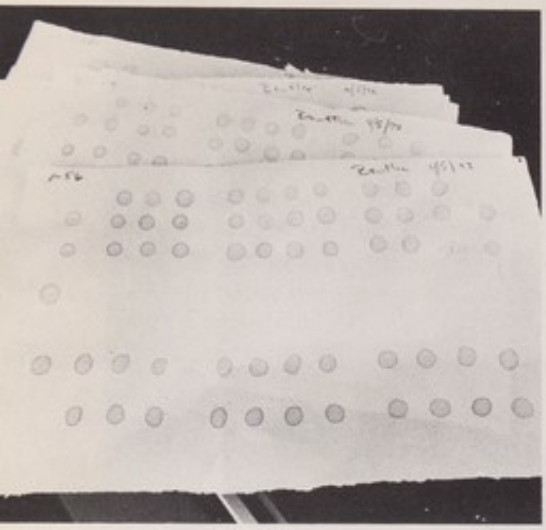
This screening procedure is the first one to be developed by our laboratory that permits detection of the "carrier" heterozygote by mass screening. It is also the first procedure developed anywhere, to our knowledge, that permits detection of S hemoglobin in the newborn infant by mass screening, in spite of the presence in such specimens of more than 90% fetal hemoglobin. Because of the recent increased interest in prevention and treatment of sickle cell disease within the United States, it is hoped that this test will receive a large-scale trial in the near future. During the past five years the screening laboratories collaborating with me have demonstrated that, with automation, six to eight of the tests described above can be carried out at little increase in cost over the single PKU test, providing a sufficient volume of specimens are available. The minimum is approximately 25,000 per year; increasing volume further reduces the cost and produces other advantages. For these reasons, we have recently been proposing development of model, or demonstration, "regional" centers for inborn errors of metabolism.

By "regional" program, I mean, first of all, one in which a sufficient number of newborn specimens are screened daily to permit application of those methods of automation already routine in a number of laboratories for the purpose of carrying out at least 8 of the 13 tests available.

Even more important for such a program is close and continuous liaison between the screening center and the medical follow-up and management of the cases detected in screening. Several programs with these features have been developed in the past few years in Europe and in New Zealand. In fact, although approximately 24 multiple-test laboratories, using four or more of my tests on the dried spots of blood collected for PKU screening, have come into existence since 1964, almost all of these are located in other countries except for the original four in the U.S., which have been working with me as part of a joint collaborative effort. However, the health departments of Ohio and Maryland recently purchased punch-index machines and have now begun multiple testing. This occurred as a



That control of phenylalanine levels in the pregnant PKU woman can ward off mental retardation in her offspring is indicated in records above of patient reported from Dublin by Allan and Brown. This patient had borne three children observed in school to be mentally retarded; her urine was strongly positive for phenylketones, that of the children negative. Detailed investigation found no explanation other than high intrauterine phenylalanine for the mental deficiency in the children. In the first phase of the study graphed above, efforts were made to determine whether a low-phenylalanine diet could effectively reduce the mother's serum levels; this was accomplished during three months of hospitalization (spikes in graph at left reflect breaches in the diet when a night attendant "took pity" on the patient and gave her a glass of milk). When she subsequently became pregnant, she reentered the hospital for dietary maintenance during the last five months. At birth the fourth child had a cord blood level of 7.6 mg% phenylalanine; at 6 hours the serum level was 5.2 and at 48 hours a normal 1.4. Tests later showed normal intelligence.



Beutler test for one type of transferase deficiency assays enzyme activity on a synthetic fluorescing substrate added to discs; spots that do not fluoresce under UV light after incubation are positive.

result of a conference we held in February 1971.

The three principal problems in our own country that have impeded expansion of PKU test programs to include tests for other conditions are:

1) Lack of liaison between screening centers and medical centers;

2) Restrictions of screening areas by state boundaries, as states often contain too small a population of newborn infants per year for efficient PKU screening, let alone screening for more rare conditions;

3) Fragmentation of screening, even within large states, when private facilities are used. An outstanding example is California, where, I am told by a health department official, newborn tests for PKU are carried out by more than 200 different private laboratories at a charge that varies from $1 to $15 per test! With the same funds, one or two regional centers could test all infants for a dozen conditions and have more than half of the money left over to assist with medical follow-up and management.

If a continued search for these and other rare or obscure biochemical anomalies has (as yet) no "practical" value, in the sense that they involve either diseases for which no treatment exists or a disease that need not be treated, it nonetheless seems to me to possess great scientific importance. I would say it has enough scientific value to encourage mass screening for as many biochemical anomalies as possible, always assuming that this can be done without dipping into funds needed for more urgent aspects of health care.

An obvious approach to this is to expand the "six-for-the-price-of-one" achievement that automation has already brought to screening into a "20-or-30-for-the-price-of-one" arrangement. Already our laboratory is working on several approaches to this goal; all of them involve not simultaneously testing several specimens from the same individual for different anomalies, as is done with the punch-index machine, but testing a single specimen simultaneously for several anomalies.

One possibility is a "multiple inhibition assay," involving the addition of two (or more) antimetabolic analogues to the culture medium in such concentrations that the culture would respond to *either* of the abnormal constituents if present in any specimen. Thus, if one added 4-azaleucine plus β-2-thienylalanine at appropriately low concentrations, the culture should, in theory, respond to either leucine or phenylalanine. Thus far, however, we have been unable to devise such tests in a form suitable for routine use.

Another, and currently more promising, approach involves two or more strains of *B. subtilis*, each one resistant to all but one of the inhibitors used. Again a response is produced to any of several abnormal serum constituents. Yet another approach, which we call the "multiple auxotroph" test, again employs mutant strains, but this time each of the mutants requires a different amino acid for growth. A mixture of strains requiring, respectively, phenylalanine, leucine, and histidine, for instance, would respond to PKU, maple syrup urine disease, and histidinemia. A cardinal difficulty here is that the mutants can interact when mixed, for reasons still obscure, producing either inhibition or stimulation of growth. However, two such multiple tests appear promising.

In all these instances, of course, a positive response could mean *any* of the several conditions the test was designed to respond to, with differential diagnosis requiring further tests. These, however, would involve only a few score or at the most a few hundred specimens, as against the tens of thousands eliminated by the initial multiple screen.

Given the further development and perfection of multiple screening tests, mass screening can become not merely a way of rapidly detecting PKU and other treatable genetic diseases but also a source of invaluable information on biochemical differences in large populations. Given the fact that screening for PKU alone has proved itself not merely medically but also economically sound – which seems to me unarguable – we can (one hopes!) safely assume that specimens will continue to be collected by the hundred thousand and screened by one or another technique. And if that is the situation, it is surely only common sense to seek screening procedures that, for the same outlay, will yield steadily increasing "fringe benefits" of data on both pathologic and benign innate metabolic differences.

23

A Critical Evaluation of PKU Screening

DAVID YI-YUNG HSIA* *Loyola University*
and
NEIL A. HOLTZMAN *Johns Hopkins University*

The era of mass screening for hereditary metabolic defects in the newborn has only just begun. At the moment, the history of this effort can be told almost entirely in terms of tests for phenylketonuria, the first genetic disease to be screened for en masse and by law in most of the United States.

As the first chapter in what surely will be a growing chronicle, the PKU story is instructive on many issues, scientific and socioeconomic. As our society becomes increasingly aware of genetic diseases and potentials for control, a critical review of PKU-testing history may be timely.

Although large-scale screening for PKU dates to the mid-1950's in England, the disease itself was discovered in Norway some 20 years before. Seeking the reason for a peculiar musty odor in her two retarded children, a mother brought them to the physician-biochemist Dr. Ashborn Følling. He discovered that their urines turned ferric chloride a bright green. Dr. Følling then surveyed several hundred mentally retarded children in institutions. He found two pairs of siblings and four other children who responded similarly on urine testing. An entirely new approach to investigations of the etiology of mental retardation ensued.

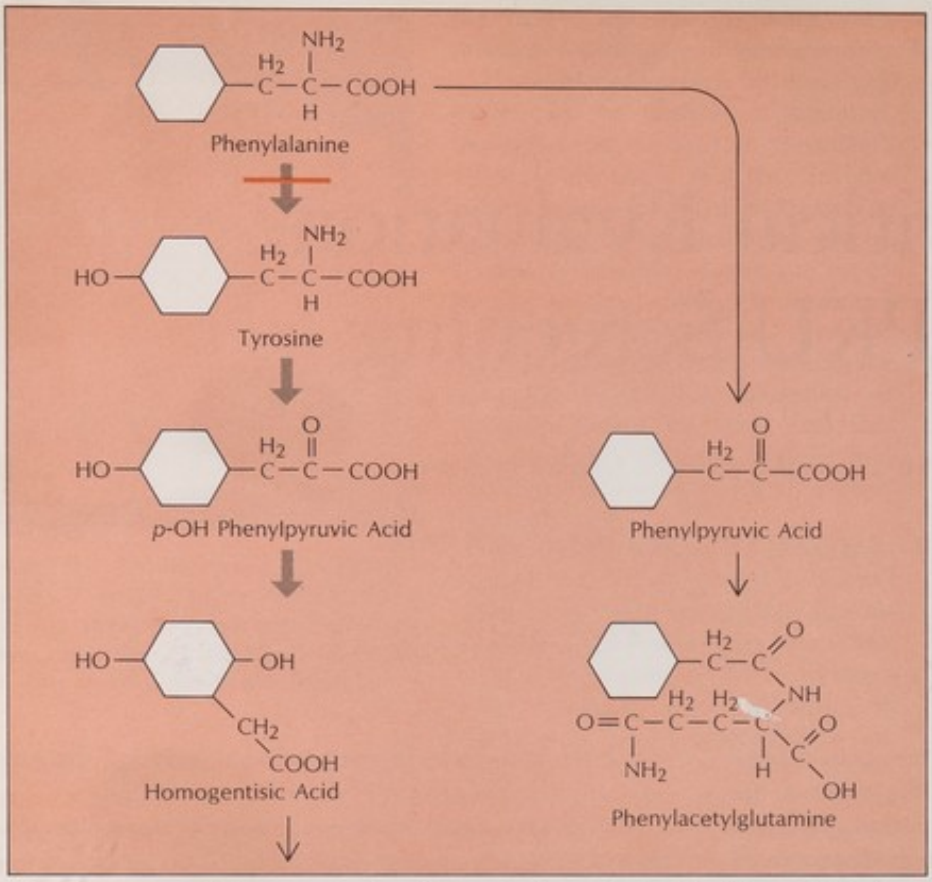
PKU is transmitted by a single autosomal recessive gene (see Knox, Chapter Six). The disease results from a deficiency of liver phenylalanine hydroxylase, as a result of which an excess of phenylalanine builds up in the blood and spinal fluid. It is not yet clear at what stage in human intrauterine development the phenylalanine hydroxylase enzyme system fully matures; due to placental transport the fetus with PKU may be spared damage. Once delivered and exposed to protein, the baby with PKU develops central nervous system damage, although its manifestations may not be appreciated for more than a year. The infant with PKU appears to be most vulnerable during the period of rapid brain growth in the first years of life.

The discovery that timely application of a low-phenylalanine diet could reduce the newborn's blood phenylalanine levels was made in 1953. Subsequent experience with dietary treatment has made it clear that if phenylalanine intake is not restricted by one to three months of age significant irreversible brain damage will occur. On the other hand, as the Collaborative Study of Children Treated for PKU has demonstrated, inception of dietary therapy before three months of age will result in normal IQ scores at least at age four.

Prompt initiation of dietary therapy requires having an effective screening test. Urinary screening studies with ferric chloride were undertaken in Birmingham, England, in the mid-1950's. The test was tried first in 50,000 older retarded children in institutions. Along with other efforts, this screening suggested that up to 1% of institutionalized persons with mental defects were PKU patients.

The ferric chloride test becomes positive in the presence of increased amounts of phenylpyruvic acid. Substances elevated in other inherited diseases can also cause a positive test. Applied to newborns, however, the test was found to be faulty. Several weeks are required after birth for maturation of the enzyme system that converts phenylalanine to phenylpyruvic acid, and during this period the test produces many false negatives. It became obvious that urine screening could only be done effectively at about one month of age. Even then, the test may yield a false negative due to the instability of phenylpyruvic acid. An-

* *The basis for this chapter is an article written by Dr. Hsia before his death. It was revised and brought up to date for publication here by Dr. Holtzman.*



Normal main metabolic route of phenylalanine is shown in simplified form at left above (heavy arrows); in PKU an inherited deficiency of phenylalanine hydroxylase blocks oxidation of phenylalanine to tyrosine. An alternate pathway through which the body may adjust to the unusual amounts of phenylalanine metabolites is also shown.

other problem arose: the first Birmingham studies showed that the chance of catching up with an infant for a test at one month of age was at best about 70% and probably much less.

The test received a mixed reaction in the United States. Some pediatricians offered it as a nicety of office practice. Between 1955 and 1964, perhaps as many as 40% of the children born in this country had the urine test. Large-scale urine testing programs continued in England until 1964. In reviewing the efficacy of the test a committee of the British Medical Research Council reported that as many as 45% of infants proved to have PKU were missed by ferric chloride screening. Today the ferric chloride test should be used only to check children who have been tested by other methods.

The most prominent successor to the FeCl test is the bacterial inhibition test devised and championed by Dr. Robert Guthrie (see Chapter 22). This blood assay obviates the delay inherent in urine testing. It tells directly whether or not blood phenylalanine is suspiciously high. A drop of blood from the neonate's heel is applied to a filter paper and air dried. A disc is punched from the blood spot and placed in an agar medium containing *Bacillus subtilis* and β-2-thienylalanine. If bacterial growth appears at the spot after 18 hours' incubation, the patient is presumed to have PKU. The basis of the test is the known fact that β-2-thienylalanine inhibits bacterial growth unless phenylalanine is present in the blood.

This semiquantitative test was devised in 1961. Dr. Guthrie won support from the U.S. Children's Bureau for a large screening study in which many hospitals and state health departments cooperated. By 1964, almost 405,000 newborns had been tested. About 275 were found to have at least 6 mg% blood phenylalanine, a presumptive positive result. Of these, 37 were found to have PKU by quantitative methods. A quantitative method is necessary not only to confirm the biochemical diagnosis but also to establish the proper phenylalanine content of the individualized diet. Paper chromatography, enzymatic methods, and spectrofluorimetry have been used to confirm Guthrie-test results and for dietary adjustment. In some places, spectrofluorimetry is employed for screening in place of the Guthrie test.

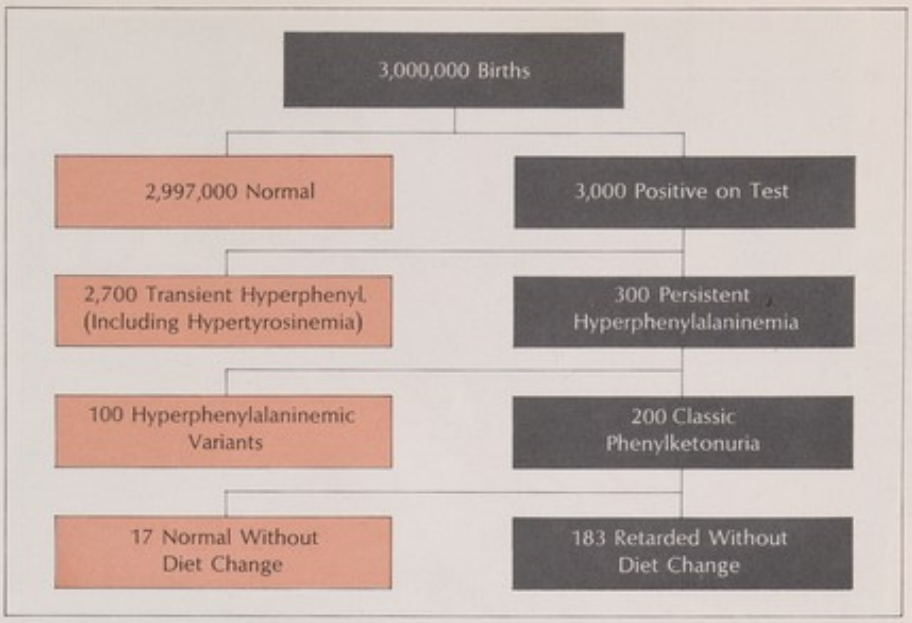
Unfortunately, the experiences with the first several hundred thousand children were never published authoritatively. Recent surveys indicate that a large proportion of infants with elevated screening tests turn out not to have PKU. Moreover, as many as 10% of infants with PKU may be missed by Guthrie screening.

It is now apparent that in infants with PKU the blood phenylalanine may be normal on the first day of life, usually rising to abnormal levels by the third or fourth day and to striking elevations by the eighth day. The amount of protein ingested and retained probably influences the rate of rise. In the United States more infants are tested on the third day of life – just before nursery discharge – than on any other. This may be too early for some infants with PKU, particularly females. A sex ratio of 30 females and 60 males, far in excess of the 100:106 ratio expected, was found in the initial sample of the Collaborative Study of Children Treated for PKU. This initial sample included all cases found by Guthrie screening (and confirmed) between October 1, 1967, and December 31, 1969, in participating states. Although the ratio of males to females has dropped among infants subsequently admitted to the study there is still a significant excess of males. Serial phenylalanine determinations during the first week of life in infants with PKU suggest that the rate of rise of blood phenylalanine in female phenylketonurics is slower than in males. There is no evidence to support another hypothesis: that among the infants meeting criteria for inclusion in the study relatively more males than females will have benign, variant forms than have classic PKU. In infants who are discharged early, in whom the test is negative, a repeat blood assay or, if this is not possible, a urine test one month after discharge, is a wise step.

The results of Guthrie-test screen-

ing so impressed the public – particularly parents of PKU children – that, in 1964, a movement to make screening for PKU compulsory by state law was launched. In the next two years, 43 states adopted such laws. Typically, these laws declare it to be state policy to test every infant within a reasonable period after birth. The health department, attending physician, or the hospital is made responsible for testing. (Generally, parental religious objections are among the few exceptions permitted to mandatory testing.)

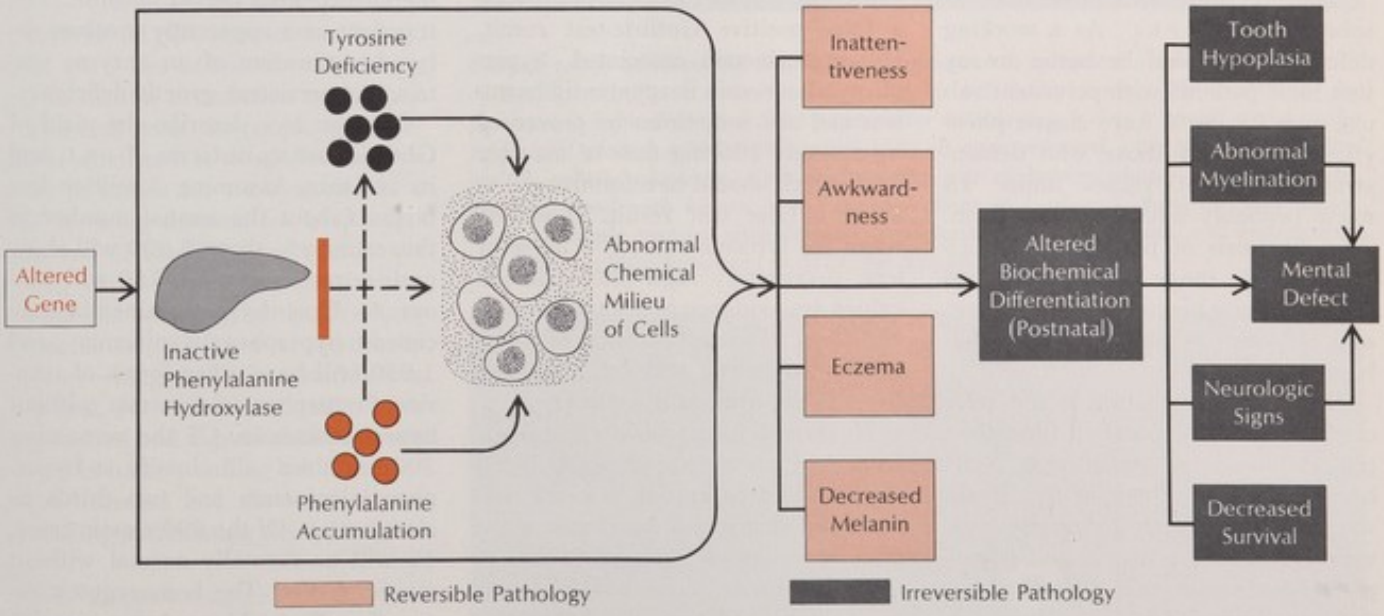
Screening seemed to show an increase in PKU frequency. In reality, however, what was being uncovered were forms of hyperphenylalaninemia milder than classic PKU. Controversy over quantitative criteria for discriminating between "dangerous" and "innocuous" forms, not entirely settled today, threw doubt on the worth of the testing programs.



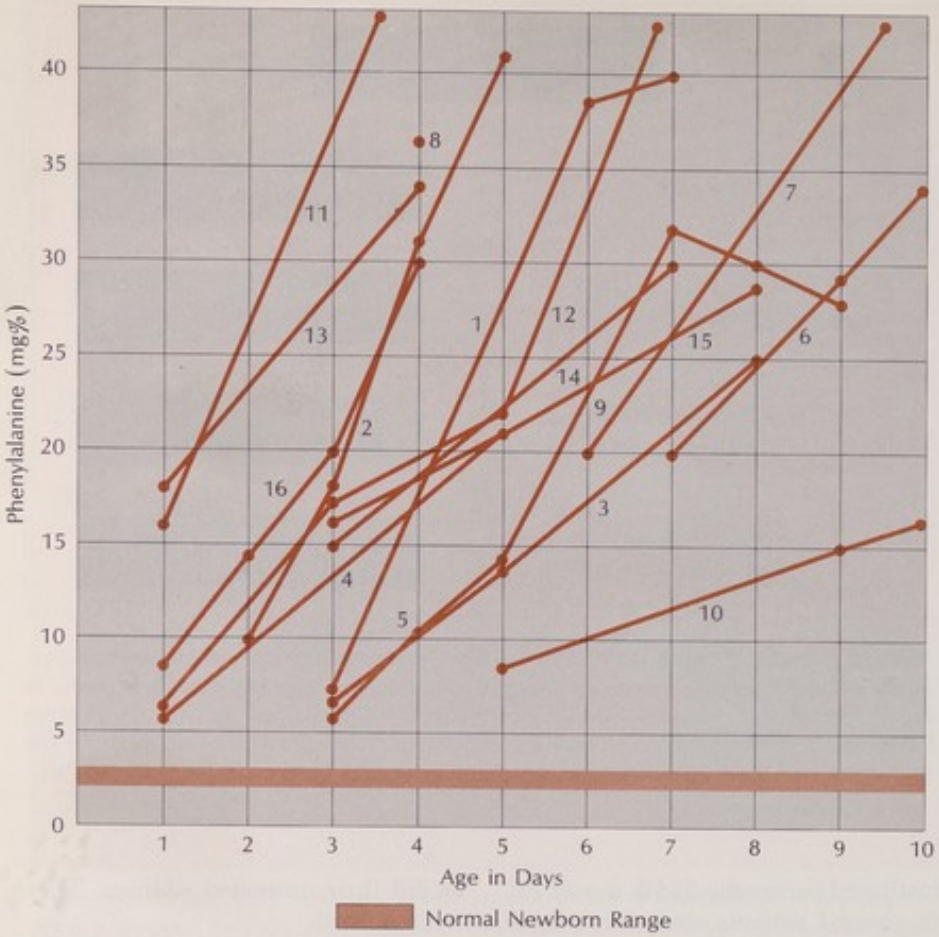
With large-scale screening for PKU, it became apparent that phenylalanine elevations in the newborn may be due to a variety of causes. Assuming 3,000,000 live births in the U.S. each year, about 0.1% will have a positive Guthrie test, but in 90% of these babies the high phenylalanine simply accompanies a transient hypertyrosinemia. Of the remainder with true phenylalanine elevations, only two thirds will have classic PKU.

Recent surveys of health departments in the United States and Canada reported the incidence of patients with values of 20 mg% or more (designated below as group I) to be about 1 in 19,000. The incidence of values between 10 and 19.9 mg% was 1 in 33,000 (group II below). The combined incidence of 1 in 12,000 was much higher than estimates based on the frequency of retardation. Diet instituted before the 335th day of life in group I patients seemed to permit normal mental development and avoid the mental retardation found in untreated or late-treated group I siblings. However, group II patients, whether on the diet or not, had the same degree of mental development as did their untreated siblings. They were normal.

Variants of PKU have been given a variety of names, such as "hyperphenylalaninemia without phenylketonuria" and "atypical PKU." Compared with classic PKU, patients with the variant forms tolerate greater phenyl-



The mechanisms through which the basic defect in PKU the absence of phenylalanine hydroxylase – operates to produce both reversible and irreversible pathology are schematized above. As phenylalanine compounds build up, the levels of other amino acids are depressed, and the developing brain is deprived of nutrients needed for normal myelination and tissue formation.



Necessity for testing phenylalanine values only after child has had adequate protein intake is suggested in graph above. Although all 16 children had values suspiciously above normal on initial testing, PKU levels were reached only after three or more days of life.

alanine challenges. Their plasma values rise slowly after birth and remain lower throughout later years.

There is no plasma value that absolutely defines PKU. As a working definition, it would be better to say that most patients with persistent values over 25 mg% have classic phenylketonuria, and those who demonstrate persistent values under 15 mg% probably have a variant form. The diagnosis of patients in the 15 to 25 mg% range depends in good part on family history, clinical course, and response to phenylalanine challenge.

But the blood picture is not fully clarified until components other than phenylalanine are considered. Newborns, especially those of low birthweight, tend to have unusually high values of serum tyrosine, and hyperphenylalaninemia may accompany hypertyrosinemia. The elevations are traceable to a transient functional deficiency of tyrosine transaminase and *p*-hydroxyphenylpyruvic acid oxidase. In general, although phenylalanine values seldom exceed 10 mg% when associated with hypertyrosinemia, this will be sufficient to produce a false positive Guthrie-test result. Hypertyrosinemia-associated hyperphenylalaninemia is apparently harmless and can sometimes be prevented by a single 100 mg dose of ascorbic acid, which should be administered to avoid a false test result. However, when the tyrosine value is 5 mg% or less, persistently high phenylalanine values are evidence of PKU or a variant form. The appearance of *o*-hydroxyphenylacetic acid and phenylketones in the urine is the clincher.

If variant forms have clinical effects, they are not apparent. As noted earlier, children appear to be of normal intelligence and the classic PKU manifestations of agitated behavior, finger posturing, along with hyperactive reflexes and muscular hypertonicity, minor choreoathetosis, and occasional true spastic paraplegia, are not observed. Untreated children with classic PKU also are of lighter complexion than other family members. Recurrent vomiting in the first month of life (occasionally associated with pyloric stenosis) and severe infantile eczema are other findings in untreated phenylketonuria but not in the variants.

Like classic PKU, persistent hyperphenylalaninemia appears to be transmitted as an autosomal recessive trait. A phenylalanine tolerance test can distinguish a group of heterozygotes from a group of normal persons, but the test is not discriminating enough to let us make individual predictions. Hence, testing parents does not allow us to predict that an offspring will have classic PKU or a variant form.

The research stimulated by screening has shown us not only the range of variant forms but also the possibility that some variant forms are transient – with good and bad implications. For example, a transient form reported first in 1967 involved a child with a positive Guthrie-test result and a blood value of 20 mg% at three weeks of age. However, at one year, the value was 3 mg% in response to protein challenge. The child has since had apparently normal mental development. The contrary situation – hyperphenylalaninemia progressing to classic PKU – may also occur and produce some mental retardation. Obviously, monitoring of patients with possible transient variants is indicated for an extended period of time. The transient case apparently involves delayed maturation of an enzyme system, not persistent genetic deficiency.

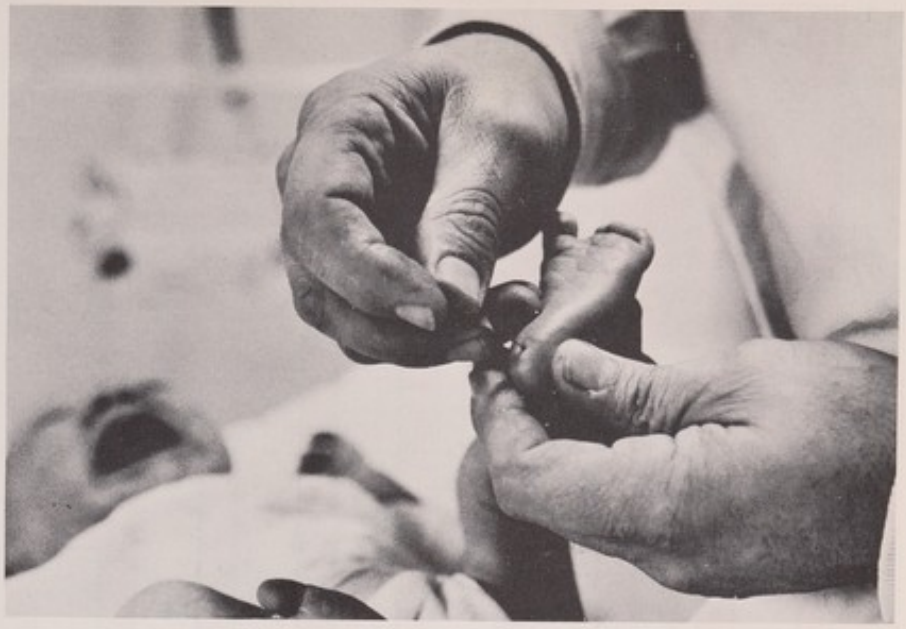
We can now describe the yield of Guthrie testing in terms of PKU and its variants. Assuming 3 million live births (about the annual number in this country), about 3,000 will show positive test results but 750 will turn out to have hypertyrosinemia-associated hyperphenylalaninemia and 1,950 will have other forms of transient hyperphenylalaninemia without hypertyrosinemia. Of the remaining 300 one third will classify as hyperphenylalaninemia and two thirds as classic PKU. Of the 200 classic cases, 17 will be mentally normal without a special diet. The homozygote's escape from mental retardation remains a mystery, but methods are lacking to identify those who might safely avoid the diet.

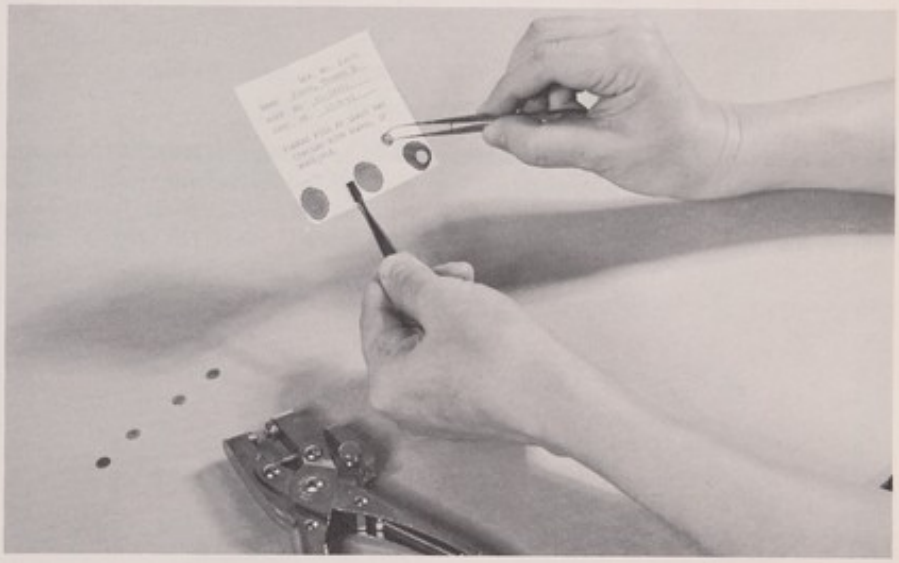
By and large, present screening, in which the neonate is tested once, enjoys the confidence of the medical community. At a meeting in mid-1970, physicians involved in PKU management agreed unanimously that such programs are effective but probably miss a few cases. Nonetheless, one participant, observing that it took from 1958 to 1964 to acquire enough evidence to throw doubt on the urine test, urged that pediatricians report to state health departments any cases missed by screening and that pediatricians test by blood assay in all cases of mental retardation and other neurologic disorders. Another recommendation was that the PKU diagnosis be verified semiannually through phenylalanine challenge. Conferees agreed that some children were on diets unnecessarily and that others, with a variant, should be placed on restricted diets and observed.

No single source offers a clear picture of the money this country spends annually on PKU diagnosis and treatment. It may be between $5 million and $10 million, the lion's share going for screening. Treatment – usually prescribed at a regional center after referral by a physician through a health department – is subsidized in many states. This includes family counseling, the work-up to establish the proper diet, and the diet, which itself may cost $60 or more monthly. The aggregate cost of treatment is not known. (Research spending probably is severalfold the cost of screening, since PKU has attracted wide attention in mental retardation research.)

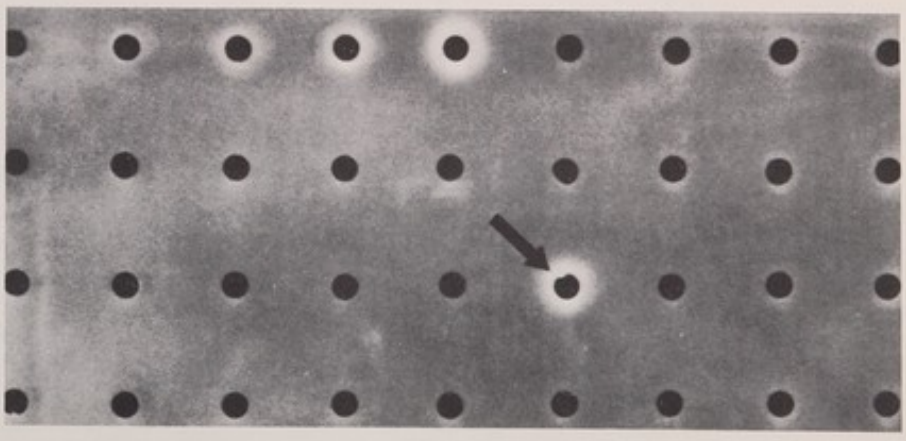
The real issue the people must decide is whether it is worth the nation's effort to screen for PKU at all. At a time of rising demand for health services and of funding shortages, PKU testing expenditures are of an order that cannot be taken lightly. We have not achieved the kind of assurance from screening that glimmered before us at first. Nevertheless, families have been spared and a great deal has been gained in understanding the nature of genetic disease and the mechanisms for finding and treating diseases of this type.

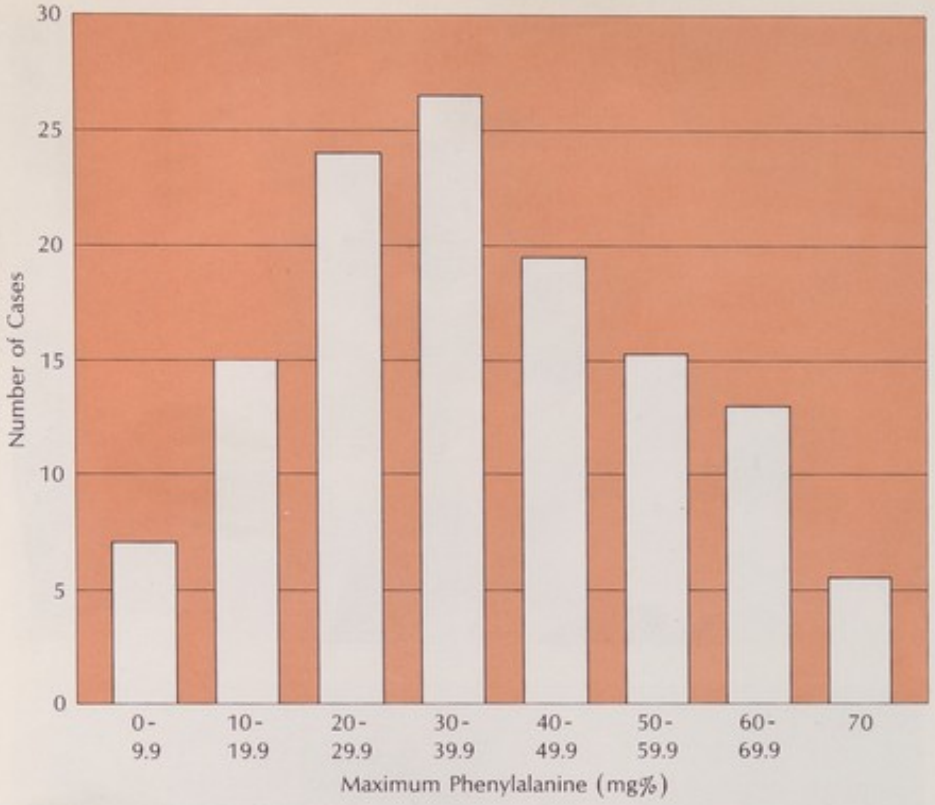
We do not suggest that PKU screening be discontinued. The consequences of *not* supporting a program that might rescue an individual from mental retardation weigh heavily in



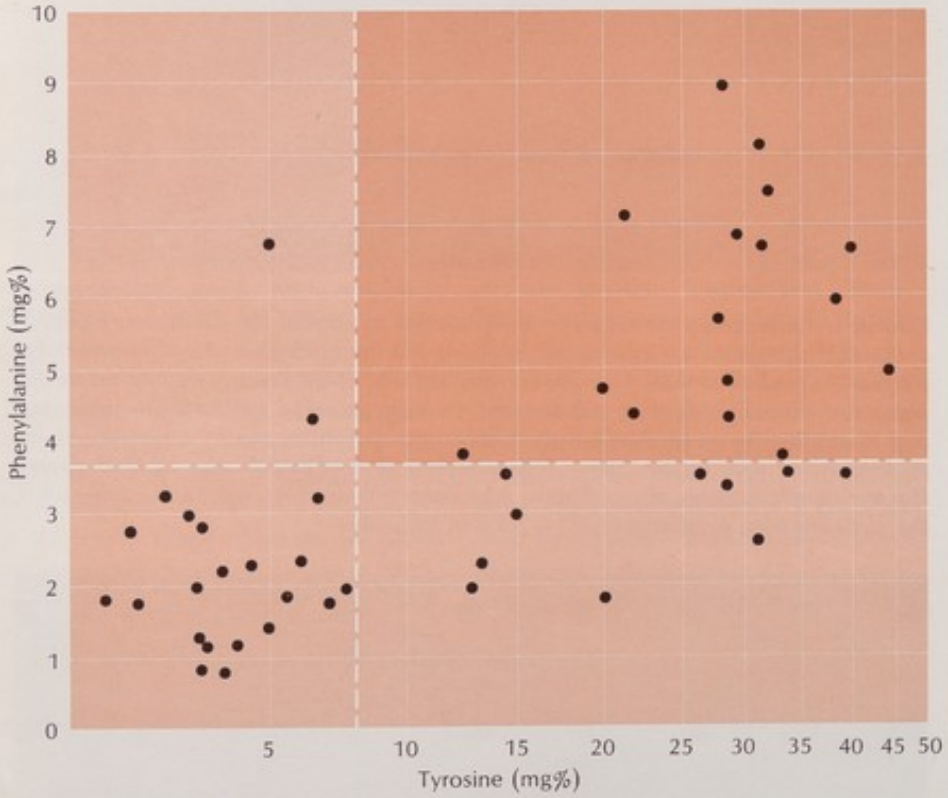


Guthrie test is based on the ability of phenylalanine to prevent the inhibition of growth of Bacillus subtilis *in a medium that contains β-2-thienylalanine. Blood obtained by puncture is applied to thick filter paper; after autoclaving, a disc is punched out of the paper and placed, with other test samples, on agar prepared with β-2-thienylalanine and* B. subtilis *inoculum. Control discs (top row of bottom photo) are prepared by spotting blood with known, rising concentrations of L-phenylalanine on filter paper. If a test sample contains phenylalanine, a bacterial growth zone (halo) will appear after the plate has been incubated at 37° C for 8 to 16 hours.*





Not all instances of high blood phenylalanine should be considered classic PKU. Graphed above is distribution of phenylalanine levels of all (nonhypertyrosinemic) patients giving a positive test result (6 mg% or more) in six states during 1946-66. In this group about 20% were hyperphenylalaninemic but not in PKU range (20 mg% or more).

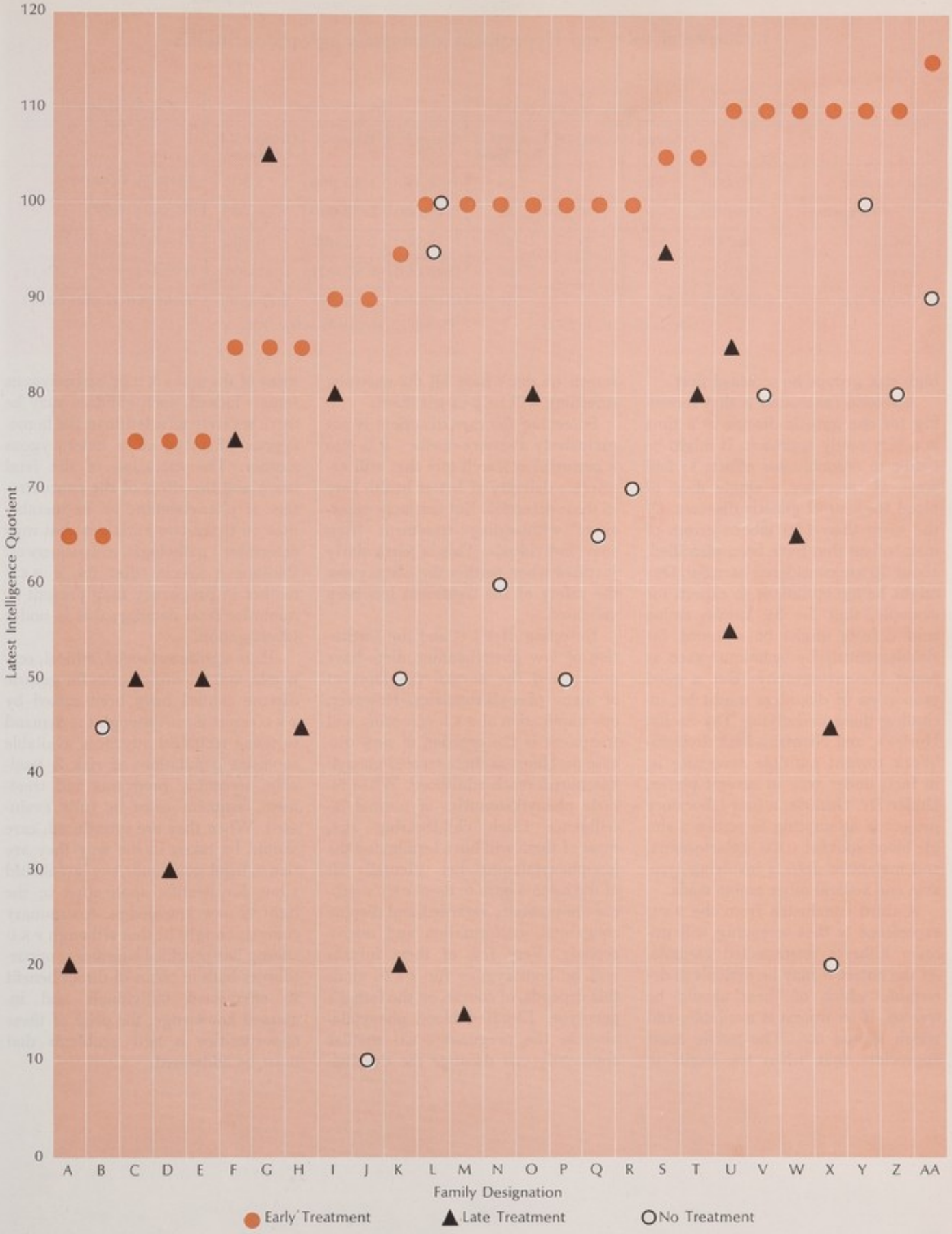


The most common cause of high blood phenylalanine in newborns is associated hypertyrosinemia, the result of a functional deficiency of both tyrosine transaminase and p-hydroxyphenylpyruvic oxidase; this is especially frequent in low-birthweight infants. Plotted above are phenylalanine and tyrosine levels of 48 such babies. Dashed lines are means plus 3 SD for full-term infants (vertical, tyrosine; horizontal, phenylalanine). This form of hyperphenylalaninemia is harmless and transitory.

the balance. We do suggest that the effort must be made more valuable. There is a growing skepticism concerning PKU expenditures in the current fiscal climate, and we believe that laymen and scientists who favor PKU testing should meet reasonable objections. From the PKU experience we conclude the following:

First, quality control of the screening procedure must improve. There is wide variation in the frequency of positive tests from state to state. A large component of this variation lies in the performance and interpretation of the Guthrie test. Consequently some states have a higher frequency of false positives while in others a greater number of false negatives can be expected. Another factor accounting for differences in the frequency of elevated tests is the nature of the population screened. PKU is less frequent among blacks but, from data in Maryland, still occurs once among approximately 25,000 black live births. With such a rare disease some areas will go for years without detecting a single patient and the procedure may appear to be wasteful. Thus the District of Columbia eliminated PKU screening in 1970. It had paid for testing 77,000 newborns in three years; not a single PKU case had been detected. The direct cost to the department for tests in hospitals was $135,000. The District Health Department, with a tight budget and massive demands on its health funds and services, justified the end of PKU screening by noting that 84 of 100 live births that it screened were nonwhite. Because the District has no PKU law, the department ended the screening without recourse to repeal legislation. In the states where PKU screening is dependent on legislative action, one could expect a far different experience through public hearings and a full legislative process.

Efforts to restrict mass screening to specific population groups may be susceptible to charges of discrimination, and PKU screening probably should be available to every infant, provided the parents consent. However, in screening for disorders such as sickle cell anemia and Tay-Sachs disease, for which blacks and Ashkenazi Jews respectively have a much higher frequency than other ethnic groups, cost benefit dictates that the



Value of low-phenylalanine diet in management of PKU children has been assessed by comparing IQ of siblings treated early (less than 30 days of age) with that of siblings treated late or not treated. As shown above, although not all untreated or late-treated PKU children are mentally deficient (and, conversely, treatment does not always prevent retardation), with one exception the IQ of the early-treated child is equal to or greater than that of his late-treated or untreated sibling.

Incidence of PKU and Hyperphenylalaninemia in Various Studies

Geographic Area	*Number of Infants Screened*	*Number With Phenylalanine Elevations*			*Incidence Rate*			*Reference*
		Group I	Group II	Group II as % of Total	Group I	Group II	Group I & II Combined	
Scotland	80,000	17	6	26	1: 4,700	1:13,000	1: 3,500	McBean (1969)
West Germany	100,000	9	3	25	1:11,000	1:33,000	1: 8,300	Bickel (1967)
Canada	36,000	2	2	50	1:18,000	1:18,000	1: 9,000	Clow et al. (1969)
Israel	178,174	7	9	56	1:25,000	1:20,000	1:11,000	Steinberg et al. (1969)
United States	6,382,283	337	193	36	1:19,000	1:33,000	1:12,000	Holtzman et al. (1973)

Group I: Classic PKU Group II: Hyperphenylalaninemia

high risk groups be screened first.

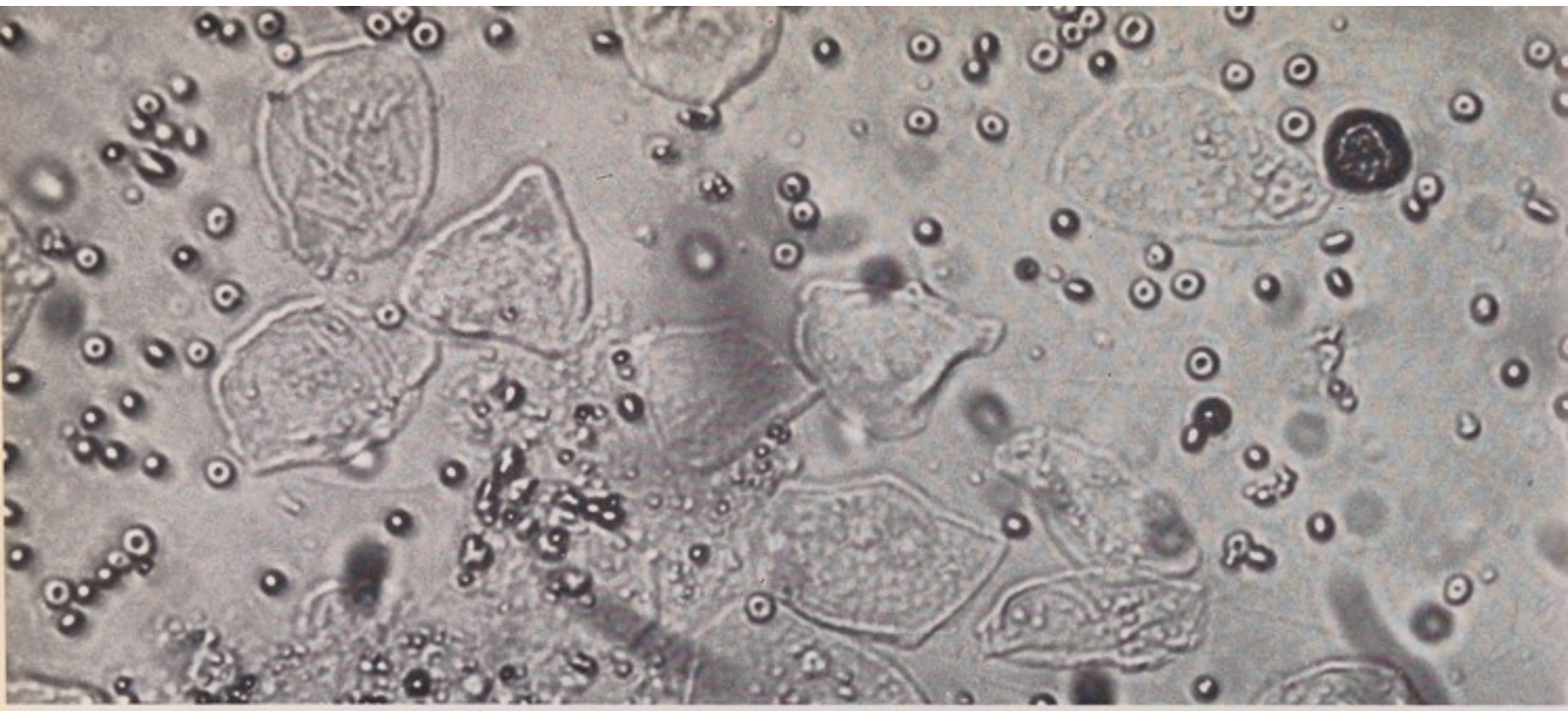
A second conclusion is that screening for one genetic disease at a time is a very costly approach. It might be better to redouble our efforts to find ways of screening a single drop of blood for several genetic diseases. Of the more than 150 inborn errors of metabolism that have been identified, about 50 are considered harmful. One might be too optimistic to expect, for example, that the 35 known amino acid defects might be screened for simultaneously by techniques such as paper chromatography. But a good proportion of disorders might be, including those underlying Tay-Sachs, Hurler's, and Niemann-Pick diseases. Work toward multiple screening is, in fact, under way in several places. Under Dr. Guthrie, a four-laboratory project is attempting to screen a single blood spot for PKU, galactosemia, and metabolic defects involving tyrosine and several other amino acids.

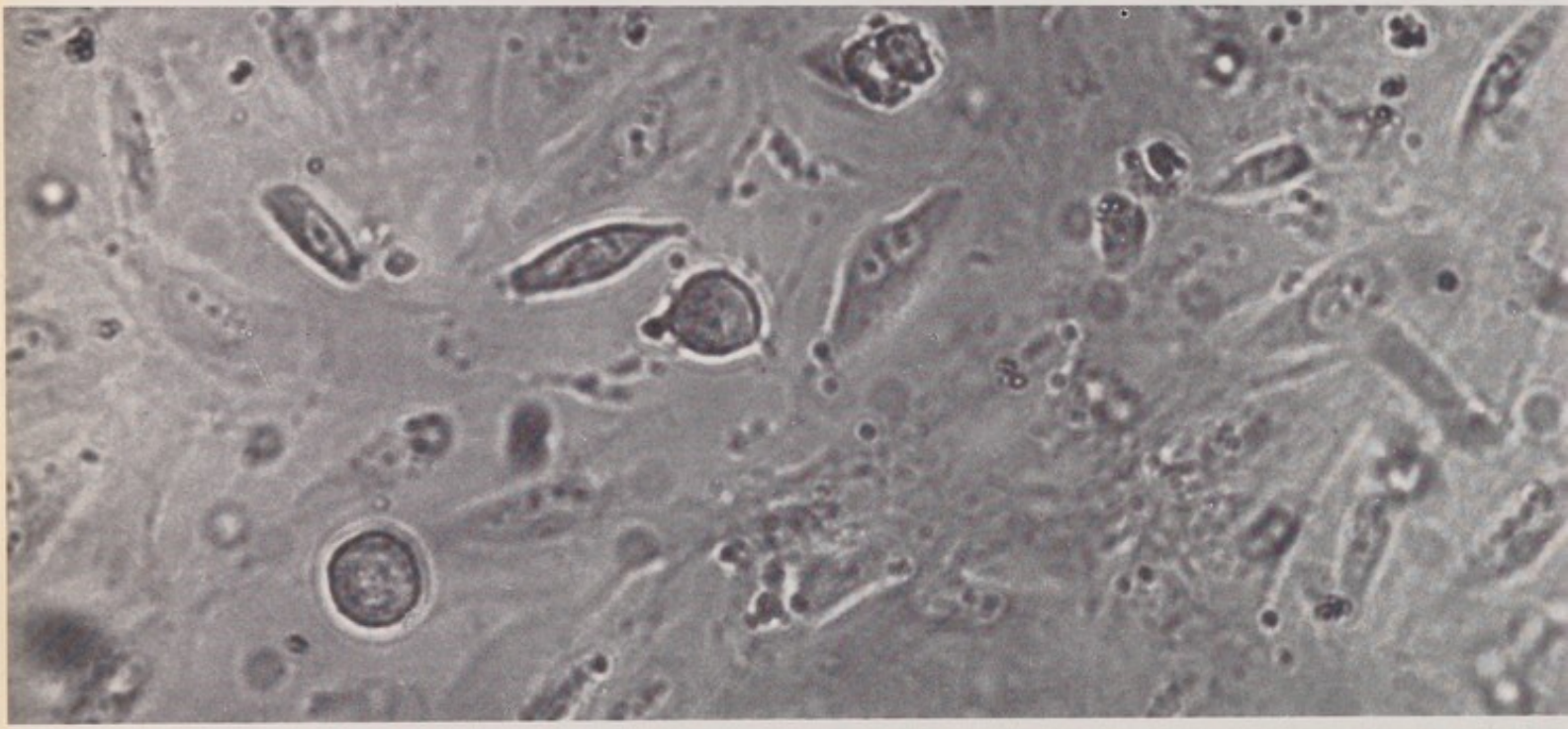
A third conclusion from the PKU experience is that screening will uncover hitherto unsuspected variants. At the outset it may be difficult to determine which of these should be treated, if treatment is available, and which should not. The public must appreciate that when we begin to screen we don't have all the answers; screening will help us get them.

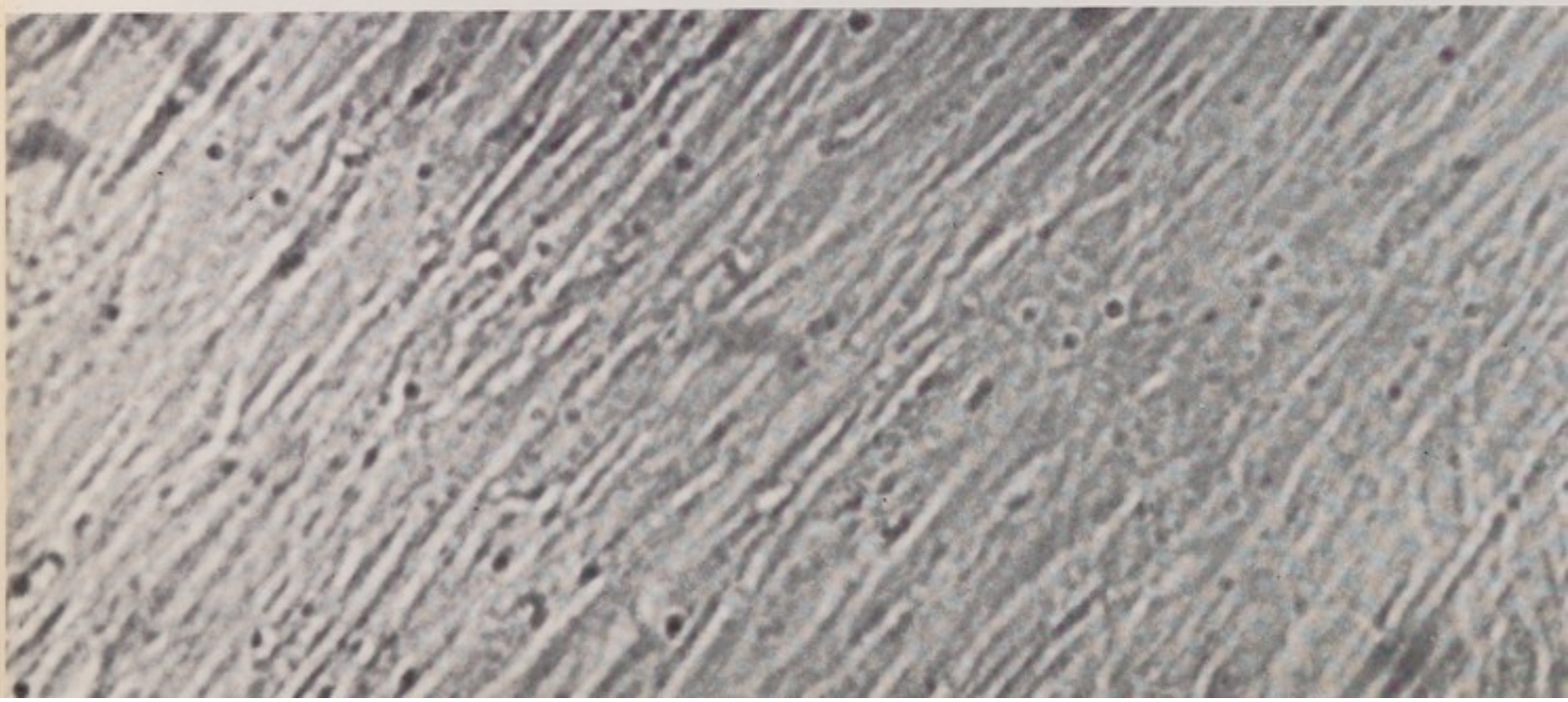
Screening for rare disorders is not exclusively a service device – it is also a powerful research tool that will accelerate delivery of better health care to those screened. But this may necessitate withholding treatment from some individuals. This is particularly justified when neither the efficacy nor the safety of the treatment has been validated.

Detection of PKU and the institution of low phenylalanine diets have guaranteed the normal development of many phenylketonurics. However, one implication of PKU screening and treatment is the creation of new disease problems as these treated phenylketonurics reach adulthood. When female phenylketonurics of normal intelligence reach childbearing age, most of them will have terminated the low-phenylalanine diet. Virtually all of the infants born to these PKU mothers are mentally retarded and display congenital malformation and microcephaly. Very few of these infants will be homozygous for PKU, since this depends, of course, on the father's genotype. Elevated blood phenylalanine in the pregnant PKU mother apparently can derange the development of the non-PKU or heterozygous fetus – indeed, such children may be more severely affected than the homozygous offspring of a heterozygous mother. The maturing of the fetal brain and the effect of the concentration of phenylalanine or its metabolites on it are two variables that may determine pathologic consequences. Perhaps a special diet for a PKU mother in pregnancy may prevent or minimize fetal damage; this is under investigation.

Thus significant social, ethical, economic, and scientific issues for genetic disease control have been raised by PKU screening. Although the demand to make technical advances available to entire populations at risk is laudable, screening programs and treatment regimens must be fully evaluated. When they are introduced, care should be taken in the way they are tied to legal sanctions. Any law should allow for flexible application in the light of new knowledge. A summary comment might be that although PKU testing has provided significant opportunities both in terms of direct benefit to affectected individuals and increased knowledge, the price of these opportunities is new problems that must be addressed.







24

The Prenatal Detection of Hereditary Defects

JOSEPH DANCIS
New York University

Until very recently the suggestion that a hereditary disorder might be detected prenatally would have drawn from most clinicians the equivalent of "why bother?" Nothing could be done about such a diseased fetus in any case, so why should anyone undertake the presumably formidable task of diagnosis in utero?

A number of developments have combined to turn a seemingly academic speculation into a promising clinical tool. In the area of hereditary defects particularly, there are clear indications that prenatal diagnosis is possible, is not enormously difficult, and can serve as the basis for effective medical intervention.

Before reviewing these findings, however, I should like to stress that, although prenatal diagnosis of hereditary disease is theoretically feasible, it has thus far been carried out in only a handful of cases. We have not yet had sufficient experience with any given condition to permit statistical statements as to its methodologic reliability, except in Rh disease, which is not, strictly speaking, hereditary. Future prospects are exciting, but between present and future lies a great deal of hard work.

Perhaps a key factor in the arrival of prenatal diagnosis was the development of amniocentesis and its use in the management of Rh disease (HOSPITAL PRACTICE, January 1967). With amniocentesis, Liley, Freda, and others established that colorimetric measurements of bilirubin in amniotic fluid could be made with sufficient sensitivity to reveal the degree of erythroblastosis and thus the effect on the fetus. In this way, the physician is afforded a guide to intervention, either by delivery of the infant (if the pregnancy is sufficiently advanced) or by in utero transfusion.

Amniocentesis in Rh disease proved that fetal diagnosis was both possible and useful. Even more important, perhaps, was the demonstration that the uterine contents could be invaded without serious risk to the continuing pregnancy. The procedure has now been performed thousands of times, with next to no mortality and minimal morbidity.

The second factor stimulating the further advance of fetal diagnosis has been the almost explosive increase in understanding of hereditary disease. (Rh disease itself, although, as noted above, not precisely hereditary, obviously has a strong genetic component.) These perceptions have, in part, stemmed from a rapid expansion in our knowledge of human chromosomes and, in part, from an increase in information concerning inborn errors of metabolism. We have witnessed a notable lengthening in the list of hereditary diseases, a much clearer comprehension of the mechanics of inheritance, and the development of techniques for identifying and differentiating various conditions by examination of very small samples of tissue.

However, since most hereditary diseases are still not treatable even when diagnosed, these developments might not have sufficed to swing the balance had it not been for the change in attitudes toward early interruption of pregnancy. In recent years, the feeling has grown among both physicians and the general public that we must be concerned not simply with ensuring the birth of a baby, but of one who will not be a liability to society, to its parents, and to itself. The "right to be born" is becoming qualified by another right: to have a reasonable chance of a happy and useful life. This shift in attitude is shown by, among other things, the widespread movement for the reform or even the abolition of abortion laws.

Amniocentesis is now a commonly performed procedure and does not require detailed description. Certainly in a discussion focused on amniocentesis as a tool for applying genetic knowledge to clinical medicine, one can avoid what is probably the only significant controversy attendant upon the technique – the advisability and safety of using ultrasound to locate the placenta before performing the amniotic tap, so that multiple pregnancies can be identified and the attendant problems for antenatal diagnosis

Amniotic fluid cells are shown in culture at (reading down) 0, 10, and 28 days; two subsequent subcultures are required for cultivation to be considered successful. Photos are reproduced through courtesy of Dr. Henry L. Nadler, Northwestern University, and J. Pediat., 74: 1, Jan. 1968, C. V. Mosby Co.

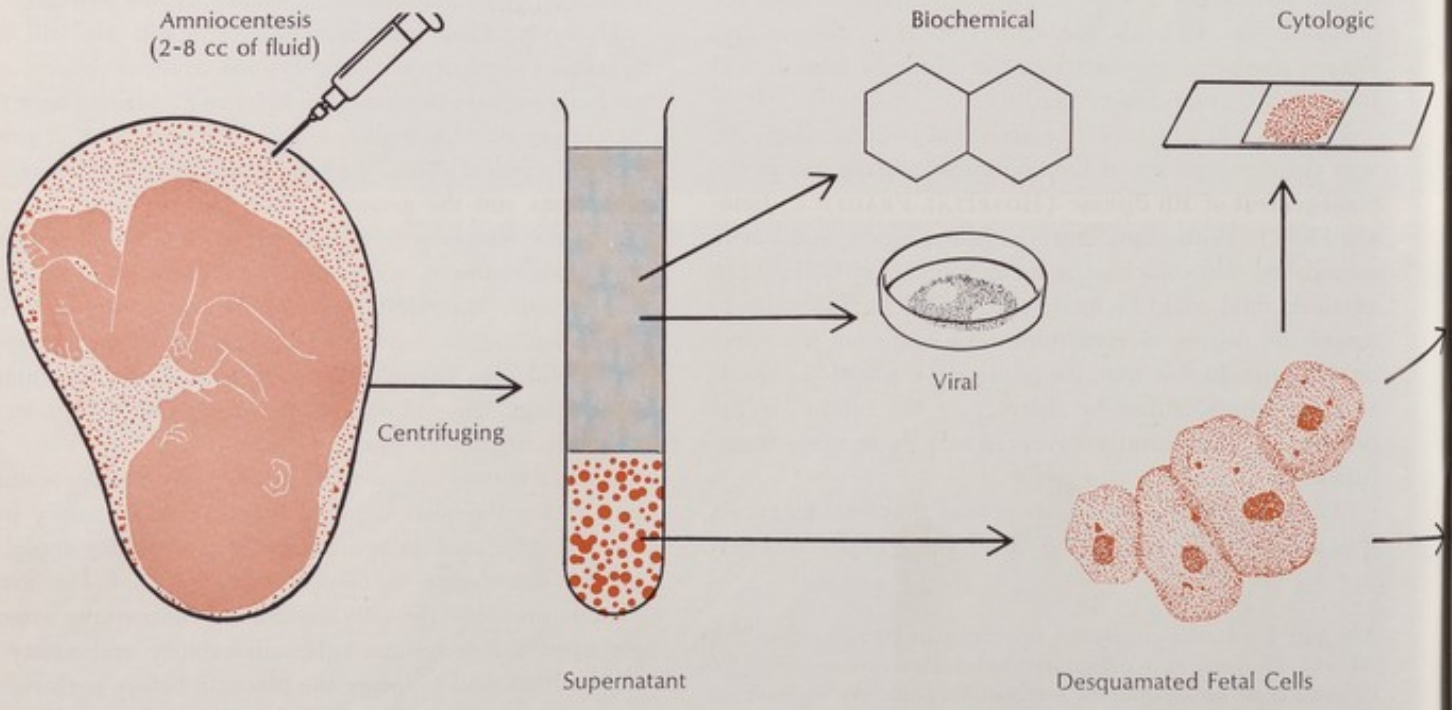
taken into account.

In evaluating amniocentesis, it is necessary to recognize that it yields two "products" – the amniotic fluid itself, and the cells suspended in the fluid. The former, apart from its diagnostic value in Rh disease, has yielded few data of significance. The hope certainly has not been realized that the contribution of the fetal kidneys to the amniotic fluid would make it possible to detect in utero those metabolic anomalies manifest in the urine after birth. The only condition for which this expectation has been borne out to date has been methylmalonic aciduria. In one case of this disorder, an excess of methylmalonic acid was measured in amniotic fluid. A more important value of amniotic fluid measurement has been in the prognostication of the respiratory distress syndrome. By determination of the lecithin:sphingomyelin ratio in the amniotic fluid, the maturation of the pulmonary surfactant system can be evaluated, providing a basis for prediction of RDS susceptibility. In this situation, the source of these lipids must be fluid originating in the pulmonary alveoli, which, of course, suggests a clue as to why the equation of amniotic fluid to postnatal urine has not proved valid. The sources and the metabolic pathways involved in the formation of amniotic fluid are multiple and complex; therefore the concentrations of metabolites in this fluid are very different than in "plain" urine.

At first it was feared that some of the cells found in the amniotic fluid might be maternal. It now appears, however, that they derive entirely from the fetus and the amniotic membrane (which itself derives from the fertilized egg and is therefore genetically identical with the fetus). This has been established by sexing the cells; in no instance has a male infant been incorrectly identified as a female, as would have been expected if there had been significant contamination by maternal cells.

The chief use of cells obtained by centrifuging the fluid has been in sex determination by noting the presence or absence of a "resting" female chromosome, or Barr body (see Ferguson-Smith, Chapter Two). Its presence implies a female fetus, its absence (except in the case of an XO chromosomal anomaly), a male. Such determinations have been used in managing pregnancies where the mother is known to be a carrier of hemophilia or muscular dystrophy, both sex-linked hereditary conditions. In several cases, pregnancies have been terminated when the fetuses were identified as male. To be sure, the son of a hemophilia or muscular dystrophy carrier has an even chance of not inheriting the disease, but anyone who has had contact with such an individual or his parents will comprehend the mother's feelings in such cases.

The most serious risk in making a final decision based on the study of freshly drawn cells lies in the possibility of contamination with maternal blood cells as the result of a traumatic tap. Such an error has been made in glycogenosis II (Pompe's disease) where the high α-glycosidase activity in contaminating maternal blood cells masked the defect in the amniotic cells. However, this kind of error can be circumvented by tissue culture techniques. In this instance, the method is to place the cells in tissue culture and repeat the assay on fibroblasts. This works because under normal tissue culture conditions the maternal blood cells will rapidly die off and, after a week or two, their contamination will be eliminated and subsequent subculture will accurately reflect the condition in the fetal cell lines. Cultures of the cells found in fluid have yielded much more useful information about hereditary diseases than the primary cells themselves. Culturing, by fostering cell division, makes possible karyotyping; in addition, it is known that "latent" charac-



Variety of studies in prenatal detection of inherited disorders can be performed with cells from amniotic fluid. Sexing of primary cells is helpful when mother is known carrier of X-linked conditions. Biochemical studies of cultured cells may reveal meta-

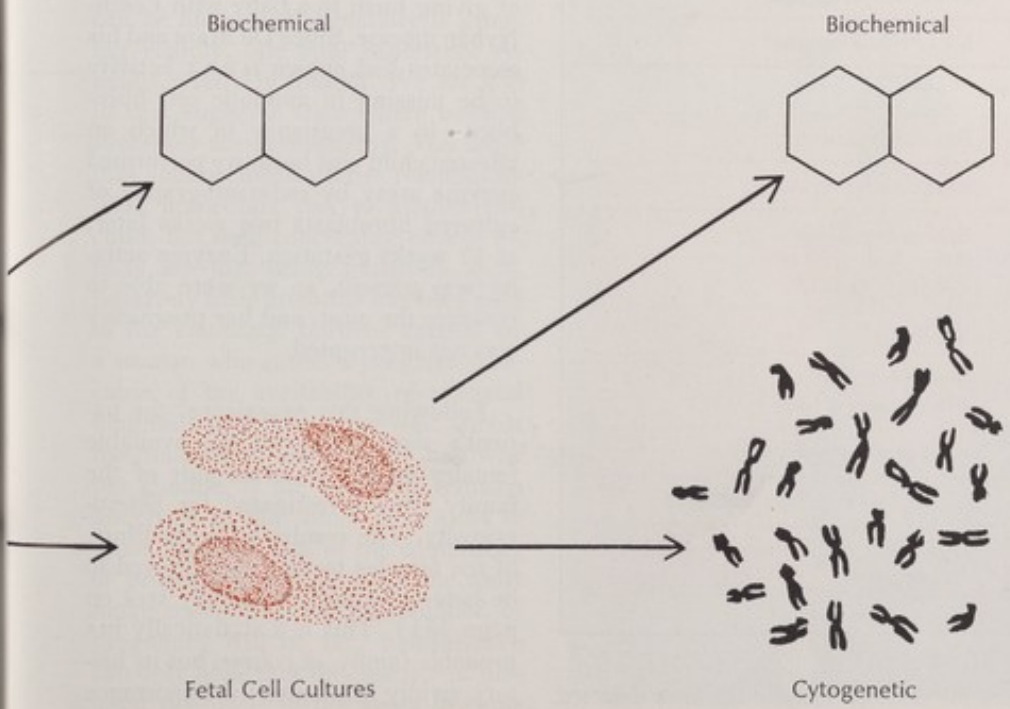
teristics of the cell may emerge during cell culture.

To date, probably the most frequent application of amniocentesis has been in the prenatal detection of chromosomal anomalies, particularly Down's syndrome. However, prenatal diagnosis of these diseases has been hampered by the difficulty, sometimes impossibility, of identifying mothers at risk of transmitting them. Almost all chromosomal anomalies are not inherited but arise de novo in a particular pregnancy. The only indications for fetal chromosome studies are the mother's previous reproductive history, and, in the case of Down's syndrome, her age, since the incidence of babies born with Down's syndrome rises steeply in mothers as they approach 40. (In this context, I should mention that I feel our experience with amniocentesis is still too limited to perform it routinely in pregnant women near 40, as some geneticists have recommended.)

Potentially even more valuable is the application of prenatal diagnostic methods to genetic disorders, those diseases that are transmitted according to Mendelian laws, and in which the risks of giving birth to an abnormal infant can be predicted accurately so that candidates for amniocentesis can be identified more easily. Prenatal diagnosis is theoretically possible if, as is true in many genetic disorders, the biochemical defect has been identified and the enzyme involved is normally demonstrable in fibroblasts grown from amniotic cells. One of the earliest demonstrations of prenatal diagnosis of an enzyme defect involved a woman who had already borne an infant with galactosemia; her subsequent pregnancy was terminated by cesarean section at 33 weeks because of toxemia. Just prior to the cesarean, a sample of amniotic fluid was obtained and the cells cultured; the cultures completely lacked the enzyme whose absence is known to cause galactosemia. This finding (some six weeks after birth because of the time required for culturing) was of no "practical" value, since the enzyme deficiency had already been detected in a sample of cord blood and appropriate dietary measures taken. Nonetheless, it demonstrated that this enzyme deficiency shows up in amniotic fluid cells and can presumably be detected long before birth. The number of genetic diseases that are theoretically detectable in amniotic fluid has grown rapidly in the past few years, and further research can be expected to continue this trend.

The brief history above points up one of the problems basic to prenatal detection of enzyme defects – how to identify heterozygous parents before they have an affected child? In one disease, gangliosidosis G_{M_2} or Tay-Sachs disease (see Kaback and O'Brien, Chapter 25), the heterozygote can be identified with considerable accuracy by a relatively simple assay on plasma. Also, the mutant gene has a relatively high frequency among Ashkenazi Jews although it is rare in the general population. This combination of circumstances has triggered an ambitious effort to identify carriers and, through prenatal diagnosis, avoid the tragedy of even one afflicted child being born. A search for heterozygotes should also prove profitable in the sex-linked recessive diseases, where the fetus is at risk if the mother carries the mutant gene, whether or not her conjugal partner has the same mutation. The case history presented below, involving a boy with the X-linked, recessive Lesch-Nyhan disease, illustrates the importance of identifying heterozygotes with these diseases and, also, many of the significant features of their prenatal diagnosis (see Seegmiller, Chapter 10).

First, a few comments about Lesch-Nyhan disease. It is a relatively recently described genetic disease in which a characteristic finding is the excretion of large quantities of uric acids – far greater than the amounts found in severe gout. A deficiency of hypoxanthine-guanine phosphoribosyl transferase (HPRT) is now known to be responsible for the disease. Clinically, the children are retarded, spastic, and slow-growing; in addition, they show the very bizarre and unpleasant symptom of self-mutilation. For reasons we do not understand, the child is under some strange compulsion to bite his lips and fingers (on occasion, a finger has actually been bitten off), so that his arms must be restrained to prevent injury. The condition is incurable. Allopurinol, which blocks the synthesis of uric acid, can reduce the serum levels of that substance, thereby minimizing kidney damage and extending the life of the child. But it does not affect the neurologic symptoms, which appear to be due to the anomalous production or accumulation of some compound other than uric acid that results from the same enzyme defect, or from some



chromasia (as in mucopolysaccharidoses) or enzyme deficiency (e.g., galactosemia). Cytogenetic analysis may detect Down's syndrome of D/G translocation type.

other subtle disturbance of metabolic equilibrium.

In cases in which familial history indicates that a risk of a child being born with Lesch-Nyhan disease exists, the needs are to determine whether the pregnant woman is a carrier (heterozygote) and whether the fetal cells show the enzymatic deficiency characteristic of the disease.

Luckily, heterozygosity for Lesch-Nyhan syndrome can now be detected. From studies on mice, Mary Lyon predicted several years ago that, because of random inactivation of the X chromosome, carriers of X-linked recessive diseases would be mosaics. In other words, the heterozygote would have both normal and abnormal cells, the performance of any particular cell being dependent on whether or not the X chromosome carrying the mutant gene had been inactivated. This mosaicism can be demonstrated in heterozygotes for Lesch-Nyhan disease by radioautography of skin fibroblasts after incubation with radioactive hypoxanthine. Gartler and his colleagues have described an ingenious approach to the same problem that avoids the delay of tissue culture. This is measurement of the enzyme activity in an individual hair root, which derives from only a few primordial cells. In the heterozygote, enzyme activity will range from zero to normal, depending, of course, upon the makeup of primordial cells in the particular area of the scalp.

Now for the case history:

The boy was referred to our clinic for diagnosis. Shortly after it had been established that he had Lesch-Nyhan disease, it came to our attention that his maternal aunt had become pregnant. We had to discover whether or not she was a carrier for the syndrome.

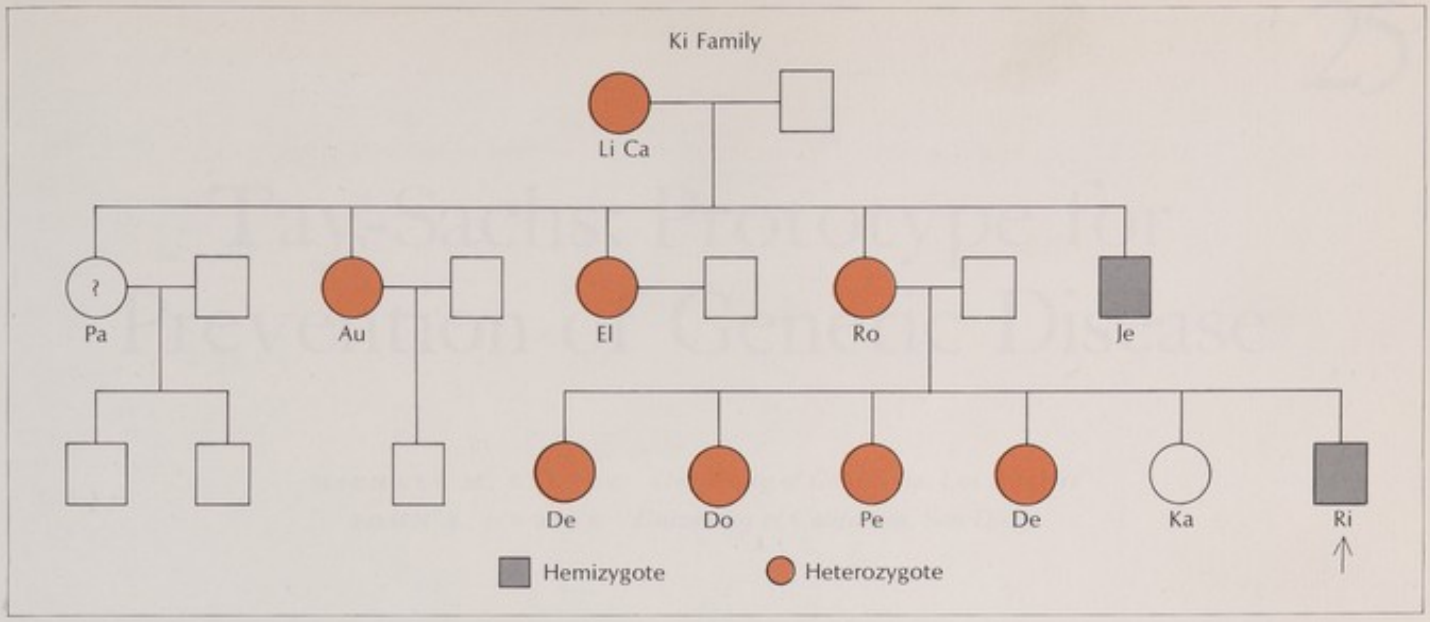
The boy's aunt proved to be a heterozygote – amniocentesis was performed in the 15th week of pregnancy (see Au in family tree on page 251). The absence of Barr bodies on the freshly drawn amniotic cells strongly suggested that the fetus was male; karyotyping fibroblasts grown in tissue culture confirmed this. Thus, the aunt faced a 50% chance at this point of giving birth to a baby with Lesch-Nyhan disease. Since De Mars and his associates had shown HPRT activity to be missing in amniotic cell fibroblasts in a pregnancy in which an affected child was born, we performed enzyme assay by radioautography of cultured fibroblasts two weeks later, at 17 weeks gestation. Enzyme activity was present, so we were able to reassure the aunt, and her pregnancy was not interrupted.

Following this experience, the patient's sisters and all the available females on the maternal side of the family were investigated for heterozygosity. The results were startling: of ten females tested, eight proved to be heterozygotes (see family tree on page 251). This is a statistically improbable family, of course, but its history vividly illustrates the importance of carefully examining potential carriers of an X-linked recessive disease.

Familial Metabolic Disorders Demonstrable in Tissue Culture

Diagnosis	*Deficient enzyme activity*
I. ENZYME HAS BEEN DEMONSTRATED IN CULTURED AMNIOTIC CELL	
Gaucher's disease	Glucocerebrosidase
GM, gangliosidosis	β-galactosidase
Arginosuccinic aciduria	Arginosuccinase
Homocystinuria	Cystathionine synthetase
Maple syrup urine disease	Branched-chain keto acid decarboxylase
Ornithine-α-keto acid transaminase deficiency	Ornithine-α-keto acid transaminase
Fucosidosis	α-fucosidase
Glycogen storage disease Type III	Amylo-1,6-glucosidase
Type IV	Branching enzyme
Mamosidosis	α-mamosidase
G-6-PD deficiency	G-6-P dehydrogenase
Xeroderma pigmentosum	DNA "repair enzyme"
II. DIAGNOSIS HAS BEEN MADE PRENATALLY	
GM_2 gangliqsidosis (Tay-Sachs)	Hexosaminidase A
Metachromatic leukodystrophy	Arylsulfatase A
Niemann-Pick disease	Sphingomyelinase
Hurler's syndrome	Iduronidase
Hunter's syndrome	Unknown
Methylmalonic aciduria	Methylmalonyl CoA Vitamin B_{12} defects
Glycogen storage, Type II (Pompe's)	α-1,4 glucosidase
Galactosemia	Galactose-1-P uridyl transferase
Lesch-Nyhan disease	Hypoxanthine-guanine phosphoribosyl transferase

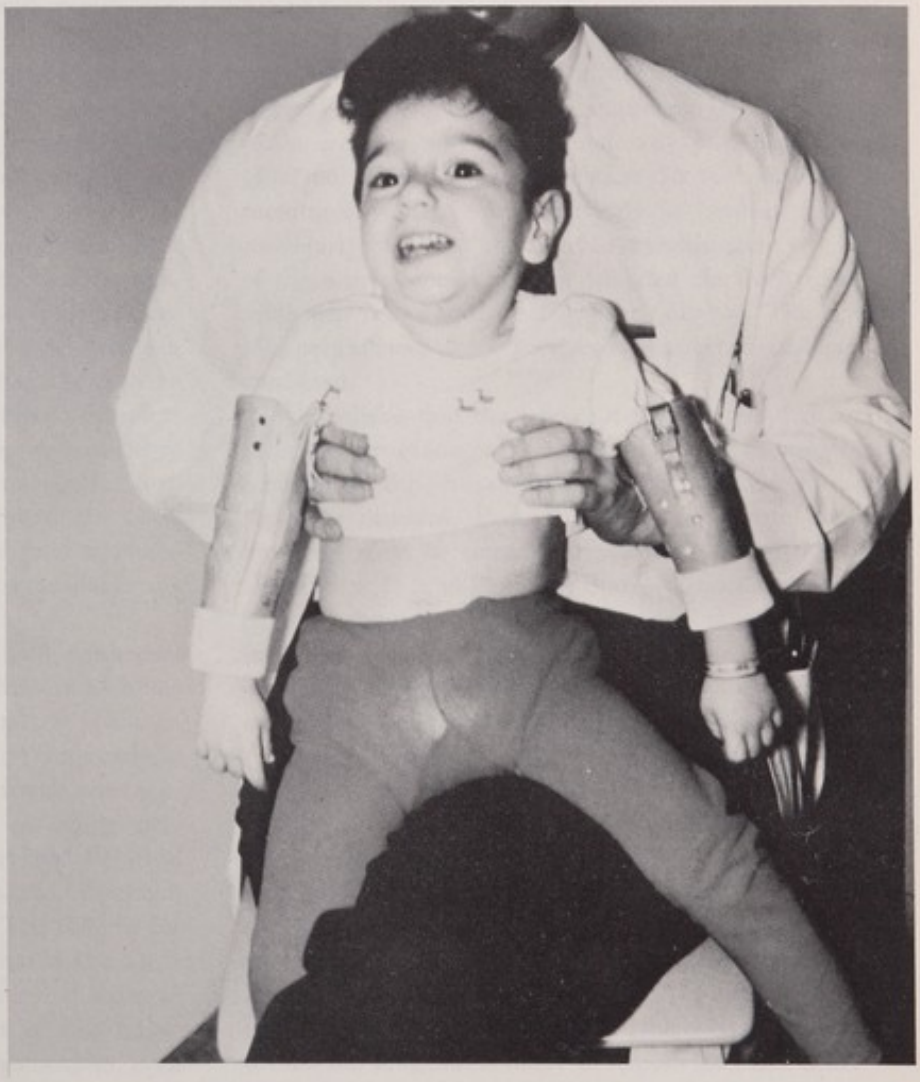
Light color indicates those disorders in which the involved enzyme has been detected in amniotic fluid cells cultured from normal fetuses so that prenatal diagnosis should be possible; darker color, those disorders which have actually been diagnosed prenatally.



Ri presented at 4 years of age with symptoms of cerebral palsy and self-mutilation. The diagnosis of Lesch-Nyhan disease was made by demonstrating the appropriate enzyme defect in the red blood cells. Uncle Je had died undiagnosed many years previously with the same clinical picture. Sisters and aunts were investigated for heterozygosity.

A son born to any of these carriers has a 50% chance of being afflicted with a seriously incapacitating disease; amniocentesis and prenatal diagnosis, coupled with intervention when necessary, can obviate this tragedy.

The case histories presented above show the impact prenatal diagnosis can have in medical genetics, but it is important to keep these exciting developments in proper perspective. Although the number of diseases that can be diagnosed prenatally is growing rapidly, our total experience remains relatively limited. We are still in that euphoric state where isolated successes are newsworthy. We know that amniocentesis carries remarkably little immediate risk to mother and child, but data concerning remote effects are just being gathered. Also, relatively little attention has been paid to the emotional burdens carried by a woman who enters a pregnancy because of the availability of prenatal diagnosis and then must wait almost five months to learn if she has won her gamble. If the diagnosis remains uncertain at that time, as it must in some cases because of the complexity of procedures, the decision she must make concerning the continuation of pregnancy will be very painful. We can only hope that the practice of prenatal diagnosis in the years to come will meet the promise offered by theory.



Child with Lesch-Nyhan syndrome (hyperuricosuria) requires restraints on the arms to prevent self-mutilation; mental retardation is also associated with the disorder.

25

Tay-Sachs: Prototype for Prevention of Genetic Disease

MICHAEL M. KABACK *University of California, Los Angeles*
JOHN S. O'BRIEN *University of California, San Diego*

One day in May 1971 some 1,800 men and women presented themselves at a Washington, D.C., synagogue as part of a pioneering experiment in the delivery of new medical knowledge. During a seven-hour period, three quarters of these people had blood drawn for a simple serum test to detect heterozygous carriers of the autosomal recessive disorder called Tay-Sachs disease.

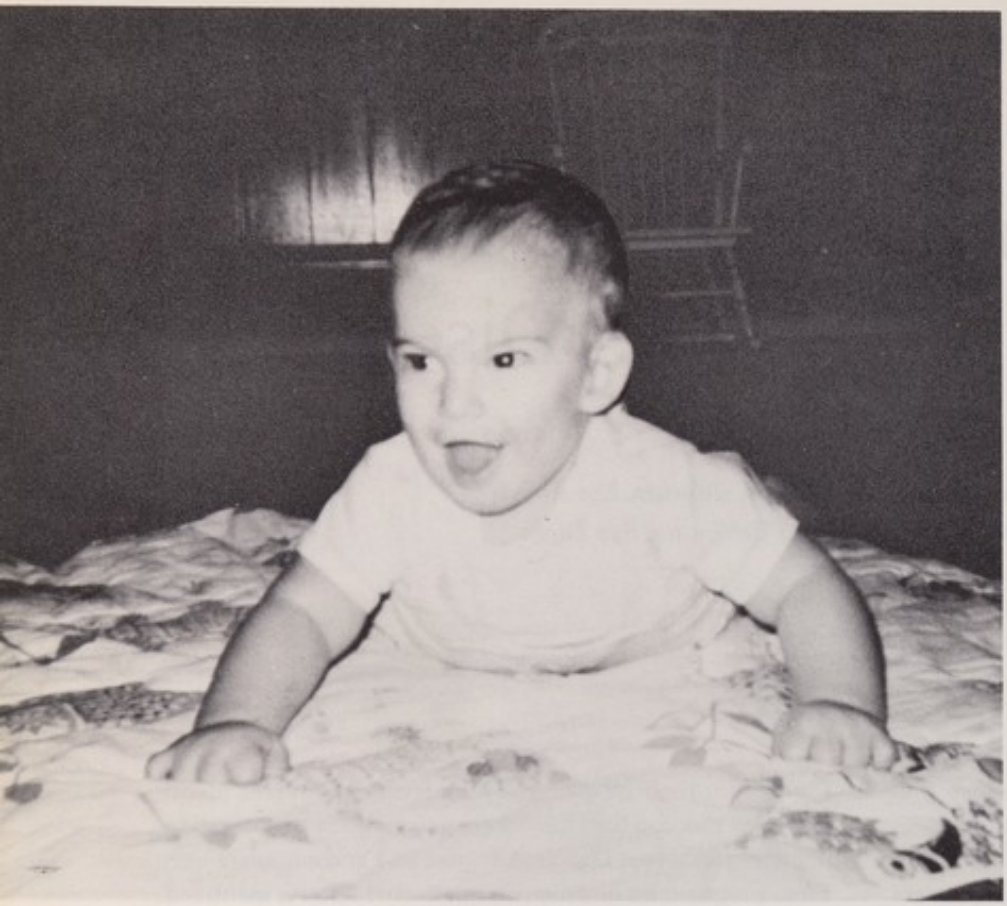
It was the first voluntary community-based attempt at mass screening of adults for carriers of a genetic abnormality. It was remarkable for another reason as well. The enzyme deficiency responsible for Tay-Sachs disease had been identified only two years earlier. Since then the enzyme assay had been automated, making it possible to complete 300 serum assays per day. In addition, the organizational and educational action required for a mass screening program had been carried out. In only two years the basic scientific understanding of this uniformly fatal neurodegenerative childhood disorder had become available for large-scale application to the population.

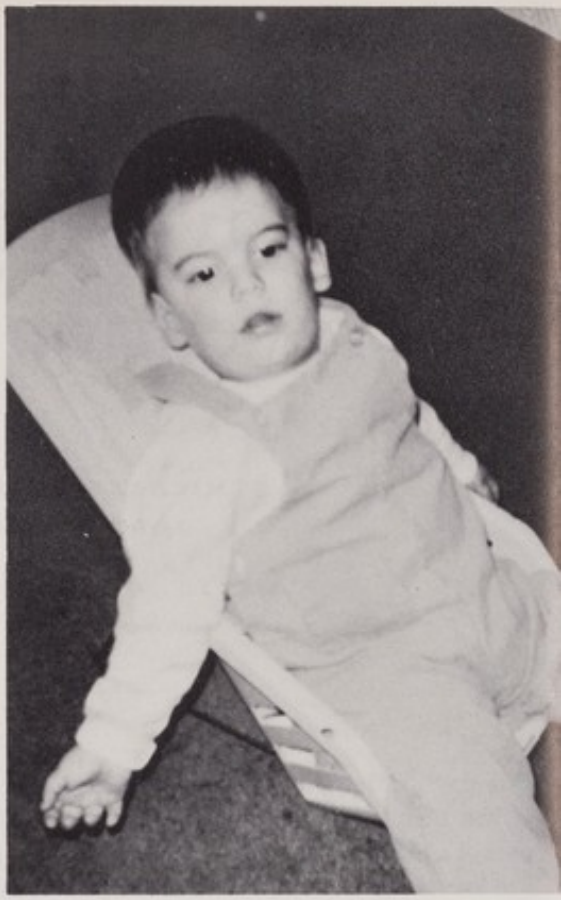
Continuation of the screening program in the Ashkenazi Jewish population in the Washington-Baltimore area has led to the testing of nearly 10,000 individuals. It is known that Tay-Sachs disease occurs 100 time more frequently in Ashkenazi Jews than in other Jewish groups and non-Jewish populations. Similar community screening programs have begun or are being planned in at least 40 cities, not only in the United States but in Canada, England, Israel, South Africa, and other countries. The reasons for widespread interest in such programs is *not* that TSD is that common, even amongst Jews, but rather that a simple blood test allows for complete prevention of this tragic genetic disease and at the same time enables couples, even if genetically at risk, to have unaffected children. Blood testing can determine whether or not a couple are at risk for TSD in their offspring. The infrequent couples identified as at risk can then have their pregnancies monitored by amniocentesis (see Dancis, Chapter 24) and complete only those pregnancies in which an unaffected fetus is identified. Since accurate antenatal detection of TSD became available, more than 100 pregnancies at risk for TSD have been monitored to date (predominantly in families that have previously had TSD children). Approximately one fourth of these pregnancies were terminated electively after the fetus was found to have Tay-Sachs disease. The remaining pregnancies have produced unaffected children, as predicted.

We believe that the basic procedures adopted for the screening programs in Washington and Baltimore provide a model for the prospective prevention of other autosomal recessive diseases. Prevention by therapeutic abortion, however, must be considered a temporary and imperfect alternative. It is to be hoped that continued research will result in the discovery of an effective treatment or a cure for Tay-Sachs disease. Until that happens, carrier identification, prenatal diagnosis, and abortion of affected fetuses can prevent the tragedy of the birth of a Tay-Sachs child.

In short, just as phenylketonuria represents a prototype for screening newborns for treatable genetic disease, (see Hsia and Holtzman, Chapter 23), Tay-Sachs provides a model for mass screening of adults so that couples at risk for recessive disorders can be located prior to the conception of affected offspring, thereby making prospective prevention possible. Since the prospects for effective treatment or cure are now remote, the need for the interim prevention program is likely to remain for some time.

Genetic or genetically related conditions account for a significant proportion, perhaps 10% to 20% of all pediatric





The progression of Tay-Sachs disease is illustrated in this sequence of photos showing the same child. At left above, at age 3½ months, he is asymptomatic and appears to be developing normally. Although a month later he was able to sit with little or no help, by age 7 months (at right above), early regression was apparent and he could no longer sit alone. He also displayed

hospitalizations, with hereditary neurologic disorders being especially prominent. Of the 420 or so autosomal recessive conditions already known, many occur more frequently in specific populations where screening might be highly effective – for example, sickle cell anemia in blacks, thalassemia in Italians, and cystic fibrosis in Anglo-Saxons.

Mass screening is not something that can be embarked upon casually however (see Guthrie, Chapter 22). The mere suggestion that an individual or a group may have "bad genes" can arouse significant anxieties. Clearly the screening test must be accurate, and experienced genetic counselors are required to guide individuals who are identified as carriers. A screening program should be able to offer those screened a positive course of action if they are shown to be at risk. It is our belief that if the program cannot offer an alternative of treatment, cure, or at least prevention without limitation of individual freedom of choice in mate selection or procreation, it should not be undertaken on a large scale. These matters will be discussed in more detail later.

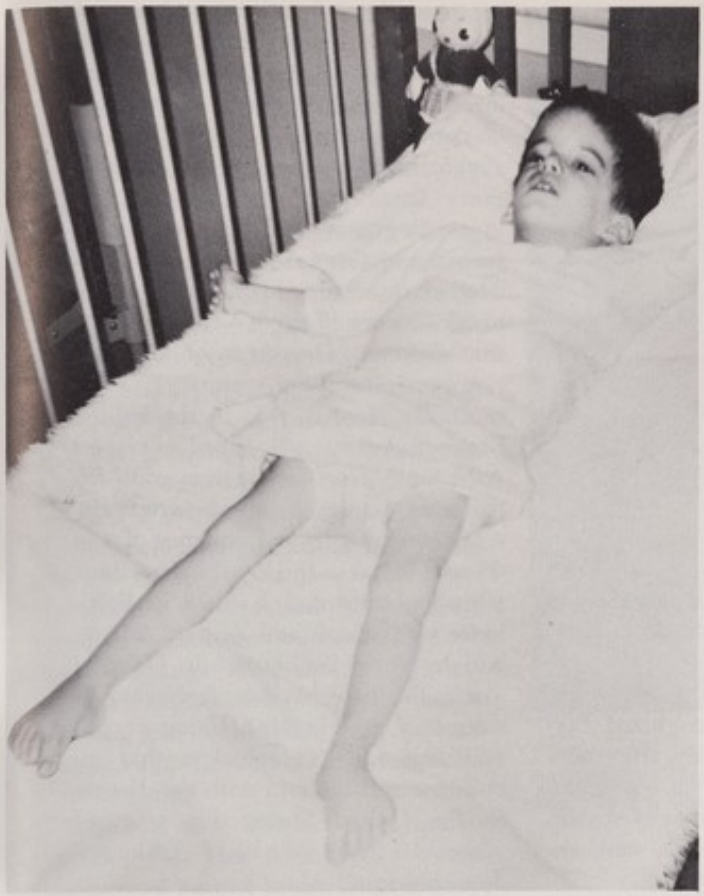
The Tay-Sachs Child

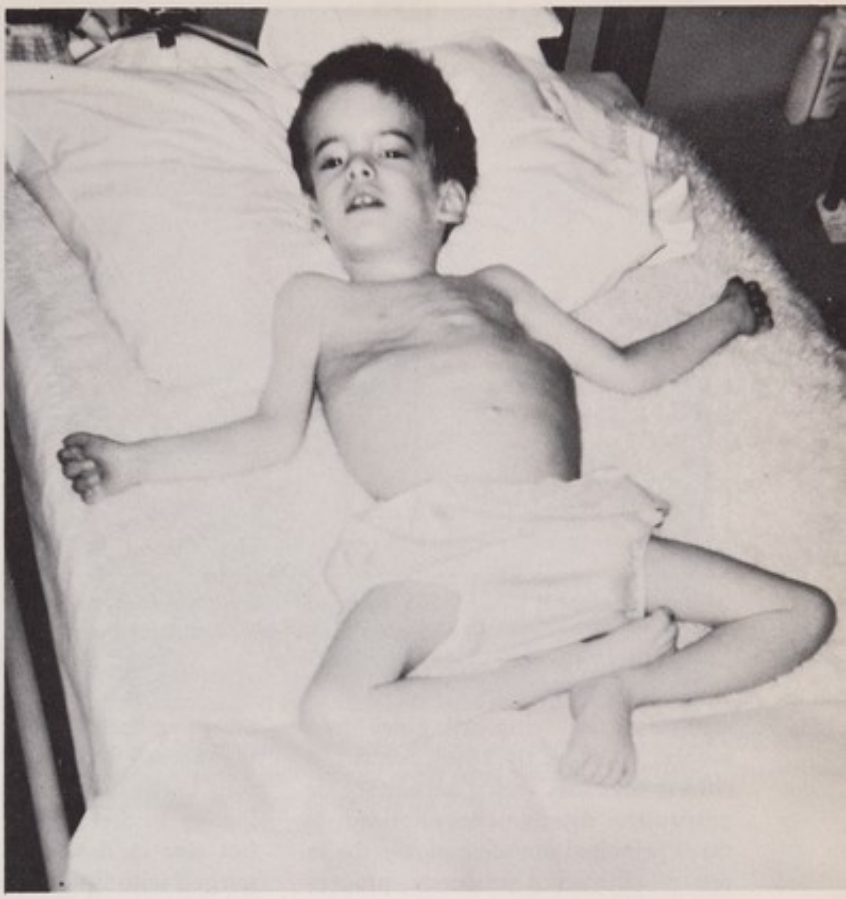
Tay-Sachs disease is the most common form of sphingolipidosis. Several thousand infants have died with this disease since it was recognized and characterized late in the 19th century. The typical course of a patient with the disease is illustrated in the photographs at the top of these pages. Most Tay-Sachs babies are beautiful. They have clear, translucent skin with pink, doll-like coloring and long eyelashes. They appear healthy at birth and develop normally for the first three to six months. Gradually, however, the central nervous system degenerates because of the progressive intraneuronal accumulation of excess amounts of the sphingolipid ganglioside G_{M2}. By the time the child is 8 to 12 months old, physical and mental deterioration is usually obvious.

Some of the first signs of deterioration noted by parents are mild motor weakness and an increased startle response to sudden sound. The child usually develops the ability to crawl and sit up unaided; he may even be able to pull to standing. But within a few months he can no longer sit unassisted and must be propped. The mother may also notice that he does not fix his gaze well; his eyes wander because of failing vision. Soon the infant ceases to smile or react to social stimuli. Ultimately, usually by 18 months, he becomes paralyzed and blind.

As the disease progresses, medical management becomes more difficult. The child must be tube fed because he has difficulty swallowing; secretions may pool in the bronchi and must be aspirated; he must be turned in bed regularly to avoid pressure sores; he may be constipated and require manual evacuation, enemas, or cathartics; eventually he develops seizures that may occur as often as once an hour. These respond initially to anticonvulsants but later become refractory.

Diagnosis is generally made within





decreased muscle strength and coordination, was less interested in his environment, his eyes tended to wander, and he exhibited a striking startle response to minimal sounds. Hospitalization became necessary at 18 months. At left above, at 38 months, seizures, blindness, and severe retardation were present. Final photo was made at 42 months; the child died two months later.

9 to 12 months after birth. The children usually die between the ages of three and five, frequently because of bronchopneumonia; the average age at death is 40 months.

Hospitalization is usually begun after a child has reached the age of about 16 months. Some parents wish to keep their child at home and can care for the problems indicated above with support and help. In either hospital or home, long-term intensive care for two to three years is required. In some facilities, 24-hour nursing care may cost as much as $180 a day. Hospital costs – when suitable facilities can be found – vary between $10,000 and $50,000 a year. Until prenatal diagnosis became a possibility, few couples dared to risk the birth of another child once they had gone through the experience of having a Tay-Sachs infant.

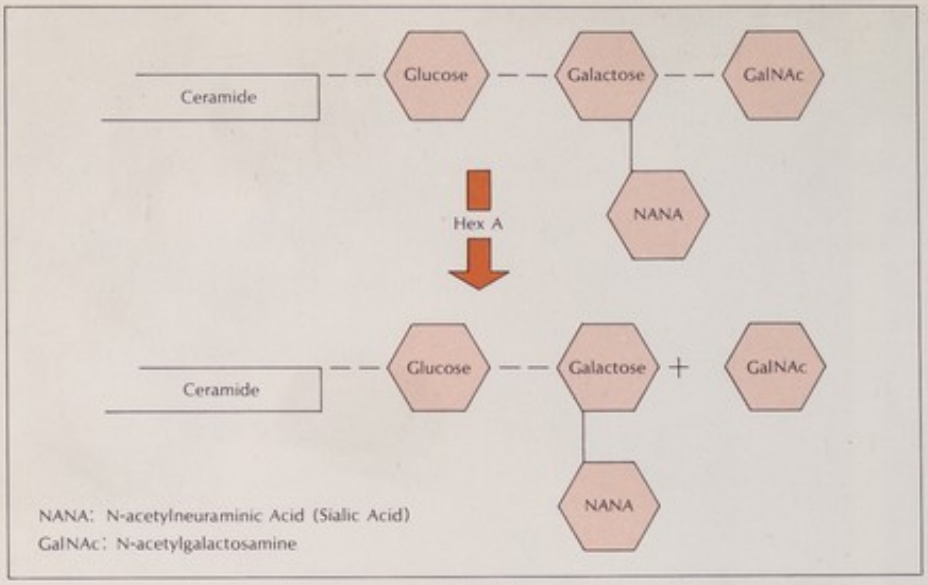
Early History and Genetics

Demographic studies by Stanley Aronson, Professor of Pathology at Downstate Medical Center, Brooklyn, and Ntinos Myrianthopoulos, at the National Institute of Neurological Diseases, demonstrated that the ancestors of the majority of Jewish cases of Tay-Sachs disease in the United States can be traced to the northeastern provinces of Poland, Western Russia, and those neighboring upon the Baltic Sea (Kovno, Suwalki, and Grodno). Very few originated in the Western Balkan zones or Germany. The reason for maintenance of the high gene frequency for TSD amongst Ashkenazi Jews is obscure. Such explanations as consanguinity, selective advantage for the heterozygote, founder effect, and genetic drift have all been invoked as contributing factors.

The modern history of the disorder began in 1881 when Warren Tay, a British ophthalmologist, described the first recorded case. He reported eye-ground changes in a one-year-old child with pronounced muscular weakness. He also described one of the characteristic features of the disease, the so-called cherry-red spot (see the illustration on page 257), as follows: "In the region of the yellow spot in each eye a conspicuous, tolerably defined large white patch, and showing in its center a brownish-red, fairly circular spot, contrasting strongly with the white patch surrounding it."

The cherry-red spot is not diagnostic for Tay-Sachs disease. While almost all children with the disorder have this abnormality, there are at least half a dozen different storage disorders in which cherry-red spots occur. The spot can be seen as early as the first few days of life. The red spot is not in itself abnormal, since it represents the normal vasculature of the foveal retina; it is the accumulation of lipid in ganglion cells adjacent to the fovea, resulting in the white halo, that is pathologic.

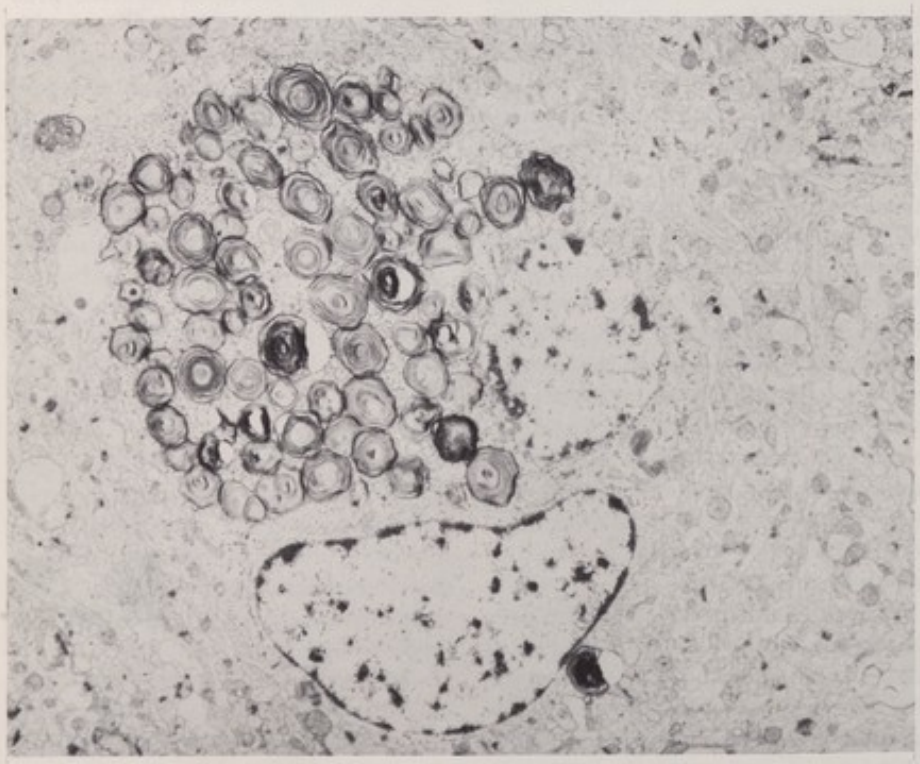
Tay subsequently saw two more children with similar symptoms in the same family, and another in a second family. Then, in 1887, the American neurologist Bernard Sachs gave the first pathologic description of the dis-



Hexosaminidase A (Hex A) is a lysosomal hydrolase, which, as illustrated above, acts to catalyze the cleavage of N*-acetylgalactosamine (Gal NAc) from ganglioside* G_{M_2}.

ease after studying 19 cases from several families. By 1898, Sachs concluded that this was a hereditary degenerative disease characterized by three principal manifestations: the arrest of all mental processes, progressive weakening of the muscles terminating in general paralysis, and rapidly developing blindness associated with changes in the macula lutea, development of the cherry-red spot, and optic atrophy. Sachs also recognized that the disease was a lipidosis, since neurons in the nervous system – predominantly in the central nervous system but also in the periphery – were engorged with lipid material.

Tay-Sachs disease is transmitted as an autosomal recessive trait. Parents of affected children are clinically normal and both have hexosaminidase A function that is intermediate between those of patients and normal subjects. Sex ratios are nearly equal. Analysis of pedigrees, correcting for incomplete ascertainment of matings of two heterozygotes, gives ratios of approximately one affected child to every three unaffected children. In utero diagnosis of fetuses conceived from the matings of two heterozygotes demonstrated that in 39 instances 9, or 23%, were affected, a value close to that expected. Drs. Aronson and Myrianthopoulos have estimated, from mortality records, that in the United States as many as 1 out of every 30 Ashkenazi Jews is heterozygous for the trait. Using enzyme assays to detect heterozygotes, a frequency of 1 in 27 was found in the Washington Jewish population and 1 in 23 in Baltimore by Kaback and Zeiger. Among non-Jews or Sephardic or Oriental Jews, the heterozygote frequency is estimated at 1 in 300. Using overall statistics one can calculate that 50 children will be born with the disease in the United States this year, of whom 40 to 45 will be of Ashkenazi Jewish origin.



Electron micrograph of a neuron (x16,000) shows accumulations of lipid (ganglioside G_{M_2}*) in form of many concentrically arranged, membranous cytoplasmic bodies.*

The Enzyme Defect

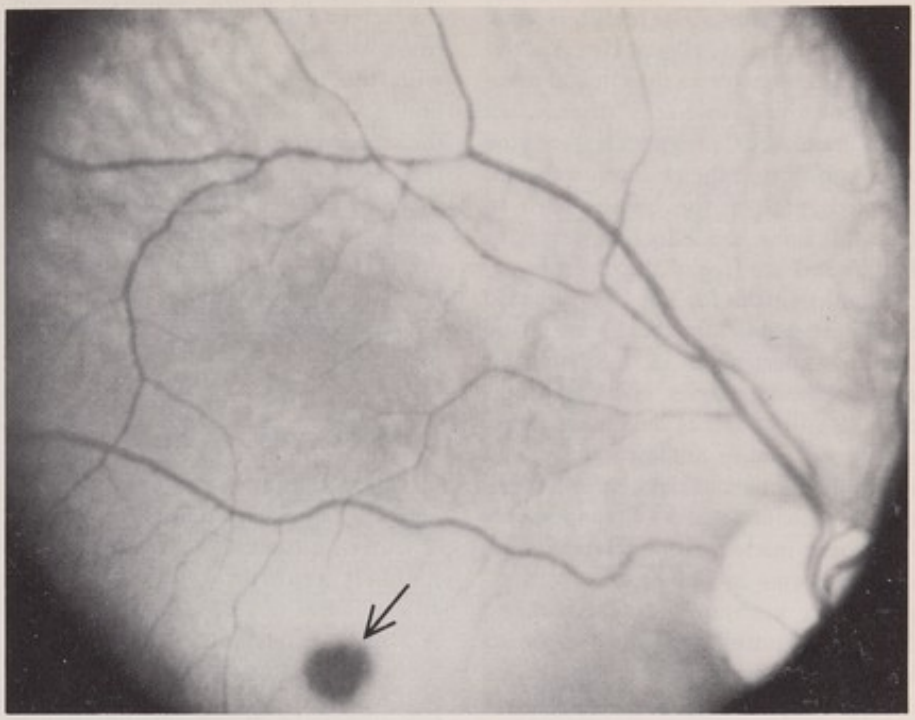
The biochemical history of Tay-Sachs disease began when the late Ernst Klenk, Professor of Chemistry at Cologne University, described a massive accumulation of gangliosides in brain tissue of affected children in 1942. Twenty years later Lars Svennerholm, Professor of Biochemistry at the University of Gothenburg, identified the specific lipid as ganglioside G_{M_2} and characterized its structure. Ganglioside G_{M_2} is one of seven or eight different gangliosides normally present in man's brain, but in Tay-Sachs disease it accumulates to levels 100 to 300 times normal.

The ganglioside accumulation could be explained by excess synthesis, diminished degradation, or both. Several pieces of evidence suggested decreased degradation as the most likely mechanism. First, ganglioside G_{M_2} accumulates in structures that resemble altered lysosomes. Second, the normal brain contains lysosomal hydrolases capable of breaking down ganglioside G_{M_2} by sequentially cleaving off sugar molecules. This made the deficiency of a lysosomal hydrolase a conceivable explanation for Tay-Sachs disease.

This hypothesis was strengthened by the findings of Dr. Roscoe Brady and his coworkers at the National Institutes of Health that in several other lipid storage diseases the absence of a specific degradative enzyme accounts for the accumulation.

When British workers in 1968 reported the existence of two hexosaminidases in normal human spleen tissue, Drs. Shintaro Okada and O'Brien, working at the University of California, San Diego, recognized that one or both of these enzymes could be involved in ganglioside G_{M2} degradation by cleaving the terminal beta-linked *N*-acetylgalactosamine (see illustration on page 256). When frozen tissue from Tay-Sachs children was examined, it was found that one of the enzymes, hexosaminidase A (Hex A), was missing. The absence of Hex A was demonstrated in freshly drawn serum and leukocytes from patients and in skin fibroblasts growing in tissue culture.

Okada and O'Brien also demonstrated that Hex A and B are found in normal human brain, liver, kidney, skin, leukocytes, serum, and cultured skin fibroblasts, and that Hex A is missing from all these tissues in patients with Tay-Sachs disease. They also established that both enzymes are present in cells cultured from amniotic fluid obtained during the second trimester of normal pregnancy and that parents of affected children have lower than normal levels of Hex A.



Funduscopic view of the eye in a child with Tay-Sachs disease shows the cherry-red spot (arrow) characteristic of this disorder and several other lipid storage conditions; typically a yellow ring surrounds the spot, first described by Tay.

Prenatal Diagnosis

These findings laid the foundation for both prenatal diagnosis and carrier detection. Just a few days after the first publication (in 1969) of Hex A data a woman telephoned O'Brien. She had borne a Tay-Sachs child eight years before and was 13 weeks pregnant. She said that if she could be confident that the baby would not have the disease she would continue the pregnancy because she and her husband wanted another child. If the answer was uncertain, she was considering terminating the pregnancy by therapeutic abortion, since she could not face the tragedy of another Tay-Sachs baby.

Fluid was obtained by amniocentesis and Hex A was found in both the amniotic fluid and the cells. O'Brien indicated that the baby very likely did not have Tay-Sachs. Five months later she delivered a baby girl. Serum assay of the umbilical cord blood showed normal Hex A activity. O'Brien examined the child when she was 14 months old and she was developing normally; serum Hex A assay demonstrated that she was heterozygous. Another laboratory in which the prenatal diagnosis for Tay-Sachs was carried out was that of Dr. Lawrence Schneck of Kingsbrook Medical Center in Brooklyn. All told, more than 100 pregnancies have been monitored for Tay-Sachs disease, 39 of them in O'Brien's laboratory alone. These are predominantly in couples who have previously had TSD children. Now, however, some at-risk pregnancies are being monitored in couples who have not had a TSD child but who have been identified in screening programs.

The fluorometric assay for serum developed by Okada and O'Brien was modified to permit assaying of Hex A in amniotic fluid, uncultured cells, and cultured amniotic cells. Amniocentesis for diagnosis of Tay-Sachs disease in utero is best carried out between the 14th and 16th gestational week. By this time there is enough fluid so an adequate sample can be obtained and yet there is still time for the cells to be cultured before the optimal time for an abortion has passed.

In laboratories with well-developed expertise in growing amniotic cells, the results of all three assays should be consistent. In our laboratories, for instance, we assay the fluid itself and cells sedimented from the fluid. We confirm the results of these two assays by assaying the cultured cells when they have grown out – in from 10 days to four weeks. The results on the cultured amniotic cells are the most reliable, since they produce a greater spread of values between affected and nonaffected individuals (heterozygotes and normal homozygotes).

Among the 39 monitored pregnancies at the University of California, San Diego, there were nine fetuses deficient in Hex A. Eight of the pregnancies could be terminated safely, and were. Diagnosis of Tay-Sachs disease was confirmed in seven by means of electron microscopy, ganglioside analysis, and enzyme assays; one fetus was unavailable. On the average there was a 30-fold increase over normal levels in cerebral ganglioside G_{M2} levels in the affected fetuses. In the ninth affected pregnancy amniocentesis was

carried out too late to end the pregnancy artificially. The child is now over two years, has absent Hex A activity, and demonstrates the clinical manifestations of Tay-Sachs disease. In the remaining 30 pregnancies, enzyme assays indicated the children were not affected; and of the babies born to date, all have had adequate levels of Hex A and are free of the disease.

Amniocentesis for Tay-Sachs should not be attempted unless 1) an obstetrician experienced in midtrimester amniocentesis carries out the procedure, 2) the parents have been counseled concerning the procedure and its implications and risks by an experienced clinician, and 3) arrangements have been made for the determination of the enzyme in a laboratory with expertise in hexosaminidase A assays. Numerous pitfalls, which include an improperly done amniocentesis, an inadequate specimen, bacterial growth in the sample, and artifactual inactivation of the enzyme due to shipment problems or poor laboratory technique, can be avoided by careful planning so that the ultimate disaster of a false-positive or false-negative diagnosis will not occur. When these requirements have been met in the past, to our knowledge no errors in diagnosis have been made. Should there be any question about these matters, the mother should be referred to one of the major centers in the United States that are capable of providing these services.

To develop an effective therapy, many laboratories are now trying to characterize the basic molecular defect responsible for the Hex A deficiency, with the ultimate aim of trying to find a way to activate production of the missing enzyme or replace it. Intravenous administration of Hex A in matched plasma from normal donors has been attempted by O'Brien and others; however, dozens of infusions in patients at various stages of illness have produced no clinical improvement.

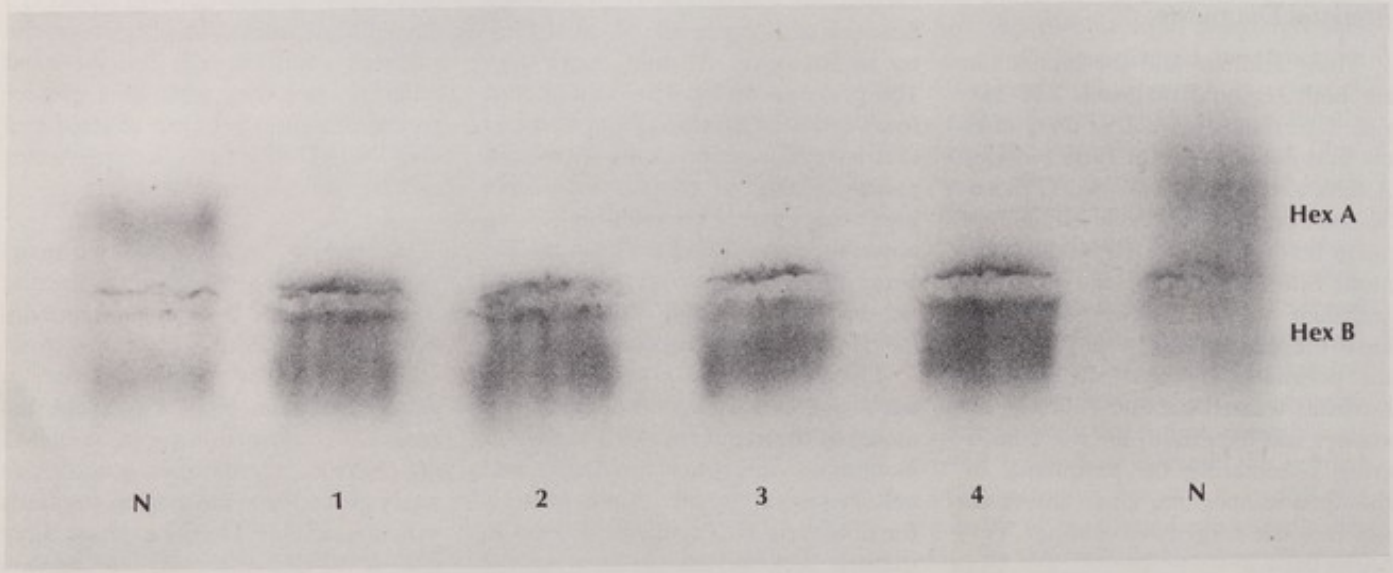
Carrier Detection

Unfortunately for the parents involved, Tay-Sachs disease is typical of other autosomal recessive diseases in that it occurs principally in families where it has not been known before. A study conducted 40 years ago by Slome, involving 88 sibships and 130 cases of Tay-Sachs disease, revealed that 82% of the cases marked the first appearance of the condition on either side of the family. Thus, if one waits for proband cases to occur, only about 20% of the total cases are preventable. In order to detect the majority of the cases prenatally it is necessary to identify at-risk matings before they reproduce.

The fundamental procedure for carrier detection is the fluorometric assay devised by O'Brien et al, which estimates the activity of both Hex A and Hex B in serum. A slightly different version of the same assay is used for leukocyte analysis. The basic assay was modified somewhat by Kaback and his associates when he was at Johns Hopkins to permit the screening of serum from large numbers of individuals.

The assay of Hex A and B depends upon the difference between the heat stability of the two enzymes. Hex A activity is rapidly destroyed by heating, while that of Hex B is not. The serum assay is so sensitive it can be performed on a drop of blood from a heel stick in an infant, although collection of a larger sample is usually preferred so that multiple tests can be carried out. While the serum analysis is simple enough to be conducted in any hospital laboratory, it requires critical internal quality control to provide accurate, reproducible results.

As the assay is used in our laboratories, primary serum screening can positively identify 97% of those tested as either carriers or noncarriers of the Tay-Sachs gene. The remaining 3% fall into an inconclusive range, arbitrarily defined to avoid false-negative and false-positive results. These "inconclusive" subjects are retested with the more accurate leukocyte assay, which gives greater than 99% confidence in genotype designation. In noncarriers, 50% to 75% of their total hexosaminidase activity is heat labile (Hex A); in carriers, this value is 20% to 45%. To be classified as a carrier, a subject must have three consecutive duplicate serum determinations in the carrier range. Two consecutive determinations in the higher



Starch gel chromatographs of liver tissue obtained from four aborted TSD fetuses (numbered) and two normal (N) aborted fetuses show that while Hex B is present in both groups, Hex A activity is missing in the livers of the affected fetuses.

range are sufficient for noncarrier designation.

Hex A activity in the serum may reflect many factors extraneous to the Tay-Sachs gene, such as medications, pregnancy, and systemic illnesses, e.g., diabetes and hepatitis, which may give unreliable data regarding carrier status. For example, the serum assay in pregnant women after the fourth week of pregnancy or in some women taking birth control medications may falsely indicate carrier Hex A levels. However, the leukocyte assay remains accurate in both instances. Although the leukocyte method is more accurate it is also more laborious and costly, making it less desirable as a screening test.

Approximately 350 individuals out of the first 10,000 screened in the Washington-Baltimore program were "inconclusive" after initial serum testing. Leukocyte assays on over 300 of these subjects allowed all but four to be accurately genotyped. These four probably represent genetic variants for Hex A, since other members of their families also had inconclusive serum and leukocyte Hex A levels. In any couple in which both partners are found to be carriers by serum assay, confirmatory leukocyte assays should be carried out and family studies conducted in order to corroborate these findings.

Mass Screening

Three criteria that make prospective prevention of TSD feasible are: 1) it occurs principally in a defined population group so that selective screening is possible; 2) there is a simple, accurate, and inexpensive carrier detection test; and 3) the condition can be detected in an affected fetus early enough in pregnancy to permit selective abortion if the parents so desire, and to enable at-risk couples to have unaffected offspring.

If mass screening can be justified because it provides a positive alternative, it still must be undertaken with human values foremost in mind. Careless planning, premature announcements, incorrect information, or improperly delivered results can have disastrous effects on those who are supposed to be helped by the program.

Those planning a voluntary community-based screening program should first determine which of two basic strategies to employ: to attempt to reach a major share of the population at risk or to make a service available on a limited scale to those who happen to learn of the service and come in spontaneously. The choice will depend not only on a given community's needs but also on the resources available. If voluntary community-wide screening is decided upon, the planners will have to be prepared for a detailed organizational effort, with carefully planned public education.

The planners should also decide just who will be screened. Because facilities and personnel were limited and because of potentially greater psychologic problems and misunderstandings in teenagers, the Washington-Baltimore testings were confined to men and women of childbearing age, and principally to married or engaged individuals. If a woman was more than four and a half months pregnant neither she nor her husband was tested because time would be inadequate to carry out the required procedures and for the information to be useful in that pregnancy. If the woman was less than four and a half months pregnant, only her husband ("pregnant husband") was tested, since her *serum* test would be likely to give a spurious result. If his serum test indicated he was a carrier or if it was inconclusive, both man and wife were immediately asked to come to the hospital for leukocyte and repeat serum assays. Serum screening of "pregnant husbands" was given priority to minimize "deadline" problems.

If the target population is sufficiently large, a community-wide program is a rational way to bring genetic screening and counseling to those who need it. This sort of program requires long and complex preparation, however, and should not be attempted unless the planners are thoroughly committed to meeting the responsibilities it will eventually generate. The strategy and planning sequence of the Washington-Baltimore program is illustrated on page 260.

One of the major foundations of a large screening project should be a strategically planned educational program. This must start many months in advance of the screening itself. In the project at Johns Hopkins, 14 months of planning and education preceded public announcements of the first screening session. The only technical work done during this period of time was development of the automated assay and trials of the technique in parents of affected children and with hospital personnel.

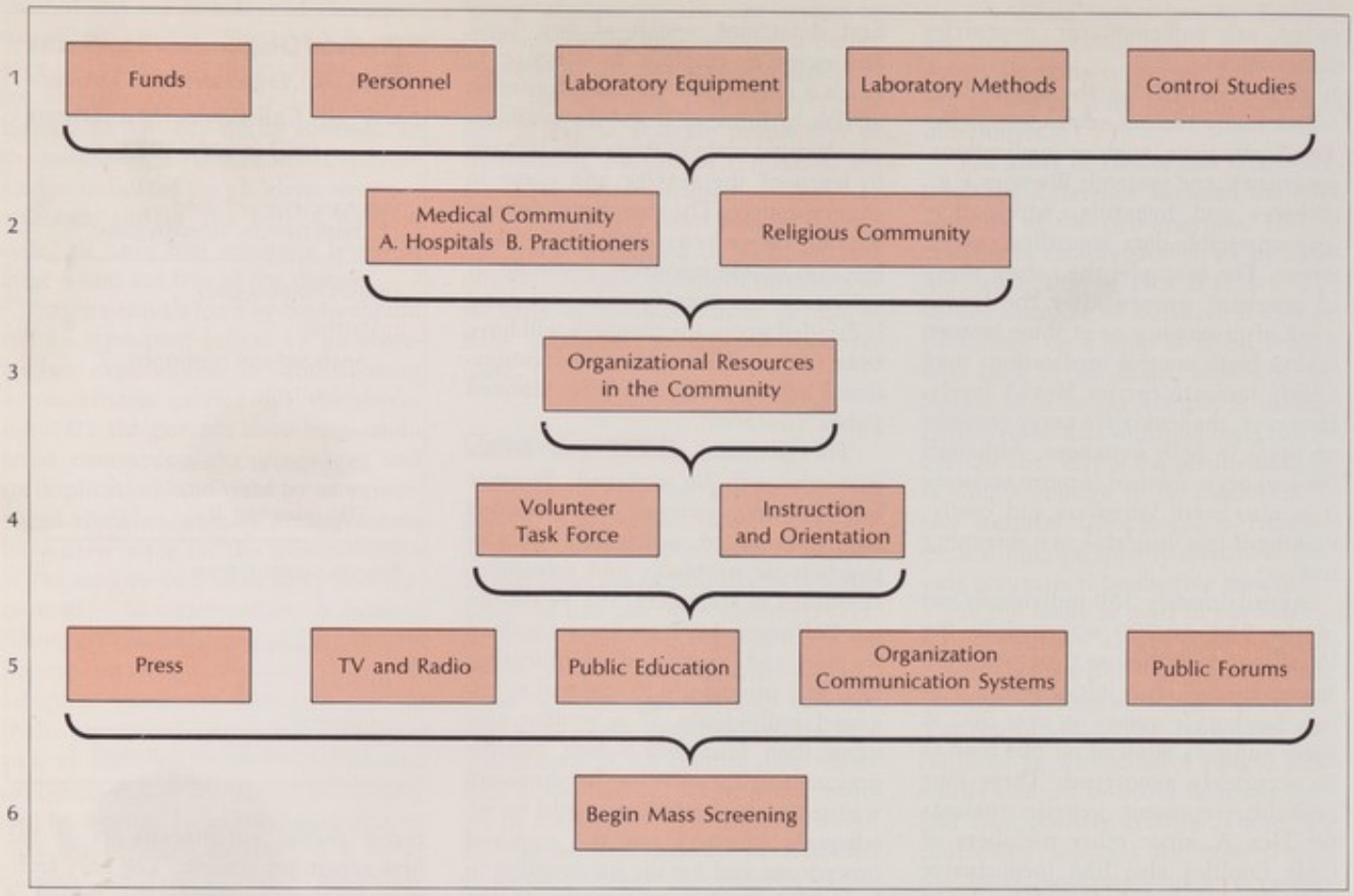
Prenatal Diagnosis of Tay-Sachs Disease: Experience at University of California, San Diego (to January 1973)

Pregnancies monitored (both parents heterozygous)		39
Fetuses homozygous		9
Aborted		8
Analyzed and confirmed	7	
Unavailable	1	
Born		1
Diagnosis at 27 weeks of gestation has clinical TSD (now age 2)		
Fetuses heterozygous		30
Born		28
(All have adequate Hex A)		
In utero		1
Aborted		1
(Chromosomal abnormality)		

Perhaps the first step in education is to reach the medical community. Most practicing physicians are unfamiliar with rare genetic conditions such as Tay-Sachs disease and unlikely to know of the latest advances associated with them. This is understandable, since even in 1963 less than one quarter of the medical schools in the United States offered any formal course work in medical genetics.

If a physician is informed, he can support the screening effort when one of his patients asks about it: "Yes, it's a rare disease and the chances are remote that it will affect you. But I believe you should consider being tested since the information is helpful however the results come out." A doctor who is uninformed, on the other hand, may dismiss the program as wasted effort on the part of ivory tower scientists.

A second audience that needs to be reached in any screening project involving minority populations is the



The strategy of a voluntary, community-wide screening program aimed at prevention of a genetic disease requires a carefully staged sequence of efforts, as indicated above. Without the basic knowledge and funds (stage 1) the necessary educational work (2) cannot be undertaken; similarly, the mobilization of volunteers and media of public communication should be delayed until their immediately preceding stages are completed. Hence, the mass screening may not be able to begin for many months.

leadership of the religious community. This is certainly true for Tay-Sachs disease. The rabbinate in Washington and Baltimore was approached months before any public announcement was made. The rabbis gave the screening effort their active as well as moral support. They delivered sermons on the topic and counseled couples and individuals. They also have a major responsibility for continuation of the program as they have the opportunity to provide educational material regarding the screening to young couples prior to marriage.

The third step in education is to reach a corps of community volunteers who can disseminate accurate program information, help generate interest when screening gets under way, and man the facilities when testings are conducted. On the order of 1,000 such volunteers, recruited from religious and community organizations, were trained in Washington-Baltimore. Through lectures and workshops they learned the basic facts about Tay-Sachs disease and then communicated them to organizations and friends at various community gatherings. Screening could not have succeeded without these volunteers. Two facts convince us this is so. First, the staff was prepared at the initial session to offer counseling to individuals who were confused or did not understand the program. But in fact little counseling was required because the volunteers had done an outstanding job of communicating what the project was all about. Second, the volunteers made it physically possible for the technicians and physicians to handle the large numbers of people who came to be screened. They took care of marking the test tubes and recording names and addresses. It is a real testimonial to the conscientiousness of these volunteers that only four tubes were mislabeled or misplaced during the screening of about 10,000 persons.

Only after this firm foundation of knowledge and volunteer support had been laid was the screening program publicly announced, approximately four to six weeks before the first community testing. Small task forces of 10 to 20 people started calling members of their parent organizations and distributing flyers and letters while the newspapers and broadcast media carried announcements of the screening data. The timing of these steps in the program was indispensable to its effectiveness; one should not start talking publicly about testing until preparations are close to final. Otherwise the interest generated will be dissipated, frustrated, and wasted.

All mass screening efforts in the Washington-Baltimore area were carried out in community facilities such as synagogues, schools, or community centers. This seemed far more appropriate than hospitals, which are not set up to deal rapidly with hundreds of "patients." These other facilities

could accommodate large numbers and were also more accessible. A schematic diagram of the layout for a typical session is shown below.

To initiate the Washington-Baltimore program, approximately $65,000 was required for equipment, laboratory and office personnel, and supplies. This sum was raised chiefly from the private community, the John F. Kennedy Institute in Baltimore, and the Maryland State Department of Health and Mental Hygiene. The program was made self-supporting thereafter by charging a voluntary fee of $5 for each individual given the serum test. This fee slightly exceeded the actual cost of the test but the difference covered the expenses of the few people who did not pay. Physicians who participated in screening sessions volunteered their time.

Genetic Counseling

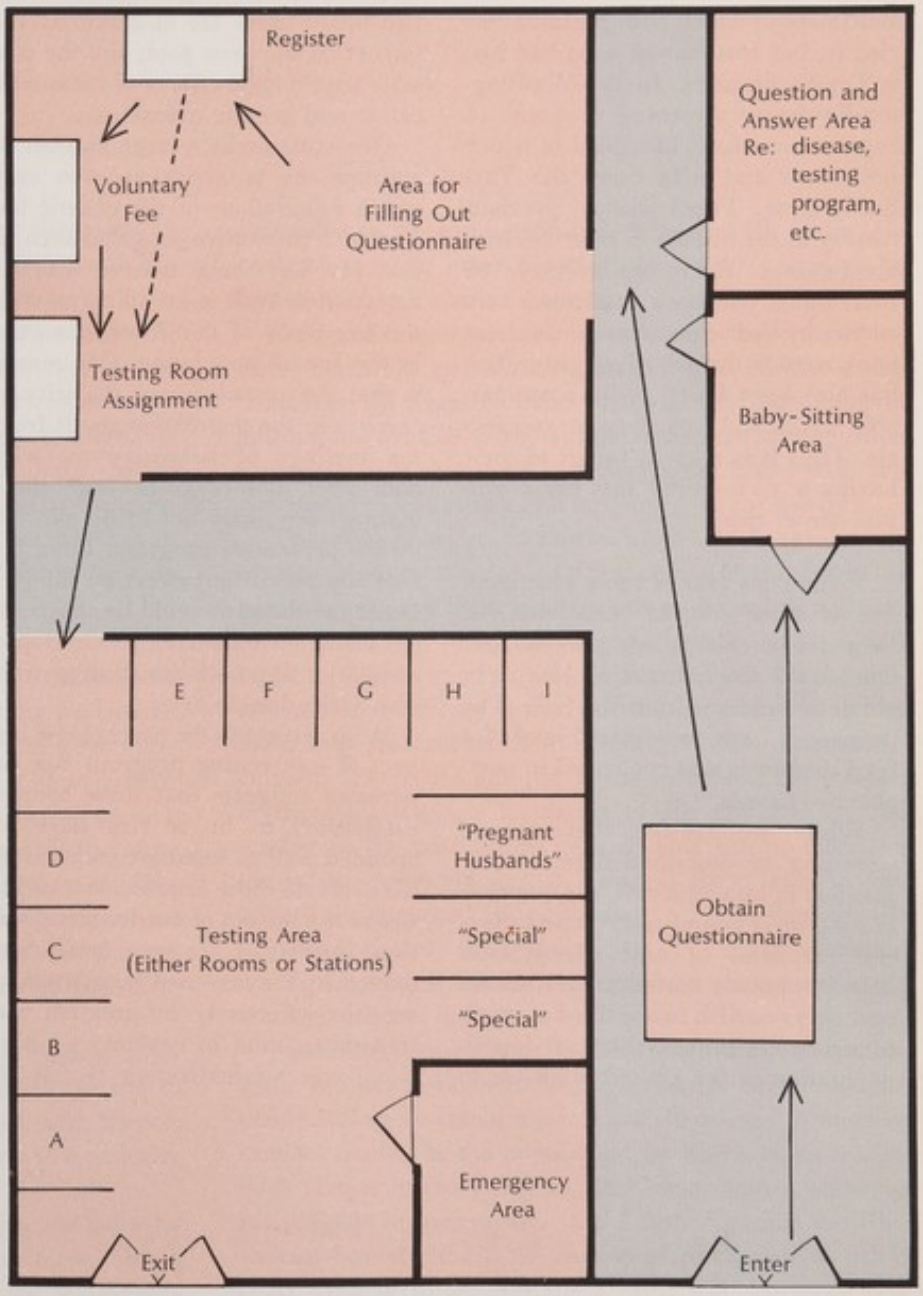
Once the technologic problems of screening for a genetic disease are solved, the question remains of helping people use the information obtained to make voluntary decisions of the highest personal importance.

Test results must be delivered in a careful and sensitive way. When there is clearly no carrier in the family a form letter will do. But in almost every other circumstance at least a telephone call from the genetic counselor is required. For example, two of the couples found at risk in the Washington-Baltimore screening had children at home under one year of age. A letter giving them the test results would have been frightening. A phone call, some verbal reassurance, and immediate testing of their children were able to allay their anxieties.

Working with members of the public under these circumstances requires a special commitment. Even before an individual or a couple is screened the mere idea of testing can be threatening. "You're going to tell me whether I have bad genes," is the general idea. Our experience suggests that intensive counseling will overcome the fears of the few couples who have serious questions before testing. And all couples in whom possible or proved risk exists must have direct contact with the counselor. Discussion must be shaped to the individuals involved. Some require more factual data, others more emotional reassurance. But this is the interface at which a screening program will succeed or fail.

Complete privacy in the delivery of results is of course essential. We have received calls from inquiring relatives of screenees, for instance. "My son-in-law was tested. Is he a carrier? I need to know because my daughter hasn't been tested." Our results are delivered only to those screened. In a program involving adults voluntarily seeking this service, we believe we are able to maintain confidentiality better than might be possible in a mandatory situation involving people of all ages or in a commercial testing setup coordinated through physicians' offices.

Once a carrier of the Tay-Sachs gene is identified, another question arises: How far should the physician go in trying to identify other carriers in the same family? We believe there is an obligation to try to do this as far as possible. In our program this is done through the identified carrier, who is provided with educational material to send or give to his appropriate relatives. This information instructs those relatives to contact the center if



Community facility, in preference to a hospital, may provide an optimal setting for mass screening programs in genetic disease prevention. Layout shown above is based on Washington-Baltimore program for Tay-Sachs disease; parking space is also needed.

they wish to be tested or obtain further information. Arrangements are then made with physicians of out-of-town relatives to have appropriate samples sent for testing. To reduce the volume of work we begin (when possible) by testing the carrier's parents to determine which side of the family carries the gene, since in most cases both are not heterozygotes.

In all of our studies of Tay-Sachs carriers (more than 250 in the East and 75 in California), we have always been able to identify at least one parent in the previous generation as a heterozygote. In fact, three older couples were found in Washington and Baltimore in which both partners carried it, but fortunately none had had a Tay-Sachs child. In the Washington-Baltimore screening program 11 couples have been identified in which both man and wife carry the Tay-Sachs gene. There was no previous history of the disease in their immediate families. Every one of these 22 individuals has been confirmed as a carrier by leukocyte assay, and at least one parent in the preceding generation has also been found to have comparable serum and leukocyte enzyme levels. Thus it is evident (short of their having a TSD child) that these couples are at risk.

Within one year of being identified, five of these couples conceived. All elected to monitor their pregnancies. One of the five showed no Hex A in amniotic fluid and cultured cells. The pregnancy was terminated and the fetal diagnosis was confirmed in post-abortive tissues.

While we believe that genetic screening as described here is fully justified in human terms, the question of its cost is bound to be raised. Several estimates of cost effectiveness have been made that indicate it would cost only one fifth to one third as much to screen the entire Ashkenazi Jewish population in the United States as to care for the affected children that would be born without a preventive program. Because the gene is so much less frequent outside the Jewish population, it would be economically unrealistic to screen the entire nation.

Implications for Other Genetic Diseases

Obviously our experience with Tay-Sachs disease has many implications for dealing with other autosomal recessive conditions once suitable carrier and fetal detection procedures have been developed. In considering the long-term effects of genetic screening, two broad areas are of concern – the impact on the gene pool, and the possible psychologic effects of mass education and genetic counseling.

One must make a large number of assumptions to arrive at even very rough calculations of the genetic impact of a preventive program such as that for Tay-Sachs disease. Carrier detection as such is not likely to alter the frequency of the Tay-Sachs gene in the Jewish population. One reason is that the overwhelming majority of carriers in the population result from the matings of heterozygotes with noncarrier homozygotes and these matings are likely not to be affected by the preventive program. Long before any significant effect on the frequency of the gene could be observed, medical science should have developed a superior approach for dealing with Tay-Sachs disease.

With respect to the psychologic impact of a screening program, our experience suggests that those couples established to be at risk have responded well to intensive counseling. To evaluate those impressions, and to assess the impact of carrier identification, interviews are now being conducted in the screened population to see how effectively information was transmitted and to evaluate whether significant stigmatization is felt by identified heterozygotes. In addition, other studies are being conducted to investigate other psychosocial considerations in mass genetic screening of this type.

A final point concerns the virtues of voluntary genetic screening as contrasted with a legally mandated form. We have been impressed with the effectiveness of the voluntary approach and we believe that the psychologic, political, and moral dangers of legislating human genetic testing far outweigh the potential medical benefits.

As a practical matter, legislation alone cannot solve genetic problems. The public must be educated for any program to work, and voluntary learning based upon enlightened self-interest is inherently more effective than coerced learning. There is also the matter of financial support. Each genetic disease program will require a substantial investment in backup services to be sure it benefits all segments of the population. Although amniocentesis has been a medically accepted procedure for some time, it is still primarily a service available to those well-off enough to pay for it. The government cannot require education about genetic disease without providing the related services.

We do not believe that every Jewish person of childbearing age should be *required* to have a Tay-Sachs enzyme assay. We do believe that every Jewish person of childbearing age should know about the disease, should know what can be done about it, and should be free to make his or her own decision whether or not to be tested. Education and individual choice is a mechanism far superior to legislation with regard to genetic programs of this sort in any population. It is hoped that the information and experience gained in Tay-Sachs screening will facilitate future implementation of preventive programs for other genetic conditions.

26

The Prevention of Rh Isoimmunization

CYRIL A. CLARKE
University of Liverpool

Considered in the context of the diseases discussed in the preceding chapters of this text, Rh incompatibility is an anomaly. It is not, strictly speaking, inherited, nor does it involve anything that can be classed as a genetic defect. The parents of an affected infant, as well as the infant itself, all possess entirely normal genotypes, whether evaluated at the microscopic (chromosome) or submicroscopic (gene) level. Rather, Rh incompatibility can be described as phenotypic abnormality produced by the interaction of two normal genotypes, proving that whether or not two wrongs can make a right, two "rights" can, in this case, make a "wrong."

Nonetheless, Rh hemolytic disease must be classed as a genetic disorder, since its ultimate, if not its proximate, cause is a particular combination of parental genotypes. As such, it is almost certainly the commonest of all genetic pathologies among Caucasian populations, having an incidence rather greater than one in 200 births. It is therefore particularly gratifying to report that we are now in sight of the virtual abolition of Rh disease. Timely treatment by techniques developed over the past 10 years can prevent all but a small fraction of these incompatibilities from developing.

My own interest in genetic disease stems from a childhood interest in butterflies. In later life, this led me to experiment with crossbreeding the insects, and thence, with my colleague Professor P. M. Sheppard, to a curiosity about how certain of their characteristics are inherited – in particular, the "mimicking" traits of form and color by which some species, otherwise palatable to birds, obtain protection from predators by simulating the appearance of unpalatable species. To omit the details, it appears that such traits are frequently transmitted by closely linked groups of genes, sometimes called "supergenes." And this brings us to Rh incompatibility, since the Rh factor (in contrast to most other blood-group traits) seems to be transmitted in a similar manner.

For practical purposes, however, the Rh "supergene" behaves much like an ordinary gene, with Rh positive dominant over Rh negative. Some 15% of the population in Europe and the U.S. is Rh–; of the remainder, a proportion something less than half is homozygous Rh+, with the rest phenotypically Rh+ but heterozygous for the trait.

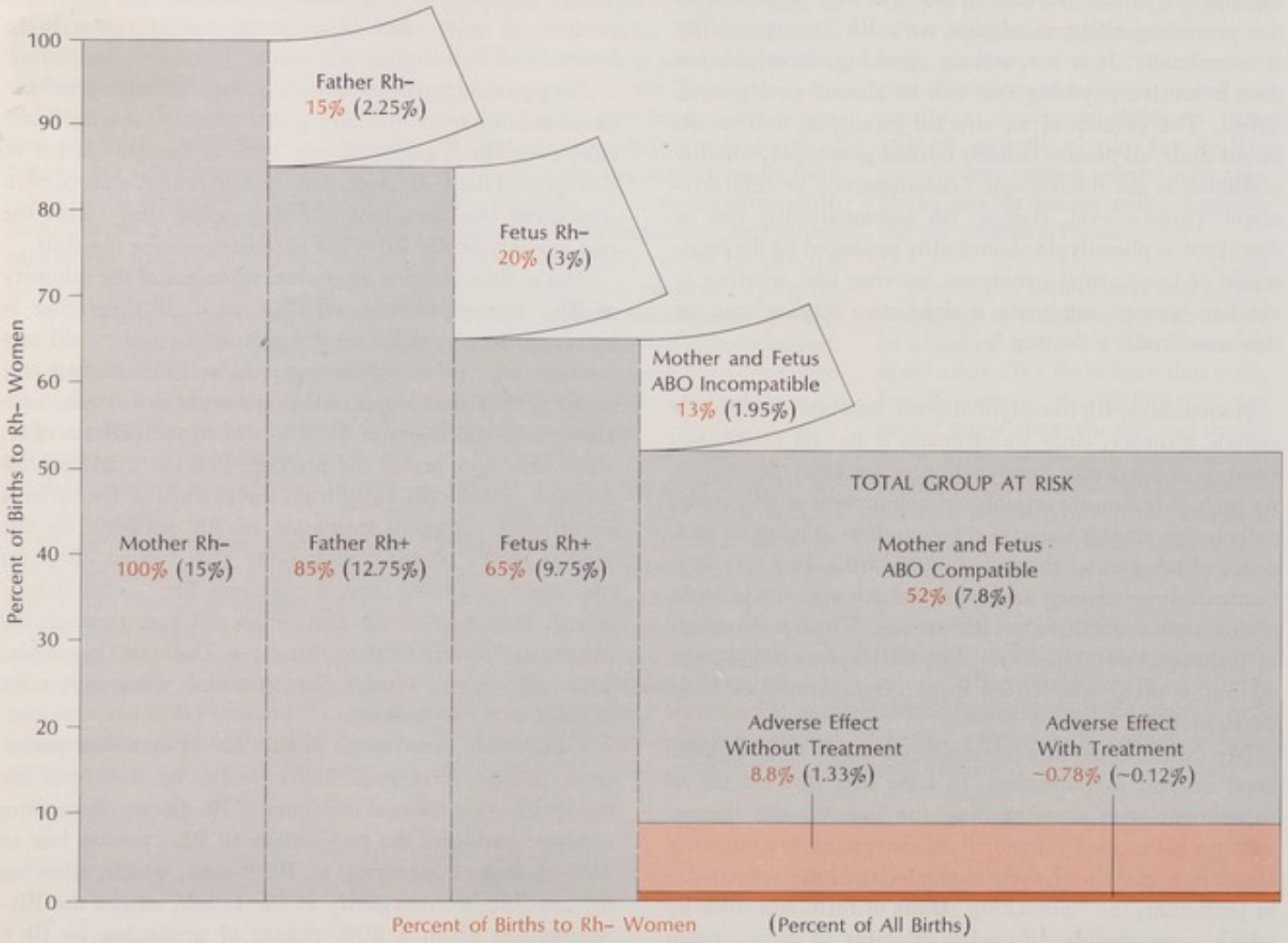
The problem begins, of course, when one of the minority of Rh– women marries an Rh+ man. If the father is homozygous, any child resulting from the union will certainly be Rh+; if he is heterozygous, half the children will be Rh+. The next step in development of the disease usually occurs at delivery of the Rh+ infant, when some of its blood may leak across the placenta into the maternal circulation. There, Rh antigen on the surface of the infant's erythrocytes triggers formation of Rh antibody by the mother in the postpartum period. This has no effect on the infant, naturally, but if a second Rh+ fetus is conceived, the maternal Rh antibodies will pass through the placenta to destroy fetal erythrocytes. The result is anemia, which, if serious enough, can produce permanent CNS damage or death in utero.

Fortunately, most steps in this rather elaborate pathogenic process work inefficiently, as can be seen from the theoretical vs the actual incidence of Rh disease. Assuming random mating in the population, an Rh– woman has an 85% chance of marrying an Rh+ man, which, allowing for possible heterozygosity in the father, means the Rh– woman has about a 60% chance of producing an Rh+ fetus. Among these, perhaps a third will represent first pregnancies and therefore not normally be at risk. The remainder, however, will still amount to something like 6% of all pregnancies. Taking the figure for British births

five years ago, before prophylaxis against Rh disease became widespread, this would mean that of the annual total of about 850,000, one would have expected some 50,000 "rhesus" babies (see below).

The actual number was only about a tenth of that figure. Most of the time, therefore, something must have gone "wrong" with the pathogenic process. Either fetal blood had not entered the maternal circulation during the first pregnancy in sufficient amounts to trigger antibody formation (which probably accounts for most of the "missing" cases) or antibodies failed to form in pathogenic quantities for other reasons. We know, in fact, that something like 20% of the time this failure occurs because the Rh incompatibility between mother and fetus is in effect cancelled out by another type of incompatibility, involving the ABO blood factors – a subject that will be discussed later. But even where this situation does not obtain, some mothers, for unknown reasons, fail to manufacture antibody. Altogether, we find that even in the absence of ABO incompatibility only about 8.5% of women in these cases will have formed detectable antibody within six months after the first delivery (little or no antibody is formed thereafter), while an equal proportion will form antibody some time during a second pregnancy – though in neither case will the antibody levels necessarily be high enough to produce damaging anemia in the second fetus. Other studies, in fact, have shown that while maternal antibody titers below a certain level (the actual figures depend on the procedures employed) mean that the fetus is pretty certainly *not* in danger, titers above that level mean merely that it *may* be endangered.

Serious Rh incompatibility has, of course, been treated for more than a generation by means of exchange transfusion immediately postpartum; it has been estimated that this procedure, despite its inherent dangers, salvages something like 70% of the endangered infants – though not all of these escape neurologic damage. More recently, improved prenatal diagnostic techniques, involving spectrophotometric examination of amniotic fluid obtained by amniocentesis, have provided much more sensitive and reliable indicators of the degree to which fetal survival is threatened by hemolysis. If a severe threat occurs late in pregnancy, labor can be induced immediately; in some other cases, the problem has been combated by intrauterine transfusion. Together, these advances



Some of the factors that explain why the incidence of Rh disease is much lower than expected are shown graphically above. Of all Rh– mothers, making up some 15% of all mothers, only about half produce an Rh+ fetus with which they are ABO compatible, i.e., are at risk of becoming isoimmunized. Of these, only a sixth will actually show detectable antibodies by the end of a second Rh+ pregnancy, even without treatment; with treatment, the incidence of immunization drops by about 90%.

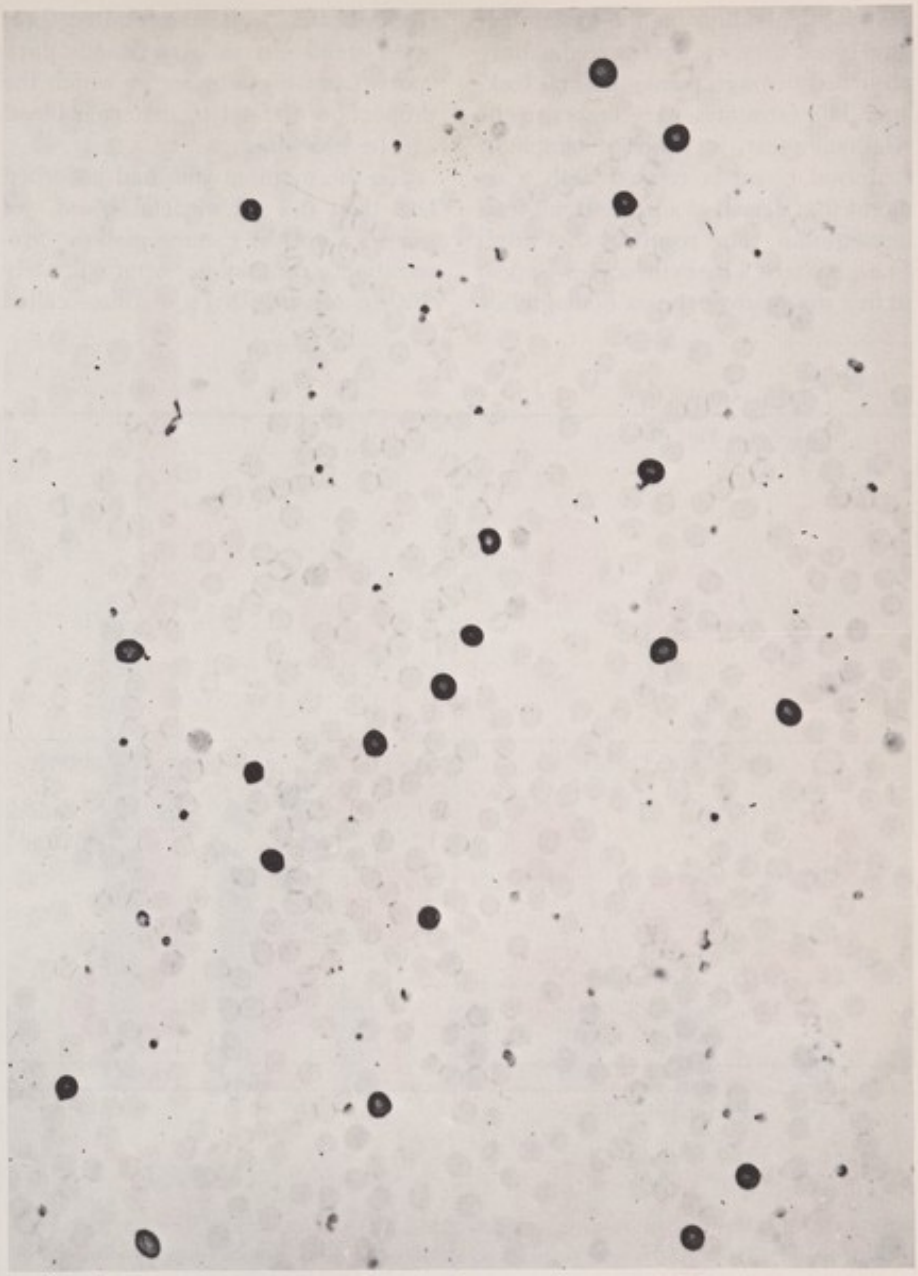
have reduced fetal wastage in Rh disease to below 10%.

All these advances were obviously encouraging, but though they minimized the Rh problem they could not be said to have solved it. Moreover, some of the procedures involved were difficult and costly, and none was completely without risk to mother, fetus, or both. A better approach was obviously needed, and the phenomenon of "protection" against Rh disease by ABO incompatibility gave some hope that such an approach could be devised. Exactly how this protection works is still in dispute; the most likely explanation seems to be that the maternal AB antibodies attack any fetal erythrocytes that leak into the maternal circulation and thereby render them immunologically ineffective. Presumably they are segregated in some part of the body (e.g., the liver) inaccessible to the mother's antibody-manufacturing cells.

But how could a similar effect be induced artificially when the mother lacked these AB antibodies? Injection of the antibodies themselves was obviously out of the question, since they would attack the maternal as well as the fetal erythrocytes. My associates, my wife, and I all pondered the problem for many months and then it suddenly dawned on us that the answer might be to give the mothers injections of anti-Rh, for if AB antibodies could neutralize incompatible cells, there seemed no reason why Rh– antibodies could not do the same to Rh+ cells.

The first attempts to test this hypothesis, mainly carried out by my colleague Dr. Ronald Finn, were, however, less than successful. The procedure involved producing anti-Rh plasma by injecting Rh+ cells into volunteer Rh– male donors, and then injecting the plasma into other male Rh– volunteers who had previously been injected with Rh+ cells labeled with radiochromium. The anti-Rh plasma did indeed knock out a large proportion of the labeled cells, but did not prevent the formation of antibody; six months later, antibody levels in the volunteers were actually higher than in untreated controls.

We persisted, however, and eventually discovered that the failure was due to our giving the "complete" antibody, which, though it destroys Rh+ cells, leaves the residue still antigenic. When we changed our methods of preparation to produce the "incomplete" antibody – a protein of considerably lower molecular weight – we found that it not only destroyed the cells but in most cases prevented antibody formation as well.



Assay for fetal blood involves treating sample of maternal blood with a reagent that dissolves adult but not fetal hemoglobin; fetal erythrocytes then show up clearly against the transparent "ghosts" of the maternal cells. This photomicrograph indicates that the mother's circulation contains approximately 1.0 ml of fetal blood.

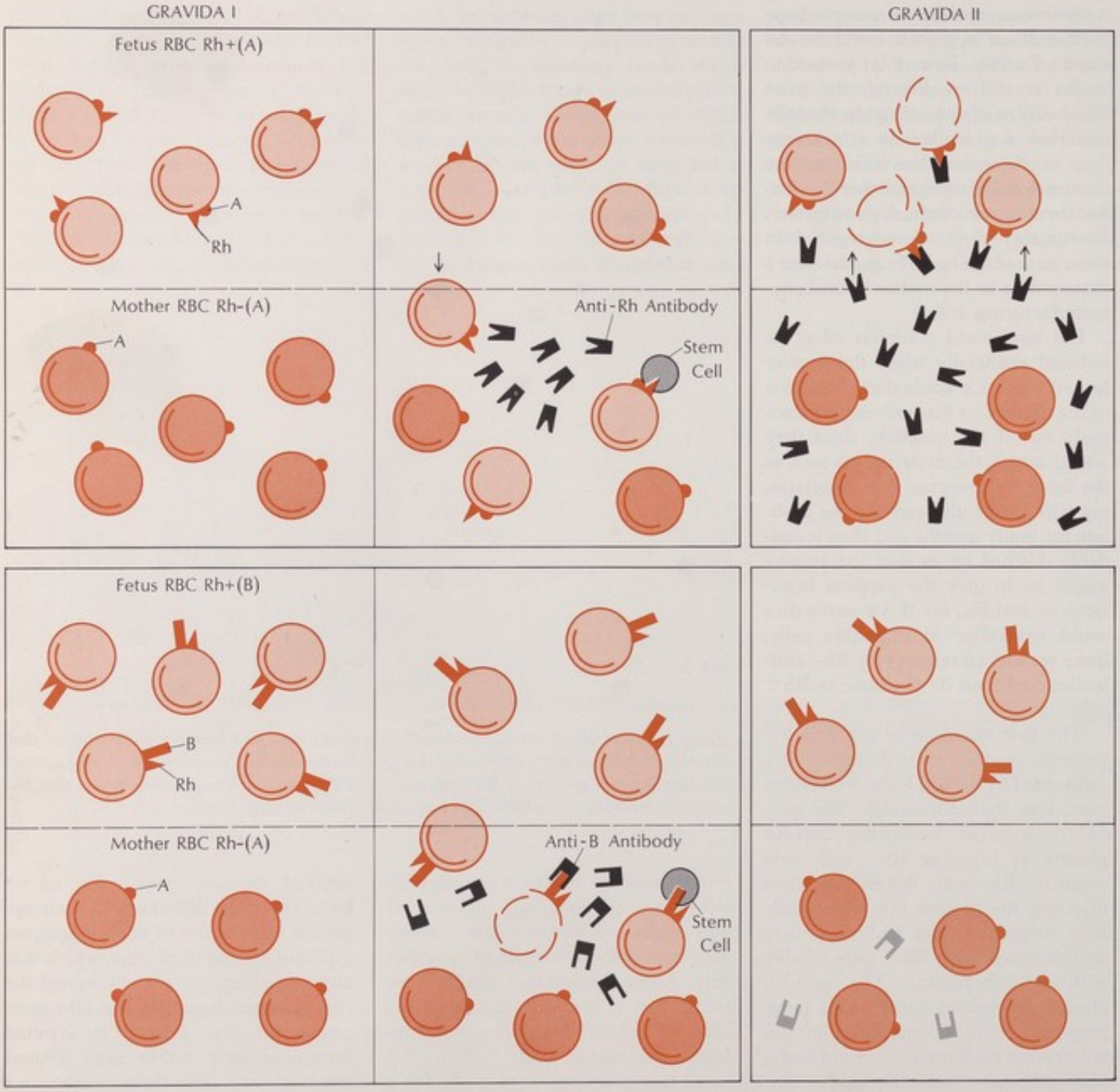
At this point we were ready for a clinical trial. Meantime, however, we had become aware of the work of a New York group composed of Vincent Freda, John G. Gorman, and William Pollack, who by a somewhat different chain of reasoning had independently reached the same conclusion as we had. The main difference in their approach was the use of anti-Rh gamma globulin instead of plasma, which was not only safer, since it eliminated the risk of serum hepatitis, but also more convenient, since it could be injected intramuscularly rather than infused intravenously. For these reasons we shifted over to gamma globulin rather than serum for clinical trials, which by this time were also going forward at other centers in several countries.

Our own clinical protocol involved separating the treated women into two

groups, depending on the amount of fetal blood they were estimated to have absorbed through transplacental leakage. The estimates were based on the Kleihauer test, in which a sample of maternal blood is treated with a reagent that dissolves adult but not fetal hemoglobin. The result is that adult (i.e., maternal) erythrocytes appear under the microscope as hemoglobin-free "ghosts," while the fetal erythrocytes stand out as conspicuous dark spots, from the number of which the proportion of fetal to maternal blood can be estimated.

To the mothers who had absorbed less than 0.2 ml of fetal blood, we gave 1 ml of the gamma globulin preparation, containing approximately 200μg of anti-Rh (sometimes called anti-D) globulin. Those who had absorbed 0.2 ml or more of fetal blood, and were therefore at presumably higher risk of developing antibodies, received five times this dose. In addition, of course, we matched both groups with untreated controls.

After four years, the results were impressive. Of the treated low-risk (less than 0.2 ml of fetal blood)



The natural protection against Rh immunization afforded by ABO incompatibility between fetus and mother is depicted. When major blood group identity exists (e.g., A:A) the leakage of fetal cells into the maternal circulation at delivery of a primipara (top) will elicit an anti-Rh antibody response in the mother. When the sensitized mother becomes pregnant again, these antibodies, boosted by new exposure, will enter the fetal circulation, causing hemolysis. However, if there is an intercurrent ABO incompatibility (e.g., A:B), any fetal red blood cells entering the maternal blood stream in the first pregnancy are likely to be immunologically neutralized too rapidly to permit Rh immunization and its sequelae in later pregnancies (lower schematic).

women, only 0.36% showed detectable antibodies six months after delivery, as against 3.6% of the controls. In the high-risk group, the incidence of immunization was still only 0.6%, as against more than 21% for the high-risk controls – the latter figure, incidentally, fully confirming that these women were indeed at a far greater risk than the others.

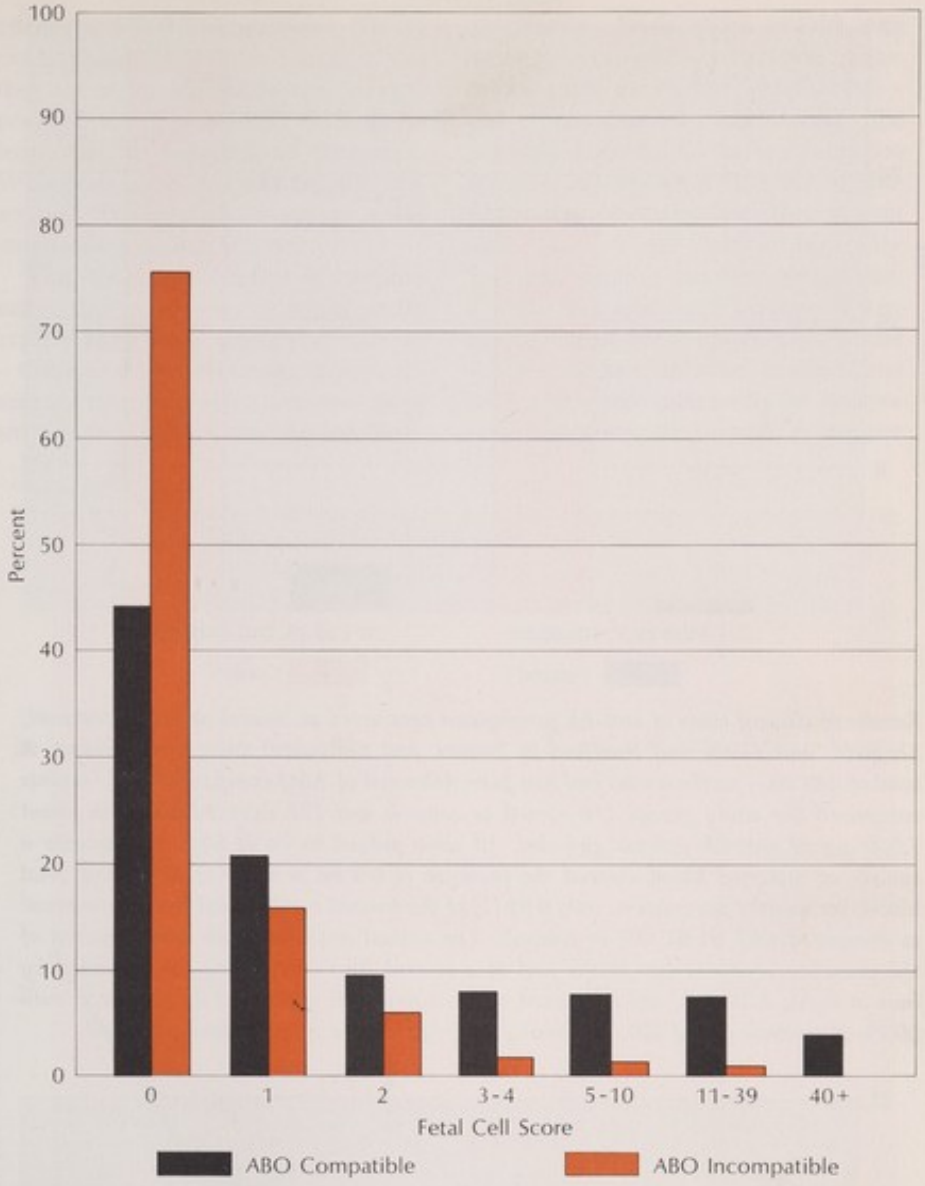
By 1971, we were able to publish a final report on both groups, this time focusing on those who had undergone a second Rh+ pregnancy. For the low-risk, second-pregnancy mothers, the total incidence of immunization, including those who had previously been immunized, was still only 1.8%, as against 10% for a control group; for the high-risk group, the proportion was 2.3%, as against 31% for high-risk controls.

Summarizing all these findings, and adding some additional data that have accumulated since the original study, we can say that in an untreated group of Rh– women, roughly 17% will be immunized by the end of their second Rh+ pregnancy; with treatment, this figure is lowered to less than 2%, a reduction of nearly 90%. When we say "immunized," of course, this does not necessarily mean immunization to a degree that will kill the second fetus, but merely sufficient antibodies to cause a significant degree of fetal hemolysis. In fact, a large proportion of this residuum can be saved by the older techniques already discussed, such as induced delivery and exchange transfusion.

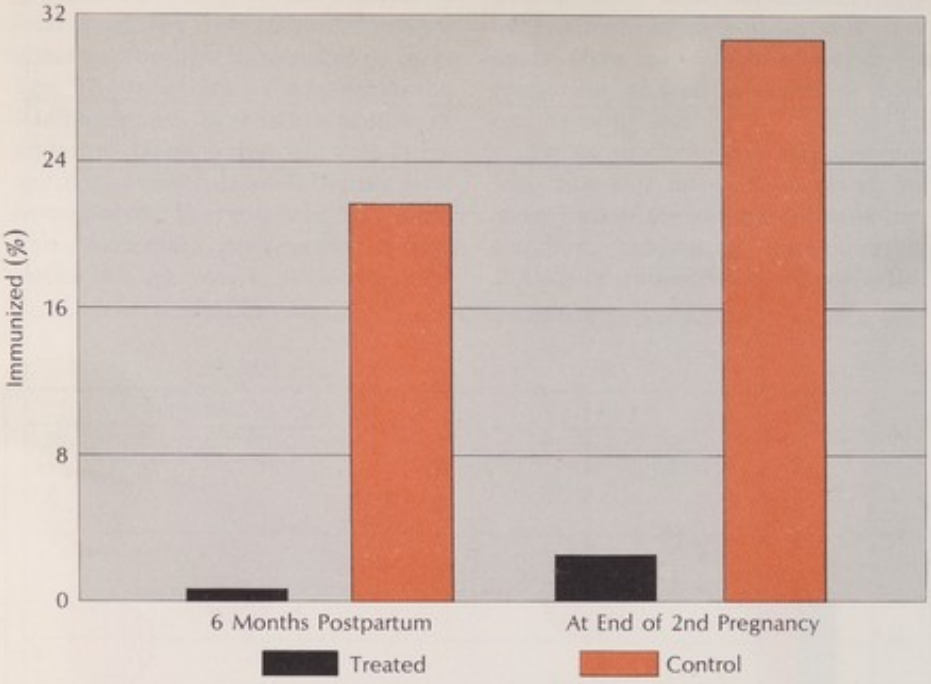
Since early 1971, prophylaxis against Rh– sensitization has been routine in Great Britain for most Rh– women giving birth to Rh+ infants, where the ABO groups are compatible between mother and child. The pregnant woman is typed during prenatal examination; if she turns out to be Rh–, her husband is typed also. If, as is likely, he is Rh+, then the infant is typed postpartum to determine whether, on the basis of both Rh and ABO incompatibility, the mother is at risk of immunization.

Should this be the case, a sample of her blood is assayed by the Kleihauer test for the presence of fetal erythrocytes. In the great majority of cases, a quick examination of the slide will reveal only an occasional fetal cell, or none, putting the woman into the lowest risk group. For these individuals, the standard dose has been reduced to 100 μg of anti-Rh injected IM into the deltoid; while this was done initially to conserve supplies of the special globulin, the failure rate seems no greater than with the earlier dose of double that figure. For the minority who cannot be screened out by superficial examination, a detailed count is run on the fetal erythrocytes to determine whether the woman needs a larger dose, and if so, how much. If necessary, she may receive as much as 1,000 μg, but the majority even in this group still require only the standard dose of 100 μg.

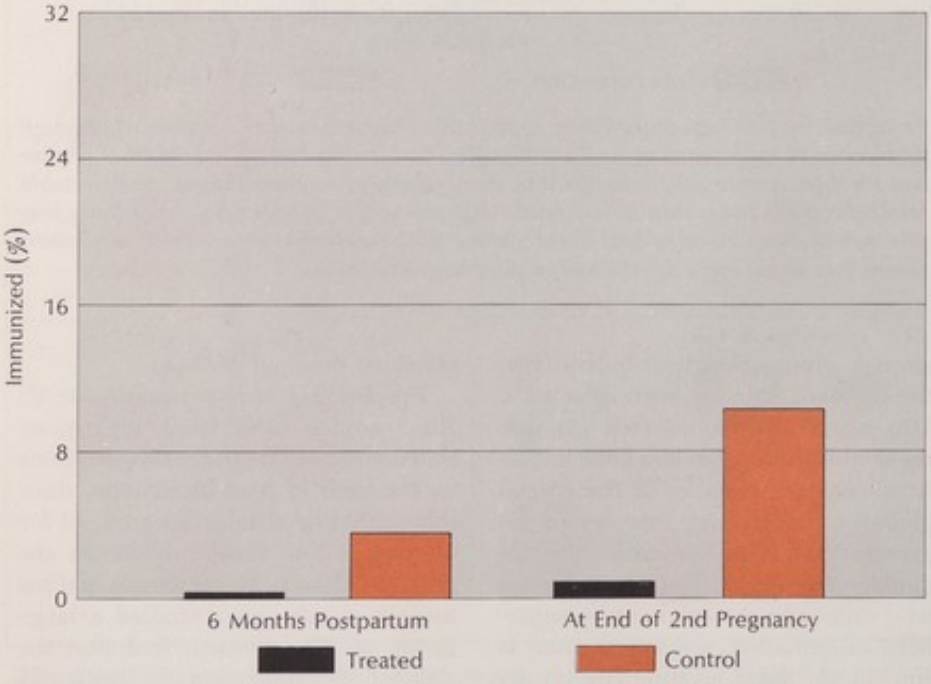
Prophylaxis is also routine for all Rh– women who have undergone abortions. Here there are no exceptions on the basis of fetal blood type, since this cannot be reliably ascertained for abortuses. For these individuals the standard dose is 50 μg. Freda and his associates, who have studied a large group of Rh– women, find that the risk of immunization is "virtually negligible" for abortions at fetal age one month or less, but becomes appreciable (about 2%) at two months and substantial (about 9%) at three months or more.



Protection by ABO incompatibility against Rh isoimmunization is achieved through destruction or segregation of fetal erythrocytes. In Liverpool study, where Rh– mother and Rh+ fetus were ABO incompatible, three fourths of mothers showed no detectable fetal cells, while fewer than 2% had fetal-cell scores of 5 or greater. (A score of 5 suggests presence of about 0.2 ml of fetal blood.) In the ABO-compatible group, 18.5% of mothers scored 5 or above and only 44% had no detectable fetal cells.



Results of clinical trials of anti-Rh prophylaxis conducted at several centers (Liverpool, Sheffield, and Leeds and Bradford in Britain, and Baltimore) are graphed above. A total of 349 Rh– mothers who had just been delivered of ABO-compatible Rh+ infants comprised the study group; 176 served as controls and 173 were treated with about 1,000 μg of anti-Rh gamma globulin. All were judged to be at high risk because a sample of maternal blood showed the presence of 0.2 ml or more of circulating fetal blood. Six months postpartum, only 0.6% (1) of the treated mothers had been immunized as compared with 21.6% (38) of controls. The crucial test came with determination of the presence of antibodies at the end of a second Rh+ pregnancy. As indicated by bars at right, 2.3% (2 of 86) of treated mothers now had antibodies as compared with 30.7% of controls (20 of 65), confirming the high degree of protection achieved.



Clinical trial at Liverpool investigated effectiveness of anti-Rh prophylaxis in low-risk mothers (circulating fetal blood less than 0.2 ml); treatment comprised about 200 μg of anti-Rh gamma globulin. While incidence of anti-Rh antibodies is lower among untreated low-risk than among high-risk mothers (see graph at top), it is still 10 times higher six months postpartum than in low-risk treated group (0.36% vs 3.6%), while at the end of a second Rh+ pregnancy, it is more than four times higher (10.2% vs 2.3%).

The gamma globulin used in Great Britain is obtained from plasma donated partly by naturally immunized women and partly by artificially immunized male and postmenopausal female volunteers. The donors, like all blood donors in our country, are unpaid, and the gamma globulin is supplied free under the National Health Service for up to two Rh+ pregnancies per patient. With the general availability of contraception, backed by abortion where necessary, very few Rh– women are likely to have additional children should they happen to be sensitized by a third, untreated pregnancy.

Similar prophylactic programs are now under way in virtually all industrialized countries, though the protocol and dosage varies somewhat from country to country and sometimes from place to place. I might note, by the way, that there have been essentially no adverse reactions to the treatment, apart from temporary soreness and sometimes swelling at the injection site. I know of only one severe reaction (in Canada), which may have resulted from a rare (about 1:1,000) absence of IgA combined with a trace of this substance in the anti-Rh IgG, leading to the formation of anti-IgA antibody in the mother. Reactions of this sort not uncommonly occur after blood or plasma transfusions, but in the case of Rh prophylaxis we are, of course, giving much smaller quantities of material – equivalent to perhaps 25 ml of plasma.

This large-scale prophylactic program can be expected to be reflected eventually in a sharply lowered death rate from neonatal hemolytic disease, but at this writing the overall figures have dropped only slightly. The main reason, of course, is that the mortality occurs chiefly among women who have had two or more pregnancies – and who would, therefore, have become immunized before prophylaxis became general. As this backlog is whittled down, we should see a considerably more marked reduction in fetal loss from erythroblastosis. I have also heard it suggested that the very slow fall in the neonatal death rate may be due in part to the more general use of premature induction and fetal transfusion, which has permitted more jaundiced infants to be born alive – but to succumb, in many cases, shortly

after birth. If this is the case, such an "excess" in neonatal deaths should have been balanced by a drop in stillbirths, but I do not know whether such a drop has in fact occurred. In any event, deaths of this sort would also be mainly residuals of the preprophylaxis era, and would therefore be expected to drop off as fewer and fewer Rh– mothers become immunized.

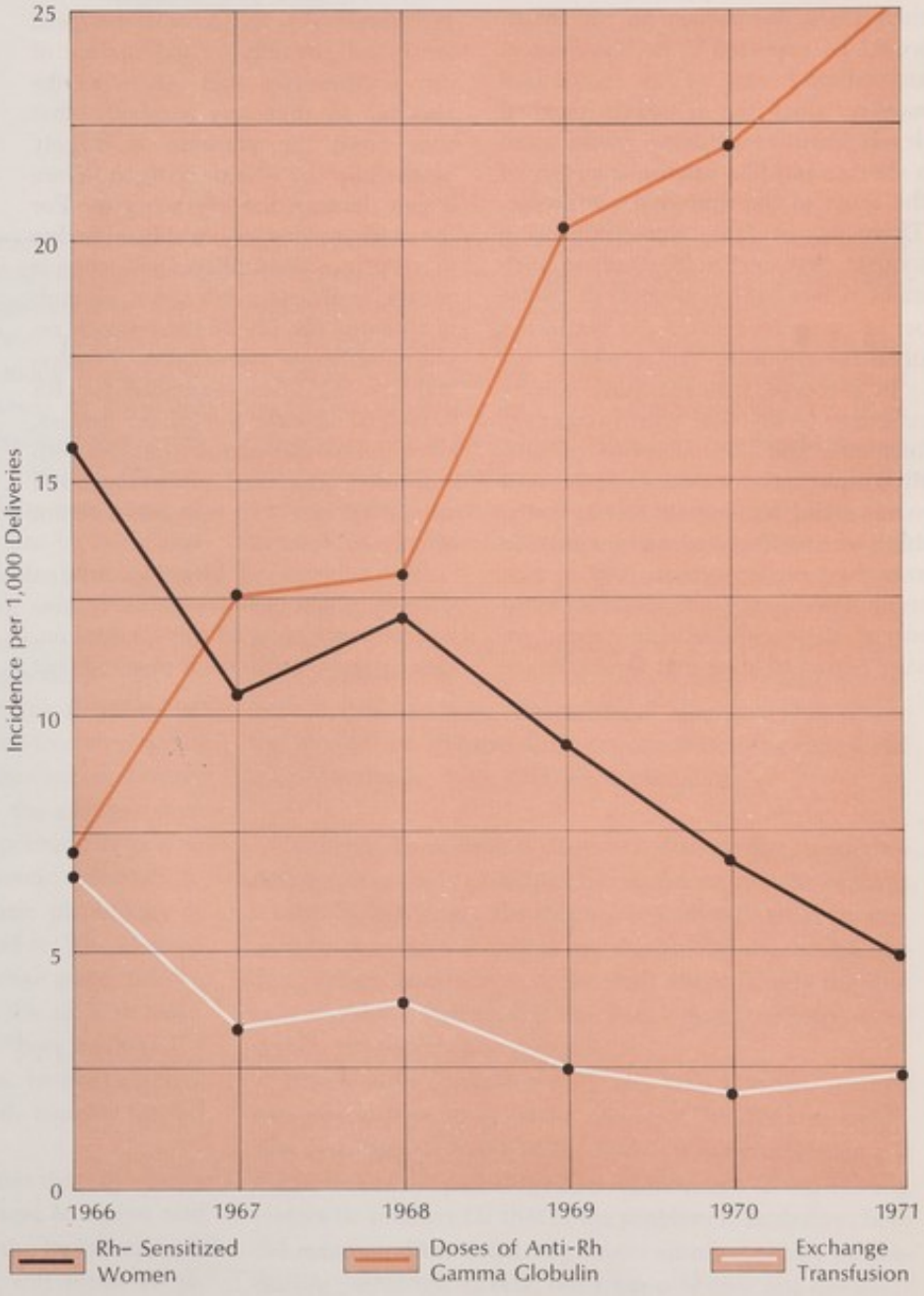
The approximately 90% success rate in our test groups – which is about the same as that found in other countries – classifies anti-Rh prophylaxis as highly effective by any standard. On the other hand, even this accomplishment still implies an eventual incidence of Rh disease of something like one birth in 2,000, which is of the same order of magnitude as most of the more common genetic diseases. One would like to do better. To consider how this might be achieved, we must examine the possible reasons for failure.

In the case of women showing what might be called primary failure – formation of antibodies within six months of the first Rh+ delivery – one likely explanation is immunization without overt antibodies ("priming") in the course of the first pregnancy, rather than postpartum, so that the prophylaxis would have been given too late. While the bulk of transplacental hemorrhage occurs around the time of delivery, it occasionally takes place earlier. The priming would have cleared the fetal erythrocytes from the mother's circulation and overt antibody formation would later develop despite subsequent administration of gamma globulin. In this connection it is interesting to note that several of these women in fact showed *no* fetal cells when tested after delivery, yet were manufacturing antibody when tested at six months.

To the extent that predelivery immunization is responsible for failure of prophylaxis, this could perhaps be combated by prophylaxis during pregnancy. A Canadian group is now engaged in testing this hypothesis by administering anti-Rh during the third trimester, and we await their findings with interest.

A certain proportion of the failures might have been due to previous Rh+ pregnancies ending in (illegal) abortion, which the women might not have cared to mention during their second (presumed first) pregnancies. To the extent that abortions become legal (as they are in the United Kingdom) and prophylaxis is provided, such problems can be expected to disappear. Where abortion remains legally proscribed, "failures" of this category will inevitably continue to occur.

The same phenomenon of immunization during pregnancy could partly explain the other category of failures – those in whom antibody, though absent six months postpartum, was present by the end of the second Rh+ pregnancy. Such cases would presumably respond to whatever measures – such as earlier prophylaxis – prove capable of minimizing the primary failures. In some instances, however, a somewhat different mechanism may be operative: the woman may have been immunologically "primed" during her first pregnancy (or possibly afterward because of partial failure of the prophylaxis), in the sense that her immune mechanisms were triggered sufficiently to remove fetal cells from the circulation but not



Rise in gamma globulin prophylaxis at Mill Road Maternity Hospital, Liverpool, is reflected in reduced incidence of Rh sensitization in mothers and of exchange transfusions in neonates; slower fall in latter is due to "backlog" of mothers sensitized before prophylaxis was available (data and graph from B. M. Hibbard).

to the point of producing detectable antibody postpartum. That this can occur is indicated by experiments with Rh– male volunteers, in some of whom labeled Rh+ erythrocytes disappear at an abnormally rapid rate even though no antibody is detectable. In a woman primed in this manner, even a minimal transplacental hemorrhage during the second pregnancy would trigger the preset antibody mechanism, so that by the end of pregnancy the antibody level could reach a potentially pathogenic level.

It is worth noting, however, that in such cases the impact on the infant would be expected to be less than in an ordinary case of an immunized mother, since the relatively gradual development of antibody would mean a shorter and less intense exposure of the fetus to the maternal antibodies. There is, in fact, some reason to suspect that erythroblastosis in such cases is less severe, though the number of pregnancies thus far is still too small for anything like certainty.

In any case, I do not think it oversanguine to say that with further refinements and improvements in anti-Rh prophylaxis, we can look forward to improving our success rate by something like another order of magnitude over the next few years. And as pre-prophylaxis mothers increasingly move out of their childbearing years, we have reason to hope that the incidence of Rh disease will ultimately drop to something like one in 20,000 births, which would place it among the rarer genetic pathologies.

It is interesting to compare the effects of ABO incompatibility with Rh disorders. The former is more than twice as common as Rh incompatibility, but only in exceptional cases does it lead to erythroblastosis fetalis. There are two reasons for this. First, the bulk of the mother's anti-A or anti-B is of high molecular weight and does not pass the placenta. Second, in something like 80% of fetuses (known as "secretors") the A or B antigens are found not only on the surface of the erythrocytes but also in the plasma, so that any antibody that does cross the placenta is largely "soaked up" by plasma antigen before it can damage the erythrocytes. For the same reasons, erythroblastosis due to ABO incompatibility, even when it occurs, is almost never severe enough to threaten the life of the fetus or require exchange transfusion. Prophylaxis on the lines suggested for Rh hemolytic disease would be useless, since the mother already has her own naturally occurring antibodies, and nature has seen to it that these are not often troublesome.

The principle of blocking antigen sites by giving passive antibody also has relevance to renal transplantation. Here, passive antibody is given so that the graft is partially coated with humoral antibody and this prevents the cellular antibody, which is the damaging one, from destroying the graft. The giving of the humoral antibody is a "holding operation," for in due course it is hoped that the recipient of the graft will himself form humoral antibody against it, and this will in its turn coat some of the antigen sites on the graft and protect it from the invading lymphocytes. The object, therefore, is not to prevent the recipient of the graft from making antibody at all but to protect the graft while he does, keeping in mind that it is the cellular antibodies that are so lethal to the graft.

In an actual case, a boy of eight who needed a kidney and whose mother was to be the donor, the procedure utilized was to inject her lymphocytes into the husband and thereby raise "anti-mother" gamma globulin. Papain treatment of the gamma globulin removed the Fc (complement fixing) piece so that the kidney cells were coated, not lysed. The boy was given this gamma globulin at the same time as, and after, the grafted kidney, and appeared to need much less immunosuppressive treatment than usual. The long-term difficulty with this boy was to establish the right balance so that enough antigen sites were left exposed to stimulate humoral antibody formation but not so many that the graft was destroyed by the cellular response.

27

Genetic Disease: The Present Status of Treatment

R. RODNEY HOWELL
University of Texas

In this collection of articles on genetic disease, a recurrent theme has been the overall inadequacy of present therapeutic methods for dealing with these diverse conditions. Nonetheless, it should be clear from a number of preceding chapters that the situation today is considerably more encouraging than it was twenty or even five years ago. Though most genetic diseases remain recalcitrant to any known form of treatment, a sizable number of them are now being tackled by various promising, if still experimental, methods, while a few can be fairly said to be substantially correctible if not curable. It is the purpose of this chapter to review the current state of the art in this area of clinical practice, as well as to examine some of the potential future developments in the field.

Essentially, all of my comments refer to disorders inherited in a recessive fashion. There are few treatments currently available for the dominantly inherited conditions. Also, the only therapeutic approach to the chromosomal disorders is by their prevention through prenatal detection.

A good way of approaching the therapeutic side of medical genetics is to review briefly the basic physiology of genetic disease, insofar as we understand it. The medical problem begins, of course, with a defective gene, that is, with an altered DNA molecule, composed like all such molecules of a long sequence of the four "base residues" – cytosine, guanine, thymine, and adenine. In the defective gene, one or more residues are altered, usually having been replaced by another of the four.

Since it is the sequences of base residues that, by means of the "genetic code," specify the sequences of amino acid residues in the proteins synthesized by the body's cells, it follows that an altered DNA molecule will mean an altered protein, containing one or more "wrong" amino acids. We now have reason to think that many such anomalies are formed but that they do not impair the protein's ability to perform its metabolic job, whatever it may be; it is only when the chemical anomaly engenders a functional anomaly that the individual is in trouble.

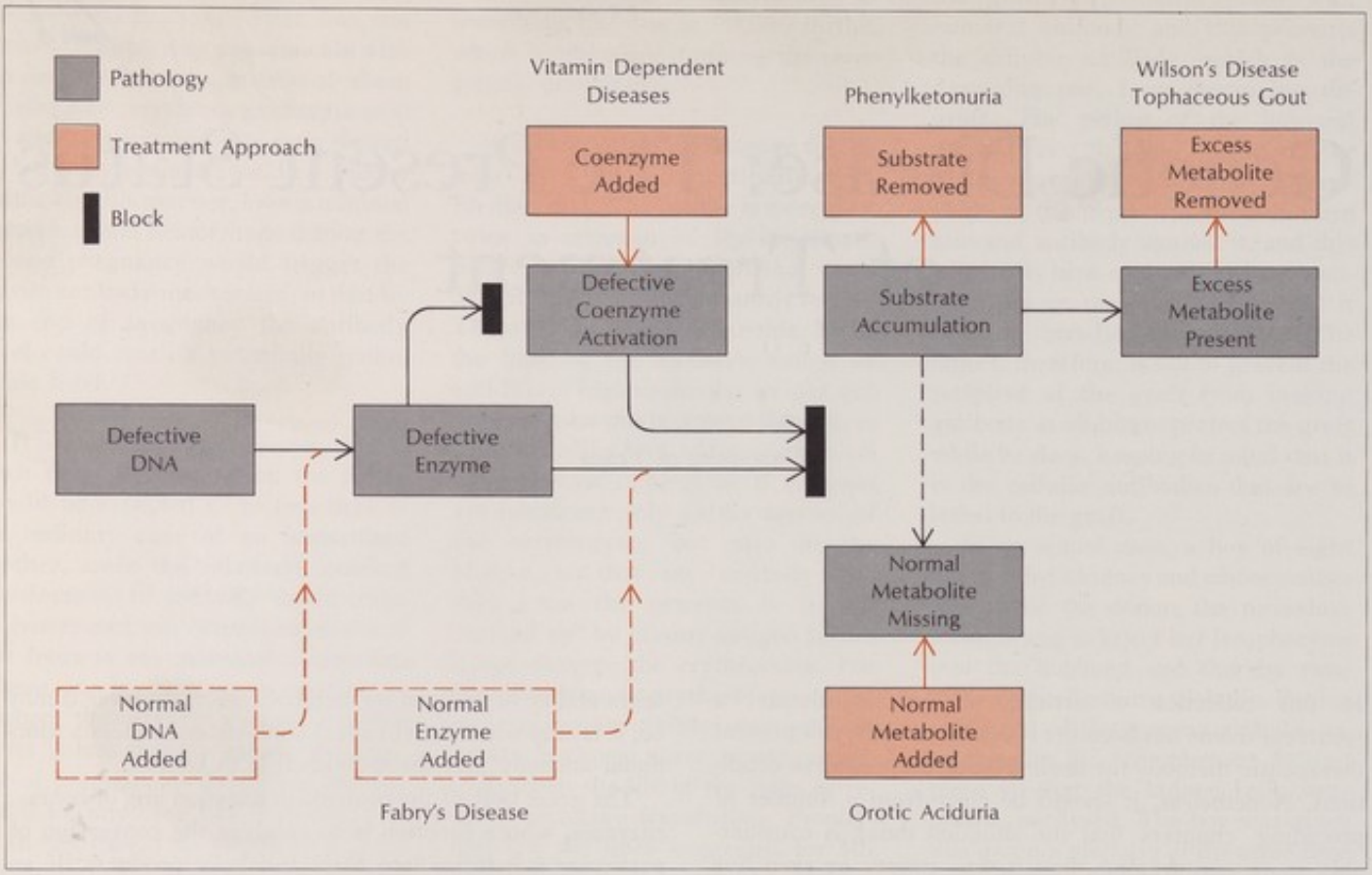
The great bulk of the proteins in question are, of course, enzymes, whose function is to catalyze the conversion of particular substrates into their metabolic products. If an enzyme is functionally inadequate, the result will be two biochemical abnormalities: an excess of the substrate and/or of abnormal quantities of products produced by its conversion via alternate metabolic paths, and a deficiency of its normal product; either phenomenon or both may produce pathologic effects on the body's tissues. (This explanation is somewhat simplified and also does not apply to dominant conditions or to the so-called multifactorial genetic diseases or to chromosomal abnormalities, where the physiologic abnormalities are much more complex and correspondingly more difficult to elucidate.)

It should be evident that genetic disease due to enzyme deficiencies can, in principle, be treated on any one of three levels: by correcting the biochemical abnormality, by correcting the enzyme defect or, finally, by correcting the gene defect. In practice, as we shall see, it is only the first two of these – principally the first – that currently show significant therapeutic potential.

Biochemical correction naturally falls into two general categories, depending on the nature of the disease: reduction of substrate levels in the body (if substrate excess is the source of the pathology) or replacement of the missing metabolic product (if that is the problem). In theory, both the excess and the deficiency might simultaneously be producing pathologic effects, but I know of only one disease – homocystinuria – in which this has been clearly demonstrated to be the case. More will be said about this later.

In the first category, the most obvious – and also the

Mechanisms in Genetic Metabolic Disease and Approaches to Therapy



Chain of events in genetic metabolic disease proceeds from DNA defect through enzyme defect to various biochemical abnormalities. In principle, therapeutic intervention can occur at any point in the chain; in practice, treatments at any but the biochemical level are mostly experimental at best. Diseases cited are examples only; for a complete list see pages 278 and 279.

most successfully treated – example is phenylketonuria. Here the basic problem appears to be high serum levels of the amino acid phenylalanine resulting from the inactivity of the enzyme phenylalanine hydroxylase. The phenylalanine excess, possibly by competitively interfering with the utilization of other amino acids in protein synthesis, hinders the growth and development of the central nervous system. In addition, some abnormal metabolites of phenylalanine are suspected of exerting various toxic effects. In any case, the disease, as we know, can be treated quite successfully by limiting dietary intake of phenylalanine during childhood – i.e., until growth of the central nervous system is complete. Since PKU treatment has been discussed in Chapter Six by Eugene Knox and Chapter 22 by Robert Guthrie, we need say no more about it here. We should note, however, that similar dietary limitation treatments are being employed, at least experimentally, in such conditions as hereditary tyrosinemia and maple syrup urine disease and are well established in galactosemia.

An alternative approach to reducing body levels of a troublesome metabolite or its abnormal by-products is by administering drugs or other substances that will in one way or another accelerate excretion of the compound(s) in question. A now classic approach of this type is the treatment of tophaceous gout with uricosuric agents. These drugs lower serum uric acid by increasing its renal excretion; the tissue deposits of uric acid salts (tophi) are then mobilized, with improvement in joint symptoms. Another version of this basic approach is employed in Wilson's disease. Although the fundamental metabolic defect in this autosomal recessive condition is not definitely known, there is no doubt that its pathologic effects are produced by accumulation of copper in the tissues. This is now corrected by administering such compounds as penicillamine, which by binding to the copper ions enables the body to eliminate them much more readily.

Turning to the second category of treatment at the biochemical level – the diseases in which the pathology stems from product deficiency rather than substrate excess – we find an excellent example in orotic aciduria. This condition is caused by the failure of two sequential enzymes to transform orotic acid into uridine monophosphate, which in turn becomes uridine triphosphate. So far as we can tell, the excess of orotic acid that gives the condition its name is benign, but the concomitant deficiency of the uridine derivatives is not. Since these compounds are essential constituents of RNA and DNA, their deficiency hinders the synthesis of the nucleic acids, which is of course essential in the functioning and multiplication of all cells. As one would expect, orotic

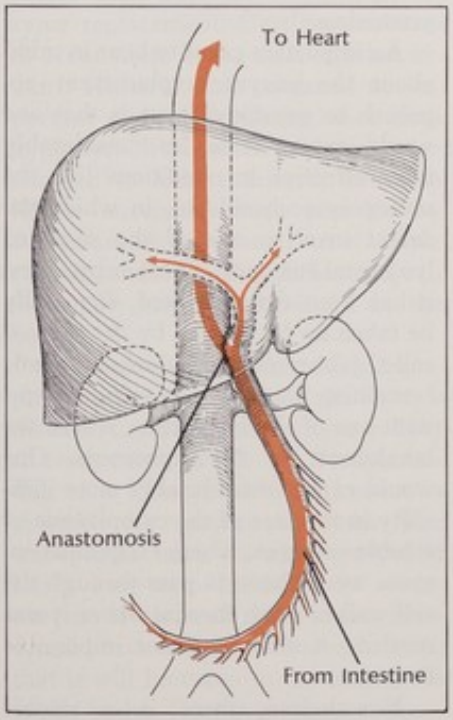
aciduria responds beautifully to replacement therapy, in which the diet is enriched by adding either uridine itself or RNA, which is a mixture of nucleotides including uridine.

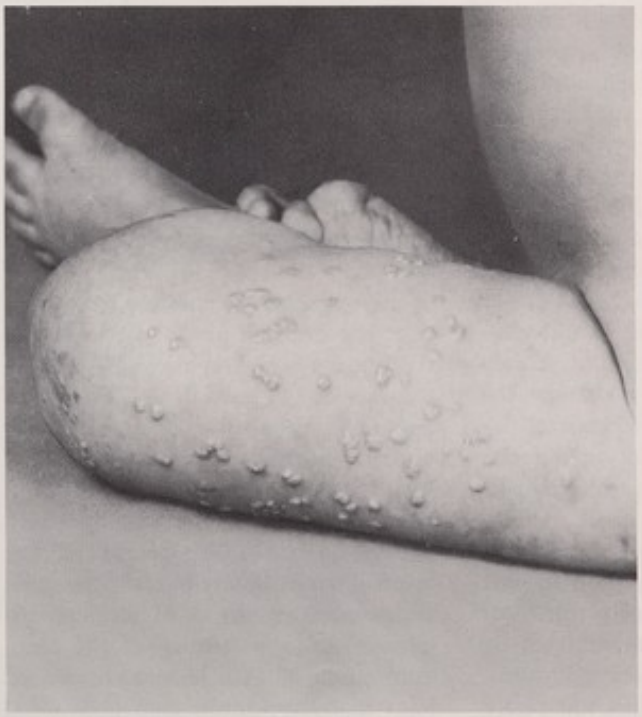
A quite different approach has been employed experimentally in two of the glycogen storage diseases, types I and III. Here the enzymatic defect prevents the breakdown of glycogen, stored in the liver, into glucose, for utilization by the tissues. Though the glycogen buildup in the liver produces hepatomegaly, the chief pathologic problem appears to be chronic hypoglycemia. Any direct replacement therapy would obviously pose insuperable practical problems, since it would involve subjecting the patient to a more or less continuous glucose drip into a part of the circulation that would carry the sugar to the tissues directly without passing through the liver where it would be "inactivated" as glycogen. Several clinicians have therefore attempted to obtain a comparable result by surgical means. In one technique, the positions of the portal vein and inferior vena cava are interchanged. Since the portal vein drains the gut, it will now carry blood rich in glucose (from food) directly into the heart, whence the blood will be pumped out to nourish the peripheral tissues. The inferior vena cava, which drains the lower limbs, is attached to the stump of the portal vein, thereby supplying the liver with blood – but blood from which most of the glucose has already been extracted. An alternative technique involves creating a portacaval shunt (usually by side-to-side anastomosis) that allows some glucose-rich blood from the portal vein to flow into the inferior vena cava and thence directly to the heart.

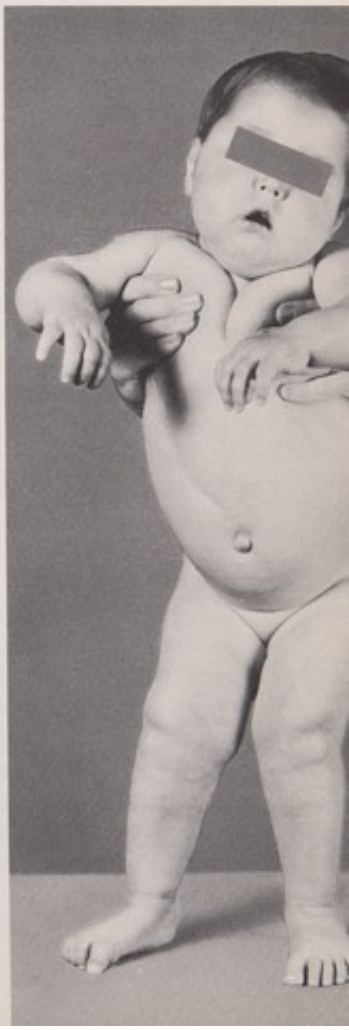
The limited experience with the surgical treatment of glycogen storage disease is distinctly encouraging. Several patients have shown marked improvement in growth, the result of the enhanced glucose supply in the peripheral tissues. In no case has the hepatomegaly been reversed, but the organ's enlargement has been checked so that in effect the patients seem to be "growing up" to the size of their livers. However, we should also note that a minority of patients have died in the immediate postoperative period from causes still poorly understood. Doubtless this problem will be solved in time, but meanwhile surgery can obviously not be considered a standard treatment – certainly not in the case of children in whom the condition is not severe enough to produce serious functional impairment.

Finally, we should cite a special category in the area of biochemical correction, the one occupied by homocystinuria. Here, the disease appears to stem *both* from the accumulation of homocystine and from the deficiency of cystine, into which it would normally be converted. At any rate, patients appear to respond better to a diet in which low levels of methionine (a metabolic precursor of homocystine) are combined with supplemental quantities of cystine.

When we consider the treatment of genetic disease on the enzymatic level, we find again that the approaches fall into two categories, depending on the nature of the enzyme defect. In the case of what may be called the binding defects, the structural abnormality in the enzyme impairs its ability to unite with the coenzyme – usually a vitamin – whose presence is required for proper enzymatic functioning. In the case of the more common catalytic defects, on the other hand, the enzyme can bind normally to its coenzyme –







Portacaval shunt (left) has been used for experimental treatment of certain of the glycogen storage diseases. Anastomosis between portal vein and inferior vena cava permits some glucose-rich blood from gut to bypass liver, where much glucose would be stored as glycogen, and flow directly to heart. Treatment can produce clear signs of improvement, as illustrated by disappearance of skin lesions in child operated on at Cleveland Clinic (photos courtesy of Dr. Robert E. Hermann, Surgery *65:3:501, 1969).*

if one is involved in the reaction – but remains inadequate to its catalytic metabolic function. A possible third category may be represented by conditions in which the problem is thought to be not with the enzyme as such but rather with regulatory substances governing its production; that is, the enzyme itself is structurally and functionally normal but is produced in subnormal quantities.

The "binding defect" conditions are, of course, the vitamin-dependent genetic diseases that have already been discussed in Chapter Seven by Leon E. Rosenberg. To summarize the matter briefly, it has been found that in many of these disorders, including such conditions as cystathioninuria and certain types of homocystinuria, the binding defect is not absolute. That is, the enzyme retains some binding capacity, which, however, is inadequate at normal serum levels of the vitamin coenzyme in question. Accordingly, the patient is given relatively enormous doses of the vitamin – e.g., up to 100 mg pyridoxine per day in cystathioninuria, on the order of 100 times the normal requirement. These can saturate the defective enzyme with vitamin even in the face of its impaired binding capacity; no sooner does one vitamin molecule "fall away" from the enzyme than another is there to replace it. Again, preliminary clinical reports are encouraging. Currently a number of clinicians are experimenting with megavitamin therapy in schizophrenia, a multifactorial disease now known to have a strong genetic component, though its biochemistry remains obscure. Although convincing data are not available, any success along these lines would classify at least certain forms of schizophrenia presumptively among the vitamin-dependent conditions.

In the case of the catalytic defects, enzymatic treatment requires replacement of the enzyme itself. Oddly enough, the outstanding example of this subtle approach is by now so well established that one hardly thinks of it in those terms. It can, indeed, be argued whether the missing clotting factors in the hemophilias are really enzymes, but they are incontestably proteins with quasi-enzymatic functions. And, of course, these diseases have long been treated by supplying the missing factors, originally by transfusions of whole blood from normal persons and currently by administering appropriate blood fractions.

A more recent development in enzyme replacement involves the mucopolysaccharidoses, a group of diseases discussed in Chapter 14 by Neufeld. Readers will recall from this chapter that this group of related conditions involves the accumulation of mucopolysaccharides in the body cells because of failure to degrade them into a form that the cells can excrete; the degradation process appears to involve the sequential action of at least six different enzymes, with the precise nature of the disease depending on which enzyme is missing.

The diseases are now being treated experimentally by infusions of plasma from normal individuals, but the results are hard to interpret. There seems no doubt that the infusions temporarily increase the body's excretion of low-molecular weight mucopolysaccharide breakdown products but the clinical implications of this fact are uncertain. On the one hand, a number of thoughtful and well-regarded investigators feel that some clinical improvement has been demonstrated. On the other hand, some equally well-regarded researchers have pointed out that the amount of enzyme presumably being administered, as estimated from in vitro tests of enzyme activity in plasma, should be only enough to "normalize" a small fraction of the body's cells. A good deal of further work will be needed before this contradiction can be resolved.

Opinions are possibly less divergent regarding a newly tried alternative approach to enzyme replacement – transfusion of leukocytes. These are generally obtained from a clinically normal relative, to minimize immunologic incompatibilities. Transfusion has been followed not only by sharp increases in urinary levels of the breakdown products but also by such rather striking signs of clinical improvement as softening of the skin, diminution of skin lesions, increased joint mobility, diminished abdominal protrusion, and decreased hepatomegaly. Moreover, though urinary excretion of MPS breakdown products rapidly returns to its original low level, the clinical improvement persists for several months after the infusion – presumably because the cells, once purged of their accumulated MPS, require considerable time to reaccumulate it to pathologic levels.

A key question concerning enzyme replacement in mucopolysaccharidoses is whether it can also alleviate the central nervous system symptoms that are a prominent feature in several types of the disease – a matter in which accurate clinical evaluation is obviously a good deal more difficult. One would not expect CNS improvement with any form of plasma infusions, since the enzymes would presumably be blocked by the blood-brain barrier. And as regards leukocyte infusion, while there is some reason to believe that these cells *may* on occasion pass through the vascular walls into the brain, there is so far as I know no evidence that they actually do so on any significant scale. If, indeed, the treatment can be shown to produce unequivocal behavioral changes, one will have to assume that the leukocytes are in fact getting to the CNS, but demonstration of such results will undoubtedly require a number of careful double-blind studies. Meanwhile, however, the prospects for effective treatment of some of the mucopolysaccharidoses must be classed as at least promising.

An important point to bear in mind about the enzyme replacement approach to genetic disease is that one would expect it to be considerably more effective in conditions like the mucopolysaccharidoses, in which the defect involves one of the so-called lysosomal enzymes. These substances, it has been demonstrated, can easily be taken up by the cell by the process called pinocytosis, in which a vacuole containing exogenous enzyme merges with one of the lysosomes, where the breakdown of MPS proceeds. One would expect considerably more difficulty in the case of the cytoplasmic or soluble enzymes, where exogenous enzyme would have to pass through the cell wall to reach the site where it was needed – a difficult if not impossible feat.

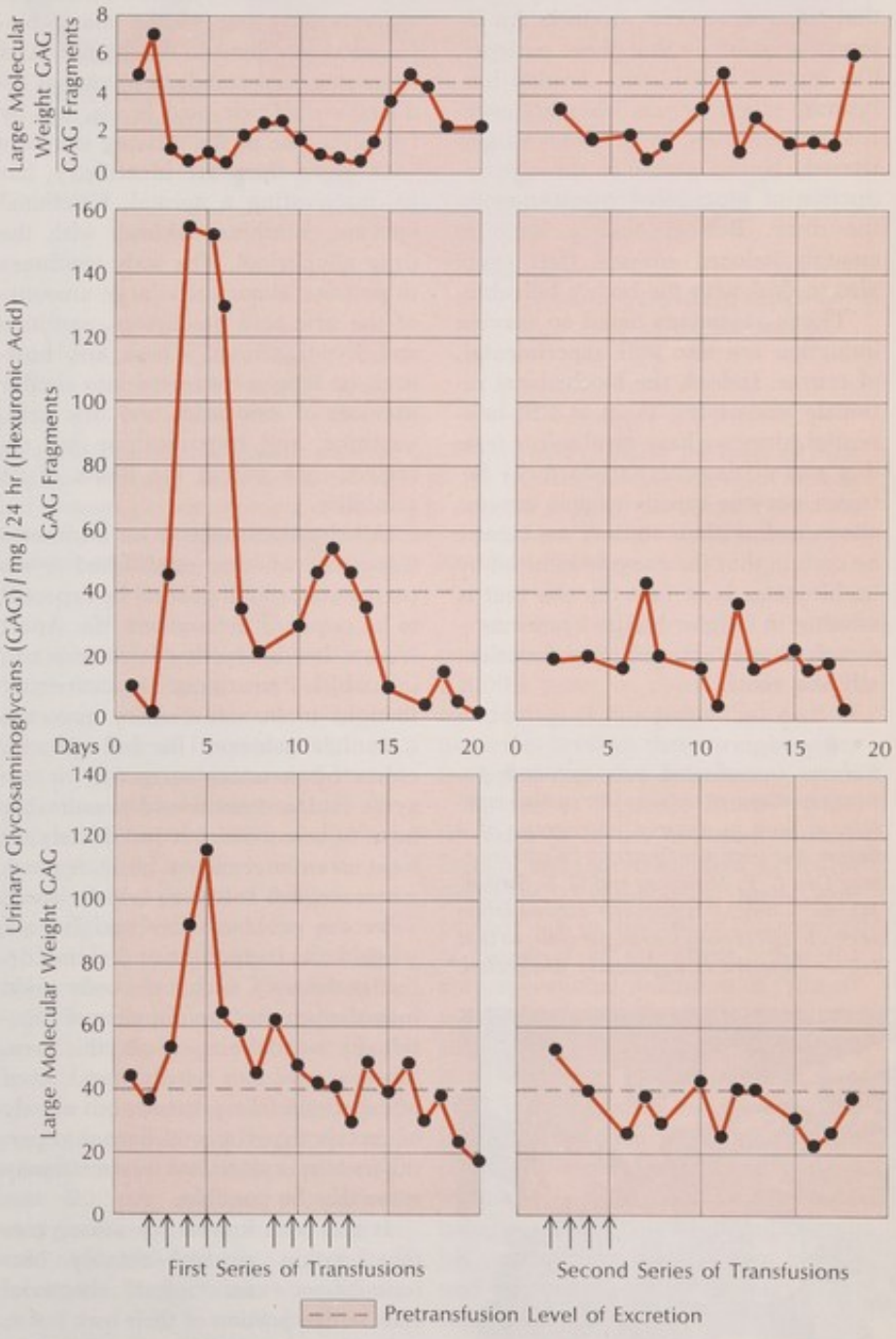
Nonetheless, there is a sizable number of other lysosomal diseases, such as Tay-Sachs, Gaucher's, and Fabry's, which should, in principle, be amenable to treatment with exogenous enzyme – provided, of course, that the

enzyme in question can be both identified and obtained in clinically useful quantities. A specially suitable candidate for research, perhaps, is cystinosis, in which – unlike most of the other lysosomal diseases – there is no CNS involvement, so that the benefits of treatment should be unequivocal. The problem here, however, is that we have no present reason to believe that normal cells (e.g., leukocytes) excrete enzymes that have any effect in cystinosis, as they evidently do in mucopolysaccharidosis.

A currently more practicable approach to the treatment of cystinosis is by kidney transplantation. In this disease the central clinical problem is kidney failure, which occurs because of accumulation of endogenous cystine within the cells of the tubules; transplantation, in effect, substitutes enzymatically normal cells for the abnormal ones. The transplanted kidney does, indeed, continue to accumulate cystine, probably through the taking-up of cystine-laden leukocytes by macrophages in the organ, but the accumulation is in the interstitial spaces rather than in the cells so that tubular function is apparently not affected. I should emphasize, however, that what is being done here is local, rather than general or systemic, enzyme replacement. Cystine continues to accumulate in the cells of other organs. Hitherto, this produced no important clinical problems other than hypothyroidism before the patient succumbed to kidney failure, but with longer survivals made possible by transplantation we must, out of simple caution, anticipate that difficulties will eventually develop elsewhere in the body.

Another condition in which transplantation appears useful is Fabry's disease. Certain workers have sought to treat this disorder by transplantation of a kidney, which is known to produce the enzyme involved. The transplants have certainly produced fairly long-term increases in serum enzyme levels, and since it is a lysosomal enzyme there is reason to hope that it will be taken up by other cells in the body. That this may be occurring at least to some extent is suggested by reports that some patients have had significant reductions in pain. But it should be stressed that these reports are preliminary, and the long-term efficacy of kidney transplantation has not yet been demonstrated.

Finally, I should mention at least briefly two diseases in which the enzyme defect is suspected of being regulatory rather than structural – i.e., the cells are presumed to possess the genetic information necessary to make a normal enzyme, but their production of it is somehow blocked. Thus a rational therapeutic approach would be to artificially induce enzyme production. In orotic aciduria it has been shown, in vitro, that this can be done by treating cultured fibroblasts from a patient with 6-azauridine. However, there has as yet been no attempt to apply this finding clinically. Another possible example of enzyme induction

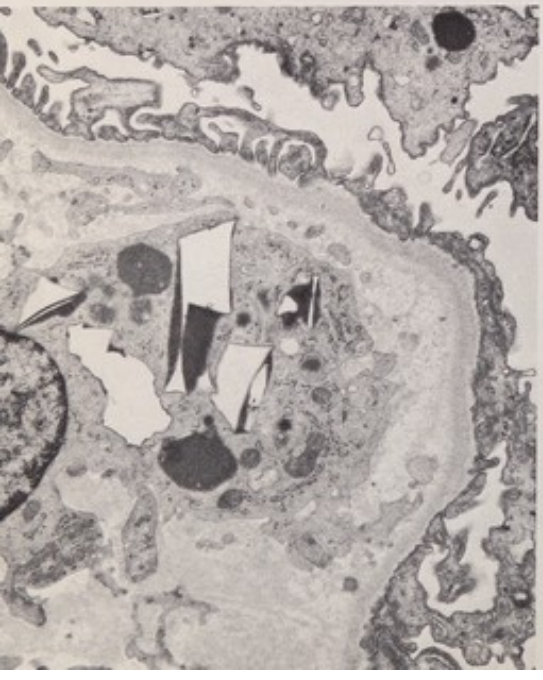


Transfusion of normal leukocytes into patient with type II mucopolysaccharidosis (Hunter's syndrome) produced a sharp increase in excretion of mucopolysaccharide breakdown products, especially those of low molecular weight, Knudson and coworkers in Houston showed; clinical improvement was also clearly apparent. The increase was far less marked following a second course of treatment, presumably because little mucopolysaccharide had reaccumulated in the patient's cells.

is in Crigler-Najjar syndrome. Here, the problem is the absence or inactivity of the enzyme glucuronyl transferase, which detoxifies serum bilirubin by conjugating it with glucuronic acid. Some Crigler-Najjar patients appear to have responded to treatment with barbiturates, whose capacity to induce enzymes is well known – one might even say notorious, since enzyme induction by barbiturates is the basis of the very dangerous tolerance that these drugs can produce. An interesting point is that these patients, like normal individuals, develop barbiturate tolerance, i.e., they are eventually not sedated by the doses administered, almost certainly through induction of glucuronyl transferase by the drug. But apparently there is enough induced enzyme "left over" also to deal with the body's bilirubin.

These treatments based on enzyme induction are also still experimental, of course. Indeed, the biochemical rationale underlying them is still inferential, since we have no absolute tests that can distinguish satisfactorily between enzyme inactivity and enzyme absence. For that matter, we cannot be certain that the enzyme induced by barbiturates is in fact the one that is missing in Crigler-Najjar syndrome – merely that it appears to be functionally the same.

Kidneys transplanted into cystinosis patients continue to accumulate cystine crystals, as seen in electron micrograph of a biopsy one year after grafting (photo courtesy Drs. C. P. Mahoney and G. E. Striker, Mason Clinic). However, accumulation seems to affect only mesangial cells, so that tubule function is apparently unaffected.



Before closing this section, I should cite two genetic diseases in which treatment is of "mixed" type – both enzymatic and biochemical. That is, the therapy involves attacking the biochemical abnormality by enzymatic means, but not through the enzymes whose inactivity is fundamentally responsible for the disease. The exocrine enzyme trypsin has long been administered in cystic fibrosis, and is generally deemed beneficial, though it corrects only one of the many biochemical problems in the disease. And in gout, a now standard treatment reduces serum uric acid levels, not by replacing one of the missing enzymes (not all of them are identified), but by inactivating a normal, functional enzyme, xanthine oxidase, with the drug allopurinol. The body continues to produce abnormally large amounts of the uric acid precursors, xanthine and hypoxanthine. These are, however, no longer converted into similar excesses of uric acid, and uric acid, xanthine, and hypoxanthine are excreted, each within the limits of its solubility.

All the treatments I have hitherto discussed, whether established or experimental, can in general be expected to be required throughout life. Apart from a few exceptions such as PKU, in which treatment is currently thought to be unnecessary once the CNS has achieved its full growth, either biochemical correction or enzyme replacement would presumably have to be carried out indefinitely, at least on an intermittent basis. A much more elegant solution, which would solve the problem once and for all, would be to correct the underlying defect in the DNA so that the body could manufacture its own supplies of functionally normal enzyme. In this area, however, we can hardly speak even of experimental treatments but merely of certain experimental data that permit us to speculate that treatment may some day be possible.

It has been known for some years that certain viruses – notably bacteriophages – can "infect" bacterial cells with a portion of their own DNA; more recently, similar DNA transduction has been observed in mammalian cells. In 1971, a group at the National Institutes of Health reported that it had infected fibroblasts from a galactosemia patient with a bacteriophage and thereby had normalized the cells' enzymatic activity. Specifically, the cells could now produce normal α-D-galactose-1-phosphate uridyl transferase, the lack of which causes classic galactosemia. As a check, the researchers also infected cells with a mutant phage that lacked this enzyme activity and found, as expected, that the cells were unchanged.

A second suggestive and most preliminary finding concerns the Shope papilloma virus. It has been known for some years that some individuals working with this virus have chronically low blood levels of the amino acid arginine, and since the virus itself is known to be rich in activity of the enzyme arginase, it is thought that they have acquired a chronic but apparently benign infection with the virus. Subsequently, a few children with genetic argininemia have been infected with the virus, and in one case the blood arginine level did fall and at last reports was still low.

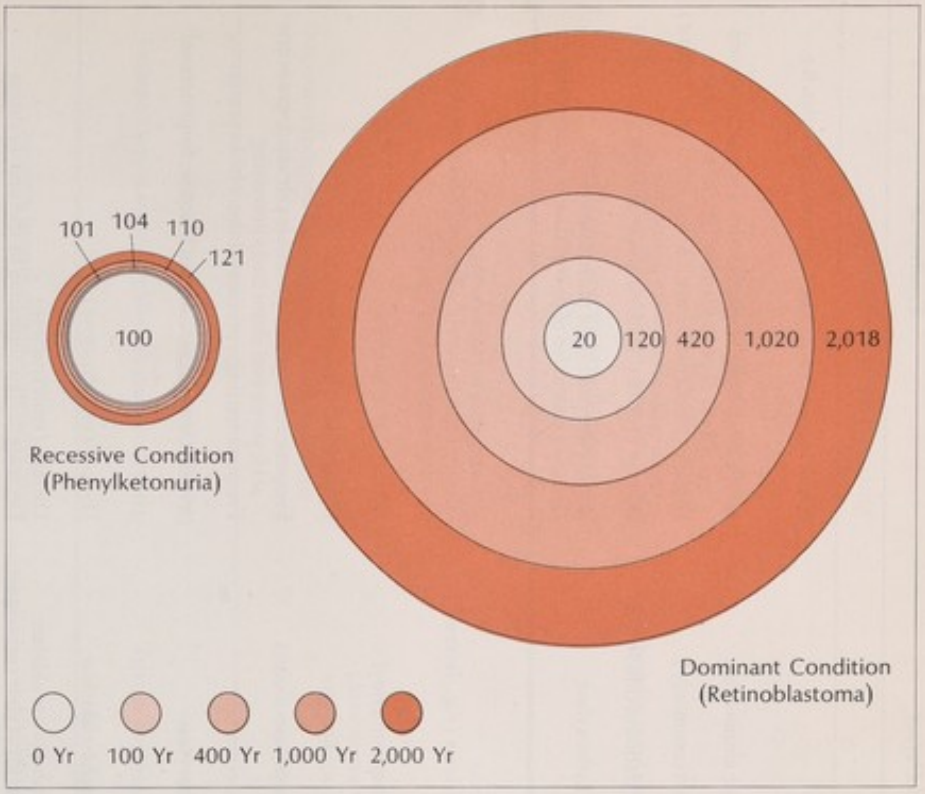
The major problem with using phages or other viruses to infect a patient with a particular, needed type of genetic information is that he would simultaneously be infected with all the rest of the organism's genetic information – which could be a great deal more than was good for him. There is obviously no guarantee that such infection will be benign even in the short run, as seems to be the case with the Shope virus, and even less certainty that introduction of a strange virus will not be oncogenic in the long run. Thus the prospect of safely treating genetic disease by introducing exogenous DNA must, I think, be regarded as currently very tenuous, though it may ultimately prove feasible in a few special cases.

A potentially safer approach might be to use the patient's own cells. These could be cultured and treated with a variety of mutagens in the hope that at least one cell would mutate in such a way as to correct the DNA defect. The odds against such a mutation would be very high, but perhaps not prohibitively so. In a considerable number of genetic diseases, it is possible, by manipulating the culture medium, to select out the one cell in ten thousand – or even ten million – that does, or does not, possess a particular enzyme; in fact, such selective

mechanisms are the basis of some standard tests for genetic disease. Thus in galactosemia, for example, one would grow mutated cells from a galactosemia patient in a medium containing galactose as the only energy source; any cell that survived would evidently be able to metabolize galactose successfully.

Assuming a "favorable" mutation could be obtained, the next step would be to find ways of colonizing the patient with a population of these cells, perhaps by transplanting leukocytes into the bone marrow. It is worth noting that such transplants have proved effective in some types of immune deficiency, even though the leukocytes are exogenous; transplantation with endogenous, though mutated, leukocytes should, from an immunologic standpoint, be even simpler.

Summing up the current situation, then, it would appear that progress in treating genetic disease can be expected to continue, but at a measured pace. The problem will almost certainly have to be solved disease by disease; the likelihood of any dramatic overall "breakthrough" seems to me very remote.



Successful treatment of genetic diseases would necessarily lead to increases in their incidence. However, these would be minor for the recessive conditions, which include most common genetic disorders, amounting to only 21% in 2,000 years (numbers on the circles show calculated incidence per million population). The calculated increase for dominant conditions would be much larger: 100-fold in 2,000 years.

The expectation that increasing numbers of individuals suffering from hitherto intractable genetic conditions can be enabled to survive and lead more or less normal lives is, from one standpoint, a very welcome development. But it carries with it certain broader implications, some of which are considerably less welcome.

A problem not infrequently raised in connection with the treatment of genetic disease is that successful treatment, by permitting affected individuals to survive and reproduce, will negate the natural process that regularly eliminates a proportion of the defective genes from the population. The result, it is held, will be a rapid increase in the population frequency of the gene in question and an even more rapid increase in the number of affected individuals who must be treated. The effect will be a drop in the "genetic fitness" of the population and – to the extent that the treatment can be expected to be lifelong – a sharp rise in the load on medical facilities and personnel.

The basic premise of this argument is obviously correct: Defective genes not eliminated by the death or reproductive incapacity of the individual possessing them will remain to increase the overall frequency of these genes in the population, and therefore, other things being equal, the number of affected individuals. What is not generally realized, however, is the extremely slow rate of this increase for most genetic diseases. For genetically dominant conditions, the increase would indeed be rapid. For example, in dominantly inherited retinoblastoma it can be calculated that if all affected individuals could be cured (not an imminent possibility) and were to reproduce at a normal rate, the incidence of the disease, currently about 20 per million population, would increase sixfold in five generations and 100-fold in a hundred generations. But dominant conditions account for only a small fraction of the really incapacitating genetic diseases. In the case of recessive conditions, the increase in incidence is far slower: in PKU, for example, under the same assumptions as above, the rise would be only 1% in five generations and only 21% in 100 generations – 2,000 years or more. It is worth mentioning at this point that each of us carries between three to eight genes for lethal recessive disorders.

So far as "genetic fitness" goes, the term is to me essentially meaningless, and indeed involves a circular definition. Evolutionists have long pointed out, in discussing the concept of "survival of the fittest," that the only operational definition of "fittest" is "those who survive." Thus if a person with a genetic disease survives, he is by definition fit. Beyond this, survival in our species has always been in part a function of technology, including in recent generations medical technology. Strip man of his technology – of his tools for manipulating his environment and his own body – and not one of us could survive for more than a few weeks. From the moment our ape-like ancestors began using tools, they set up a genetically selective process that facilitated the survival of individuals who were "unfit" to survive in a nontechnologic state of nature. And yet I do not know of

Genetic Diseases That Are Currently Treatable

Disorders of	*Main Signs and Symptoms*	*Biochemical Defect*	*Treatment*
AMINO ACID METABOLISM			
CYSTINOSIS (SEVERAL TYPES)	Widespread tissue deposits of cystine, producing Fanconi's syndrome and renal failure; crystal deposits in cornea	Unclear	Alkali to correct acidosis; renal transplantation for renal failure
CYSTINURIA (SEVERAL TYPES)	Recurrent renal stones (cystine)	Defect in reabsorption of basic amino acids	Increase solubility of cystine with large urine volume and alkali; D-penicillamine for resistant cases
HOMOCYSTINURIA (SEVERAL TYPES)	Ectopia lentis, osteoporosis, dolichostenomelia, pes cavus, retardation, arterial and venous thrombosis	Cystathionine synthase (most commonly)	Dietary (restriction of methionine, addition of cystine); certain cases respond to pyridoxine
MAPLE SYRUP URINE DISEASE	Vomiting, hypertonicity, mental retardation, characteristic odor of urine, death	Defect in oxidative decarboxylation of 3-branched chain amino acids	Dietary (restriction of leucine, isoleucine, and valine)
PHENYLKETONURIA	Retardation, seizures, hyperactivity	Deficiency of phenylalanine hydroxylase	Dietary (restriction of phenylalanine)
TYROSINEMIA	Cirrhosis, renal tubular defects	Unknown	Dietary (restriction of phenylalanine and tyrosine)
CARBOHYDRATE METABOLISM			
DISACCHARIDASE DEFICIENCIES	Abdominal fullness, bloating, cramping, diarrhea	Various intestinal disaccharidases (e.g., lactase, sucrase)	Dietary (elimination of offending disaccharide)
GALACTOSEMIAS			
TYPE I	Failure to thrive, hypoglycemia, jaundice, hepatomegaly, cataracts, Fanconi's syndrome	Deficiency of galactose-1-phosphate uridyl transferase	Dietary (omit galactose)
TYPE II	Cataracts	Deficiency of galactokinase	Dietary (omit galactose)
GLYCOGEN STORAGE			
TYPE I (VON GIERKE'S DISEASE)	Massive hepatomegaly, hypoglycemia, acidosis, hyperlipidemia, hyperuricemia	Deficiency of glucose-6-phosphatase	Frequent feedings; allopurinol for elevated serum uric acid; portacaval shunt looks promising
TYPE III (DEBRANCHING DEFICIENCY)	Hepatomegaly, hypoglycemia, ketosis	Deficiency of debranching enzyme	Frequent feedings; portacaval shunt looks promising
TYPE V (MCARDLE'S DISEASE)	Painful muscle cramps	Deficiency of muscle phosphorylase	Avoid excessive exercise; oral fructose; isoproterenol
HEREDITARY FRUCTOSE INTOLERANCES			
TYPE I	Failure to thrive, jaundice, hepatomegaly and cirrhosis, seizures, rickets	Deficiency of fructose-1-phosphate aldolase	Dietary (omit fructose)
TYPE II	Hypoglycemia, hepatomegaly	Deficiency of fructose-1,6-diphosphate aldolase	Dietary (omit fructose)
PRIMARY HYPEROXALURIA	Early recurrent renal calculi (oxalate), renal failure	Deficiency of α-ketoglutarate:glyoxylate carboligase; less commonly D-glycerate dehydrogenase	Decrease oxalate formation by diet and/or drugs (pyridoxine); reduce stone formation by high oral phosphate intake
THE UREA CYCLE			
ARGINOSUCCINICACIDURIA	Mental retardation, seizures, intermittent ataxia, hepatomegaly, and trichorrhexis nodosa	Deficiency of arginosuccinase	Dietary (protein restriction)
CITRULLINEMIA	Vomiting, lethargy, irritability, coma, mental retardation	Arginosuccinic acid synthetase deficiency	Dietary (protein restriction)
HYPERAMMONEMIA			
TYPE I	Vomiting, lethargy, irritability, coma, mental retardation	Carbamyl phosphate synthetase deficiency	Dietary (protein restriction)
TYPE II	Vomiting, lethargy, irritability, coma, mental retardation	Ornithine carbamyl transferase deficiency	Dietary (protein restriction)
HYPERARGININEMIA	Vomiting, lethargy, irritability, coma, mental retardation	Arginase deficiency	Dietary (protein restriction); Shope virus infection has been tried
PURINE AND PYRIMIDINE METABOLISM			
GOUT	Acute arthritis, sometimes tophi	Usually not known; sometimes partial deficiency of hypoxanthine-guanine phosphoribosyl transferase	Probenecid; allopurinol
LESCH-NYHAN SYNDROME	Cerebral palsy, self-destructive behavior, gouty arthritis	Deficiency of hypoxanthine-guanine phosphoribosyl transferase	Allopurinol (doesn't affect neurologic manifestation)
OROTIC ACIDURIA	Megaloblastic anemia, crystalluria, failure to thrive	Deficiency of orotidylic pyrophosphorylase and orotidylic decarboxylase	Oral uridine

IMMUNE DEFICIENCIES AND DISORDERS OF CIRCULATING PROTEINS			
AGAMMAGLOBULINEMIA ACCOMPANIED BY LYMPHOPENIA, ABSENCE OF CELLULAR IMMUNITY	Bacterial, fungal, and viral infection	Failure of normal thymic embryogenesis	Transplant of histocompatible bone marrow
CHOLINESTERASE DEFICIENCY	Prolonged apnea when given muscle relaxant (suxamethonium)	Deficiency or abnormality of cholinesterase	Avoid muscle relaxants
GENETIC DEFECTS IN GAMMA GLOBULIN SYNTHESIS	Recurrent pyogenic infections	Defect in gamma globulin synthesis	Gamma globulin administration
HEREDITARY DISORDERS OF HEMOSTASIS	Excess bleeding	Deficiencies of many factors, such as factor VIII (classic hemophilia), IX (Christmas disease)	Replacement of deficient factor either with blood or as concentrates
ENDOCRINE GLANDS			
ADRENOGENITAL SYNDROME (SEVERAL CATEGORIES)	In various types: virilization, salt loss, hypertension	In various types: 21-hydroxylase deficiency 11-hydroxylase deficiency 3-β-hydroxysteroid dehydrogenase deficiency 17-hydroxylase deficiency	Cortisol or its analogues
FAMILIAL GOITER	Retarded growth, development	Defects in synthesis of or response to thyroid hormone	Thyroid hormone (not all genetic types respond)
PITUITARY DWARFISM (SPECIFIC TYPES)	Short stature	Deficiency of human growth hormone	Administration of human growth hormone
LIPID STORAGE			
FABRY'S DISEASE	Fever; severe pain in extremities; vascular lesions of skin, conjunctiva, oral mucosa; renal dysfunction	Deficiency of ceramide trihexosidase activity	Renal transplant (to supply enzyme to other tissues)
REFSUM'S DISEASE	Retinitis pigmentosa, peripheral polyneuropathy, cerebellar ataxia, nerve deafness	Defective oxidation of phytanic acid	Dietary (restriction of phytanic acid)
HEMATOLOGY			
HEMOLYTIC DISEASE OF NEWBORN	Jaundice, anemia	Blood group incompatibility between mother and fetus	Exchange transfusion; phototherapy; when due to Rh incompatibility, administer Rh antibody to mother
HEREDITARY HEMOLYTIC ANEMIAS (GROUP OF SIMILAR DISORDERS) DUE TO ENZYME DEFICIENCIES	Hemolytic anemia (without spherocytes)	Deficiency of a variety of red cell enzymes (such as pyruvate kinase, hexokinase)	Improvement after splenectomy
HEREDITARY METHEMOGLOBINEMIA	Methemoglobinemia (must be differentiated from cyanosis)	Methemoglobin reductase	Methylene blue or ascorbic acid
HYPERBILIRUBINEMIAS AND CRIGLER-NAJJAR SYNDROME	High indirect bilirubin; in infancy, severe brain damage	Defective bilirubin conjugation	Some respond to phenobarbital
MISCELLANEOUS			
CYSTIC FIBROSIS	Chronic bronchiolar obstruction with infection, steatorrhea	Unclear	Antibiotics; pancreatic exocrine enzymes
FAMILIAL HYPOPHOSPHATEMIC RICKETS	Rickets	Renal tubular defect in reabsorption of phosphorus	Phosphate salts; high doses of vitamin D
HEMOCHROMATOSIS	Widespread accumulation of iron with impairment of heart, pancreas (diabetes); cirrhosis of liver	Increased absorption of iron	Regular venesection
RETINOBLASTOMA	Tumor of the eye, with invasion	Unknown	Radiation and chemotherapy
WILSON'S DISEASE	Cirrhosis of liver, degenerative changes in brain (particularly basal ganglia), Kayser-Fleischer corneal rings are pathognomonic	Unclear	D-penicillamine
*MENKE'S DISEASE (KINKY-HAIR SYNDROME)	Progressive brain damage and death, hypothermia, pili torti, subdural hematomas	Defect in copper absorption	Parenteral copper may be useful
*MUCOPOLYSACCHARIDOSES TYPE I (HURLER'S SYNDROME) TYPE II (HUNTER'S SYNDROME)	Skeletal abnormalities, growth and mental retardation, coarse facies (corneal clouding in Type I)	Lysosomal defects (various) in mucopolysaccharide degradation (1-iduronidase deficiency in Type I)	Plasma infusion and leukocyte infusions reported to be of benefit

* *Currently available therapy may be useful*

anyone who regrets this development.

Nonetheless, quite apart from the matter of long-term effects on population genetics, the prospects for increasingly successful treatment of genetic disease do raise serious questions of public policy. These stem, first, from the fact that the treatment of most genetic diseases can be expected, at least for the near future, to require life-long therapy and, second, from the fact that even in our own affluent country medical resources are not infinite and never will be. One is then forced to ask – to put the matter in its crudest form – whether we are justified in employing medical equipment and personnel to treat a handful of sufferers from rare genetic anomalies when the same resources, otherwise employed, could deal with the simpler, but equally incapacitating, health problems of ten or a hundred times as many people.

In the last analysis, this is a moral and ethical problem and, as with most such problems, the answers are not simple – in fact, there is a whole range of answers depending on the disease under discussion. At one extreme we have conditions such as PKU. Here, the untreated sufferer can be expected to survive for an average of something like 50 years, but with functional impairment severe enough to require institutionalization in almost all cases. And as others have pointed out, if we weigh the cost of institutionalization against the cost of detection and treatment, the latter, simply as a matter of economics, makes far more sense. This probably remains true even if we take into account the growing problem of prenatal PKU, which it now appears can be expected to threaten a large majority of the infants born to phenylketonuric mothers, although it may be controlled by reinstituting the mother's low-phenylalanine diet during pregnancy (if one were certain of identifying all such mothers).

At the other extreme, we have such conditions as maple syrup urine disease, in which the untreated sufferer will frequently survive no more than a few years, while treatment is difficult, expensive, and certainly not completely effective. Here, treatment is unjustified on economic grounds alone. Considering what the same medical resources could accomplish in the way of human betterment if applied to the treatment of other, less intractable health problems, we must also wonder if it makes ethical sense.

Without attempting to answer the question categorically, I would like to put forward two important considerations that seem to me relevant in evaluating any program for treating rare conditions of this sort. The first is that the program should – if not must – include a research as well as a therapeutic component. That is, the program should reflect, in both its structure and its specific goals, a concern not only with the essential aim of treating the disease but also with studying it intensively, to the end of achieving a better understanding of its mechanisms – and, it is to be hoped, those of genetic disease generally. On the practical side, the research should aim at devising treatments that are not only maximally effective but also make minimal demands on medical resources.

The second consideration, I believe, concerns the responsibility of physicians involved in such programs, whether as researchers or as clinicians, to ensure, insofar as lies in their power, the birth of unaffected children to families at risk for genetic disease. This means that genetic counseling is a cornerstone of current genetic therapy, by helping families prevent the recurrence of serious genetic disease. We must clearly point out to such couples that there are several alternatives to their having another affected child who must be treated, particularly in view of the less-than-perfect results achieved even with the most "treatable" genetic conditions (e.g., galactosemia, PKU). They can instead decide to have no more children, or to have them by artificial insemination, or to adopt them. Or – if the condition is one of the increasing number of genetic diseases that can be diagnosed prenatally – they can proceed to conceive in the normal way but on the understanding that should the fetus be shown to be affected the pregnancy can be terminated. (In this connection, it should be obvious that physicians concerned with genetic disease have a very legitimate concern with ensuring that state laws do not block pregnancy termination under these circumstances.)

Thus, for all the recent and prospective progress in treating genetic disease, I believe that we must still consider treatment a measure of last resort, to be obviated if at all possible. This approach, I believe, can to some extent free us from confronting the ambiguous and thorny ethical problems cited above, since we will be operating in accordance with the well-tried maxim that prevention is better than cure.

Selected References and Index

Selected References

Chapter 1

Hamerton JL: Human Cytogenetics, Vol I and II. Academic Press, Inc., New York, 1971

Hirschhorn K: Chromosome identification. Annu Rev Med (in press)

Nusbacher J, Hirschhorn K: Autosomal anomalies in man. Advances in Teratology 3:11, 1968

Redding A, Hirschhorn K: Guide to human chromosome defects. Birth Defects: Original Articles Series Vol IV, No 4, September 1968. The National Foundation – March of Dimes

Conference on Standardization in Human Cytogenetics, Chicago, 1966. Birth Defects: Original Articles Series Vol II, No 2, December 1966. The National Foundation – March of Dimes

Conference on Standardization in Human Cytogenetics, Paris, 1971. Birth Defects: Original Articles Series Vol VIII, No 7, October 1972. The National Foundation – March of Dimes

Chapter 2

Brown WM, Price WH, Jacobs PA: Further information on the identity of 47,XYY males. Br Med J 2:325, 1968

Ferguson-Smith MA: Karyotype-phenotype correlations in gonadal dysgenesis and their bearing on the pathogenesis of malformations. J Med Genet 2:142, 1965

Ferguson-Smith MA: Sex chromatin, Klinefelter's syndrome and mental deficiency. The Sex Chromatin. Moore KL, Ed. W. B. Saunders Company, Philadelphia, 1966

Ferguson-Smith MA: X-Y chromosomal interchange in the aetiology of true hermaphroditism and of XX Klinefelter's syndrome. Lancet 2:475, 1966

Ferguson-Smith MA: Testis and intersexuality. Paediatric Endocrinology. Hubble D, Ed. Blackwell, Oxford, 1969, p 359

Gartler SM, Waxman SH, Giblett E: An XX/XY human hermaphrodite resulting from double fertilization. Proc Natl Acad Sci USA 48:332, 1962

Hamerton JL: Human Cytogenetics, Vol I and II. Academic Press, Inc., New York, 1971

Lyon MF: Sex chromatin and gene action in the mammalian X-chromosome. Am J Hum Genet 14:135, 1962

Pearson PL, Bobrow M, Vosa CG: Technique for identifying Y-chromosomes in human interphase nuclei. Nature (Lond) 226:78, 1970

Race RR, Sanger R: Blood groups in twinning, chimaerism and dispermy. Blood Groups in Man, 5th Ed. Blackwell, Oxford, and F. A. Davis Co., Philadelphia, 1968

Ratcliffe SG, Stewart AL, Melville MM, Jacobs PA, Keay AJ: Chromosome studies on 3,500 newborn male infants. Lancet 1:121, 1970

Sanger R, Tippett P, Gavin J: Xg groups and sex abnormalities in people of northern European ancestry. J Oral Surg 8:417, 1971

Conference on Standardization in Human Cytogenetics, Paris, 1971. Birth Defects: Original Articles Series, Vol VIII, No 7, October 1972. The National Foundation – March of Dimes

Chapter 3

Caspersson T, De la Chapelle A, Schröder J, Zech L: Quinacrine fluorescence of metaphase chromosomes: identical patterns in different tissues. Exp Cell Res 72:56, 1972

Caspersson T, Farber S, Foley GE, Kudynowski J, Modest EJ, Simonsson E, Wagh U, Zech L: Chemical differentiation along metaphase chromosomes. Exp Cell Res 49:219, 1968

Caspersson T, Gahrton G, Lindsten J, Zech L: Identification of the Philadelphia chromosome as a number 22 by quinacrine mustard fluorescence analysis. Exp Cell Res 63:238, 1970

Caspersson T, Haglund U, Lindell B, Zech L: Radiation-induced non-random chromosome breakage. Exp Cell Res 75:541, 1972

Caspersson T, Hulten M, Jonasson J, Lindsten J, Therkelsen A, Zech L: Translocations causing non-fluorescent Y-chromosomes in human XO/XY mosaics. Hereditas 68:2:317, 1971

Caspersson T, Issler P, Lomakka G: A TV-based technique for optical analysis of chromosome regions. Exp Cell Res 75:543, 1972

Caspersson T, Lindsten J, Zech L, Buckton KE, Price WH: Four patients with trisomy 8 identified by the fluorescence and Giemsa banding techniques. J Med Genet 9:1, 1972

Caspersson T, Lomakka G, Möller A: Computerized chromosome identification by aid of the quinacrine mustard fluorescence technique. Hereditas 67:1:103, 1971

Caspersson T, Lomakka G, Zech L: The 24 fluorescence patterns of the human metaphase chromosomes: distinguishing characters and variability. Hereditas 67:1:89, 1971

Caspersson T, Zech L, Johansson C: Analysis of human metaphase chromosome set by aid of DNA-binding fluorescent agents. Exp Cell Res 62:490, 1970

Caspersson T, Zech L, Johansson C, Modest EJ: Identification of human chromosomes by DNA-binding fluorescent agents. Chromosoma 30:215, 1970

Licznerski G, Lindsten J: Trisomy 21 in man due to maternal non-disjunction during the first meiotic division. Hereditas 70:1:153, 1972

Möller A, Nilsson H, Caspersson T, Lomakka G: Identification of human chromosome regions by aid of computerized pattern analysis. Exp Cell Res 70:475, 1972

Chromosome Identification: Techniques and Applications in Biology and Medicine. Nobel Symposium No 23, Stockholm, September 25-27, 1972. Caspersson T, Zech L, Reio V, Eds. Nobel Foundation, Stockholm, 1973

Chapter 4

German J: Bloom's syndrome. I. Genetical and clinical observations in the first twenty-seven patients. Am J Hum Genet 21:196, 1969

German J: Genes which increase chromosomal instability in somatic cells and predispose to cancer. Progress in Medical Genetics, Vol VIII. Steinberg AG, Bearn AG, Eds. Grune & Stratton, New York, 1972

Chapter 5

Carrell RW, Lehmann H: The unstable haemoglobin haemolytic anaemias. Semin Hematol 6:116, 1969

Conley CL, Charache S: Mechanisms by which some abnormal hemoglobins produce clinical manifestations. Semin Hematol 4:53, 1967

Heller P: Hemoglobinopathic dysfunction of the red cell. Am J Med 41:799, 1966

Huehns ER, Bellingham AJ: (annotation) Diseases of function and stability of haemoglobin. Br J Haematol 17:1, 1969

Jaffe ER, Heller P: Methemoglobinemia in man. Progress in Hematology. Brown EB, Moore CV, Eds. Grune & Stratton, New York, 1964

Lehmann H, Huntsman RG: Man's Hemoglobins. J. B. Lippincott Company, Philadelphia, 1966

Lehmann H, Perutz MF: Molecular pathology of human haemoglobin. Nature (Lond) 219:902, 1968

Stamatoyannopoulos G, Bellingham AJ, Lenfant C, Finch CA: Abnormal hemoglobins with high and low oxygen affinity. Annu Rev Med 22:221, 1971

The Metabolic Basis of Inherited Disease, 2d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1966

White JM, Dacie JV: The unstable hemoglobins: molecular and clinical features. Prog Hematol 7:69, 1971

Chapter 6

Harris H: Garrod's Inborn Errors of Metabolism, 1908 Ed. (with a supplement). Oxford University Press, Inc., Oxford, 1963

Knox WE: Sir Archibald Garrod's Inborn Errors of Metabolism, with introduction by Sir Rudolph Peters. Am J Hum Genet 10:3, 95, 249, 385, 1958

The Metabolic Basis of Inherited Disease, 2d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1966

Phenylketonuria. Lyman FL, Ed. Charles C Thomas, Springfield, Ill., 1963

Chapter 7

Barnes ND, Hull D, Balgobin L, Gompertz D: Biotin-responsive propionicacidaemia. Lancet 2:244, 1970

Glorieux F, Scriver CR: Loss of a parathyroid hormone-sensitive component of phosphate transport in X-linked hypophosphatemia. Science 175:997, 1972

Lonsdale D, Faulkner WR, Price JW, Smeby RR: Intermittent cerebellar ataxia associated with hyperpyruvicacidemia, hyperalaninemia and hyperalaninuria. Pediatrics 43:1025, 1969

Mahoney MJ, Rosenberg LE: Inherited defects of B_{12} metabolism. Am J Med 48:584, 1970

Mahoney MJ, Rosenberg LE, Mudd SH, Uhlendorf BW: Defective metabolism of vitamin B_{12} in fibroblasts from children with methylmalonic aciduria. Biochem Biophys Res Commun 44:2: 375, 1971

Mudd SH, Levy HL, Abeles RH, et al: A derangement in B_{12} metabolism leading to homocystinemia, cystathioninemia and methylmalonic aciduria. Biochem Biophys Res Commun 35:121, 1969

Rosenberg LE: Inherited aminoacidopathies demonstrating vitamin dependency. N Engl J Med 281:145, 1969

Rosenberg LE, Lilljeqvist AC, Hsia YE, Rosenbloom FM: Vitamin B_{12} dependent methylmalonic aciduria: defective B_{12} metabolism in cultured fibroblasts. Biochem Biophys Res Commun 37:607, 1969

Scriver CR: Vitamin dependency syndromes: their larger significance. Pediatrics 37:553, 1966

Scriver CR, MacKenzie S, Clow CL, Delvin E: Thiamine-responsive maple-syrup-urine disease. Lancet 1:310, 1971

Seashore MR, Durant JL, Rosenberg LE: Studies of the mechanism of pyridoxine-responsive homocystinuria. Pediatr Res 6:187, 1972

Williams HE, Smith LH Jr: Primary hyperoxaluria. The Metabolic Basis of Inherited Disease, 3d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972, p 196

Chapter 8

Asatoor AM, Cheng B, Edwards KD, Lant AF, Matthews DM, Milne MD, Navab F, Richards AJ: Intestinal absorption of two dipeptides in Hartnup disease. Gut 11:380, 1970

Baerlocher K, Scriver CR, Mohyuddin F: Ontogeny of iminoglycine transport in mammalian kidney. Proc Natl Acad Sci USA 65:1009, 1970

Berlin RD: Specificities of transport systems and enzymes. Science 168:1539, 1970

Christensen HN: Some special kinetic problems of transport. Adv Enzymol 32:1, 1969

Groth U, Rosenberg LE: Transport of dibasic amino acids, cystine, and tryptophan by cultured human fibroblasts: absence of a defect in cystinuria and Hartnup disease. J Clin Invest 51:2130, 1972

Korn ED: Current concepts of membrane structure and function. Fed Proc 28:6, 1969

Pardee AB: Membrane transport proteins. Science 162:632, 1968

Schultz SG: Mechanisms of absorption. Biological Membranes. Dowben RM, Ed. Little, Brown and Company, Boston, 1969, p 59

Scriver CR: Familial iminoglycinuria. The Metabolic Basis of Inherited Disease, 3d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972, p 1520

Scriver CR, Hechtman P: Human genetics of membrane transport with emphasis on amino acids. Advances in Human Genetics, Vol I. Harris H, Hirschhorn K, Eds. Plenum Publishing Corp., New York, 1970, p 211

Thier SO, Segal S: Cystinuria. The Metabolic Basis of Inherited Disease, 3d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972, p 1504

Chapter 9

Evans DAP: Clinical Pharmacogenetics. Recent Advances in Medicine. Baron DN, Compston N, Dawson AM, Eds. JA Churchill, Ltd., London, 1968, and The Williams & Wilkins Company, Baltimore, 1968, p 203

Kalow W: Pharmacogenetics: Heredity and the Response to Drugs. W. B. Saunders Company, Philadelphia, 1962

La Du BN: Pharmacogenetics: defective enzymes in relation to reactions to drugs. Annu Rev Med 23:453, 1972

Motulsky A: Pharmacogenetics. Progress in Medical Genetics, Vol III. Steinberg AG, Bearn AG, Eds. Grune & Stratton, New York, 1967, p 49

Vesell ES: Recent progress in pharmacogenetics. Adv Pharmacol Chemother 7:1, 1969

Vesell ES: Pharmacogenetics. N Engl J Med 287:904, 1972

Vesell ES: Introduction: pharmacogenetics (symposium). Fed Proc 31:1253, 1972

Weber WW: The relationship of genetic factors to drug reactions. Drug-Induced Diseases, Vol IV. Meyler L, Peck HM, Eds. Excerpta Medica, Amsterdam, 1972, p 3

Pharmacogenetics (symposium). La Du BN, Kalow W, consulting Eds. Ann NY Acad Sci 151:691, 1968

Experimental and clinical aspects of pharmacogenetics (symposium). Fed Proc 24:1259, 1965.

Chapter 10

Seminars on the clinical features of the Lesch-Nyhan syndrome. Bland JH, Ed. Fed Proc 27:1021, 1968

Fox IH, Kelley WN: Phosphoribosylpyrophosphate in man: biochemical and clinical significance. Ann Intern Med 74:424, 1971

Greene ML: Clinical features of patients with the "partial" deficiency of the X-linked uric aciduria enzyme. Arch Intern Med 130:193, 1972

Kelley WN: Biochemistry of the X-linked uric aciduria enzyme defect and its genetic variants. Arch Intern Med 130:199, 1972

Kelley WN, Greene ML, Rosenbloom FM, Henderson JF, Seegmiller JE: Hypoxanthine-guanine phosphoribosyltransferase deficiency in gout. Ann Intern Med 70:155, 1969

Kelley WN, Wyngaarden JB: The Lesch-Nyhan syndrome. The Metabolic Basis of Inherited Disease, 3d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972, p 969

Lesch M, Nyhan WL: A familial disorder of uric acid metabolism and central nervous system function. Am J Med 36:561, 1964

Nyhan WL: Clinical features of the Lesch-Nyhan syndrome. Arch Intern Med 130:186, 1972

Seegmiller JE: Diseases of purine and pyrimidine metabolism. Duncan's Diseases of Metabolism, 6th Ed. Bondy PK, Rosenberg LE, Eds. W. B. Saunders Company, Philadelphia, 1969

Seegmiller JE: Biochemical and genetic studies of an X-linked neurological disease. Harvey Lectures, Series 65. Academic Press, Inc., New York, 1971

Seegmiller JE: New prospects for understanding and control of genetic diseases. Arch Intern Med 130:181, 1972

Seegmiller JE, Rosenbloom FM, Kelley WN: Enzyme defect associated with a sex-linked human neurological disorder and excessive purine synthesis. Science 155:1682, 1967

Chapter 11

Granick S: The induction in vitro of the synthesis of delta-aminolevulinic acid synthetase in chemical porphyria: response to certain drugs, sex hormones and foreign chemicals. J Biol Chem 241:1359, 1966

Granick S, Urata G: Increase in activity of delta-aminolevulinic acid synthetase in liver mitochondria induced by feeding of 3,5-dicarbethoxy-1,4-dihydrocollidine. J Biol Chem 238:821, 1963

Kappas A, Bradlow HL, Gillette PN, Levere RD, Gallagher TF: A defect of steroid hormone metabolism in acute intermittent porphyria. Fed Proc 31:1293, 1972

Kappas A, Granick S: Steroid induction of porphyrin synthesis in liver cell culture. II. The effects of heme, uridine diphosphate glucuronic acid and inhibitors of nucleic acid and protein synthesis on the induction process. J Biol Chem 243:346, 1968

Levere RD, Kappas A: Biochemical and clinical aspects of the porphyrias. Adv Clin Chem 11:133, 1968

Marks GS: Heme and Chlorophyll: Chemical, Biochemical and Medical Aspects. Van Nostrand Co., Ltd., London, 1969

Marver HS, Schmid R: The porphyrias. The Metabolic Basis of Inherited Disease, 3d Ed. Stanbury JB, Wyngaarden JB and Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972

Strand LJ, Felsher BF, Redeker AG, Marver HS: Heme biosynthesis in intermittent acute porphyria: decreased hepatic conversion of porphobilinogen to porphyrins and increased delta aminolevulinic acid synthetase activity. Proc Natl Acad Sci USA 67:1315, 1970

Tschudy DP, Perlroth MG, Marver HS, Collins A, Hunter G Jr, Rechcigl M Jr: Acute intermittent porphyria: the first "overproduction disease" localized to a specific enzyme. Proc Natl Acad Sci USA 53:841, 1965

Waldenstrom J: The porphyrias as inborn errors of metabolism. Am J Med 22:758, 1957

Watson CJ: The problem of porphyria: some facts and questions. N Engl J Med 263:1205, 1960

Watson CJ, Runge W, Taddeini L, Bossenmaier I, Cardinal R: A suggested control gene mechanism for the excessive production of types I and 3 porphyrins in congenital erythropoietic porphyria. Proc Natl Acad Sci USA 52:478, 1964

Chapter 12

Arias IM: Hepatic aspects of bilirubin metabolism. Annu Rev Med 17:257, 1966

Arias IM: Inheritable and congenital hyperbilirubinemia. Models for the study of drug metabolism. N Engl J Med 285:1416, 1971

Arias IM: Transfer of bilirubin from blood to bile. Semin Hematol 9:55, 1972

Arias IM, Gartner LM, Cohen M, Ben Ezzer J, Levi AJ: Chronic nonhemolytic unconjugated hyperbilirubinemia with glucuronyl transferase deficiency. Am J Med 47:395, 1969

Fleischner G, Arias IM: Recent advances in bilirubin formation, transport, metabolism and excretion. Am J Med 49:576, 1970

Fleischner G, Robbins J, Arias IM: Immunological studies of Y protein: a major cytoplasmic organic anion-binding protein in rat liver. J Clin Invest 51:677, 1972

Gartner LM, Arias IM: Hormonal control of hepatic bilirubin transport and conjugation. Am J Physiol 222:1091, 1972

Levi AJ, Gatmaitan Z, Arias IM: Two hepatic cytoplasmic protein fractions, Y and Z, and their possible role in the hepatic uptake of bilirubin, sulfobromophthalein, and other anions. J. Clin Invest 48:2156, 1969

Litwack G, Ketterer B, Arias IM: Ligandin: a hepatic protein which binds steroids, bilirubin, carcinogens and a number of exogenous organic anions. Nature (Lond) 234:466, 1971

Schmid R: Hyperbilirubinemia. The Metabolic Basis of Inherited Disease, 2d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1966

Chapter 13

Cohen KL, Rubin PE, Echevarria RA, Sharp HL, Teague PO: Alpha-1-antitrypsin deficiency, emphysema, and cirrhosis in an adult. Arch Intern Med (in press)

Ericksson S: Studies in alpha-1-antitrypsin. Acta Med Scand [Suppl] 432:1, 1965

Fagerhol MK, Laurell CB: The Pi system-inherited variants of serum alpha-1-antitrypsin. Prog Med Genet 7:96, 1970

Glasgow JFT, Hercz A, Levison H, Lynch MJ, Sann-Kortsak A: Alpha-1-antitrypsin deficiency with both cirrhosis and chronic obstructive lung disease in two siblings. Pediatr Res 5:427, 1971

Lieberman J, Mittman C, Gordon HW: Alpha-1-antitrypsin in the livers of patients with emphysema. Science 175:63, 1972

Porter CA, Mowat AP, Cook PJL, Haynes DWA, Shilkin KB, Williams R: Alpha-1-antitrypsin deficiency and neonatal hepatitis. Br Med J 3:435, 1972

Sharp HL, Bridges RA, Krivit W, Freier EF: Cirrhosis associated with alpha-1-antitrypsin deficiency: a previously unrecognized inherited disorder. J Lab Clin Med 73:934, 1969

Sharp HL, Mathis R, Krivit W, Freier E: The liver in noncirrhotic alpha-1-antitrypsin. J Lab Clin Med 78:1012, 1971

Stevens PM, Hnilica VS, Johnson PC, Bell RL: Pathophysiology of hereditary emphysema. Ann Intern Med 75:672, 1971

Talamo RC, Allen JD, Kahan MG, Austen KF: Hereditary alpha-1-antitrypsin deficiency. N Engl J Med 278:345, 1968

Welch MH, Reinecke ME, Hammarsten JF, Guenter CA: Antitrypsin deficiency in pulmonary disease: the significance of intermediate levels. Ann Intern Med 71:533, 1969

Chapter 14

Bach G, Friedman R, Weissmann B, Neufeld EF: The defect in the Hurler and Scheie syndromes: deficiency of L-L-iduronidase. Proc Natl Acad Sci USA 69:2048, 1972

Danes BS, Bearn AG: Hurler's syndrome: a genetic study in cell culture. J Exp Med 123:1, 1966

DiFerrante N, Nichols BL, Donnelly PV, Neri G, Hrgovcic R, Berglund RK: Induced degradation of glycosaminoglycans in Hurler's and Hunter's syndromes by plasma infusion. Proc Natl Acad Sci USA 68:303, 1971

Dorfman A, Matalon R: The mucopolysaccharidoses. The Metabolic Basis of Inherited Disease, 3d Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972

Fratantoni JC, Hall CW, Neufeld EF: The defect in Hurler and Hunter syndromes. II. Deficiency of specific factors involved in mucopolysaccharide degradation. Proc Natl Acad Sci USA 64:360, 1969

Fratantoni JC, Neufeld EF, Uhlendorf BW, Jacobson CB: Intrauterine diagnosis of the Hurler and Hunter syndromes. N Engl J Med 280:686, 1969

Knudson AG Jr, DiFerrante N, Curtis JE: Effect of leukocyte transfusion in a child with type II mucopolysaccharidosis. Proc Natl Acad Sci USA 68:1738, 1971

Kresse H, Neufeld EF: The Sanfilippo A corrective factor: purification and mode of action. J Biol Chem 247:2164, 1972

McKusick VA: Heritable Disorders of Connective Tissue, 3d Ed. The C. V. Mosby Company, St. Louis, 1966

Neufeld EF, Cantz MJ: Corrective factors for inborn errors of mucopolysaccharide metabolism. Ann NY Acad Sci 179:580, 1971

O'Brien JS: Sanfilippo syndrome: profound deficiency of alpha acetylglucosaminidase activity in organs and skin fibroblasts from type-B patients. Proc Natl Acad Sci USA 69:1720, 1972

Van Hoof F, Hers HG: L'ultrastructure des cellules hepatiques dans la maladie de Hurler (gargoylisme). Comptes Rendus Acad Sci (Paris) 259:1281, 1964

Chapter 15

Costom BH, Grumbach MM, Kaplan SL: Effect of thyrotropin-releasing factor on serum thyroid-stimulating hormone. An approach to distinguishing hypothalmic from pituitary forms of idiopathic hypopituitary dwarfism. J Clin Invest 50:2219, 1971

Daughaday WH: Sulfation factor: regulation of skeletal growth. Am J Med 50:277, 1971

Daughaday WH: The adenohypophysis. Textbook of Endocrinology, 4th Ed. Williams RH, Ed. W. B. Saunders Company, Philadelphia, 1968, p 27

Daughaday WH, Laron Z, Pertzelan A, Heins JN: Defective sulfation factor generation: a possible etiological link in dwarfism. Trans Assoc Am Physicians 82:129, 1969

Goodman HG, Grumbach MM, Kaplan SL: Growth and growth hormone. II. Comparison of isolated growth-hormone deficiency and multiple pituitary hormone deficiencies in 35 patients with idiopathic hypopituitary dwarfism. N Engl J Med 278:57, 1968

Laron Z: The hypothalamus and the pituitary gland (hypophysis). Paediatric Endocrinology. Hubble D, Ed. Blackwell, Oxford, 1969

Merimee TJ, Rimoin DL, Cavalli-Sforza LL: Metabolic studies in the African pygmy. J Clin Invest 51:395, 1972

McKusick VA, Rimoin DL: General Tom Thumb and other midgets. Sci Am 217:102, 1967

Rimoin DL, Merimee TJ, Rabinowitz D, McKusick VA: Genetic aspects of clinical endocrinology. Recent Prog Horm Res 24:365, 1968

Rimoin DL, Merimee TJ, Rabinowitz D, Cavalli-Sforza LL, McKusick VA: Peripheral subresponsiveness to human growth hormone in the African pygmies. N Engl J Med 281:1383, 1969

Rimoin DL, Schimke RN: Genetic Disorders of the Endocrine Glands. The C. V. Mosby Company, St. Louis, 1971

Chapter 16

Krooth RS, Sell EK: The action of mendelian genes in human diploid cell strains. J Cell Physiol 76:311, 1970

Littlefield JW: Control mechanisms in animal cell cultures. Arch Biochem Biophys 125:410, 1968

Mellman WJ, Cristofalo VJ: Human diploid cultures: their usefulness in the study of genetic variations in metabolism. Growth, Nutrition, and Metabolism of Cells in Culture, Vol I. Rothblatt GH, Cristofalo VJ, Eds. Academic Press, Inc., New York, 1972, p 327

Migeon BR: Somatic cell hybrids: applications relevant to genetic disease. J Pediatr 79:887, 1971

Nadler HL, Chacko CM, Rachmeler M: Interallelic complementation in hybrid cells derived from human diploid strains deficient in galactose-1-phosphate uridyl transferase activity. Proc Natl Acad Sci USA 67:2:976, 1970

Ruddle FH: Linkage analysis using somatic cell hybrids. Advances in Human Genetics, Vol III. Harris H, Hirschhorn K, Eds. Plenum Publishing Corp., New York, 1972, p 173

Tedesco TA, Mellman WJ: Galactose-1-phosphate uridyltransferase and galactokinase activity in cultured human diploid fibroblasts and peripheral blood leukocytes: I. Analysis of transferase genotypes by the ratio of the activities of the two enzymes. J Clin Invest 48:12:2390, 1969

Chapter 17

Benacerraf B, McDevitt HO: Histocompatibility-linked immune response genes. Science 175:273, 1972

Klein J, Shreffler DC: The H-2 model for the major histocompatibility systems. Surface Antigens on Nucleated Cells, Vol VI Transplantation Reviews. Möller G, Ed. Munksgaard, Copenhagen, and The Williams & Wilkins Company, Baltimore, 1971, p 3

Klein J, Shreffler DC: Evidence supporting a two-gene model for the H-2 histocompatibility system of the mouse. J Exp Med 135:924, 1972

McDevitt HO, Benacerraf B: Genetic control of specific immune responses. Adv Immunol 11:31, 1970

Mintz B: Genetic mosaicism in vivo: development and disease in allophenic mice. Fed Proc 30:935, 1971

Sela M: Antigenicity: some molecular aspects. Science 166:1365, 1969

Chapter 18

Bach FH: Transplantation: pairing of donor and recipient. Science 168:1170, 1970

Bach FH: The major histocompatibility complex in transplantation immunology. Transplant Proc (in press)

Bach FH, Amos DB: Hu-1: major histocompatibility locus in man. Science 156:1506, 1967

Bach FH, Solliday S, Stambuk M: Detection of disparity in the mixed leukocyte culture test: a more rapid assay. Histocompatibility Testing, 1970. Munksgaard, Cophenhagen, p 643

Bach FH, Widmer M, Segell M, Bach ML, Klein J: Genetic and immunological complexity of major histocompatible regions. Science 176:1024, 1972

Ceppellini R: Advances in transplantation. First International Congress of the Transplantation Society – Proceedings. Dausset J, Hamburger J, Mathe G, Eds. The Williams & Wilkins Company, Baltimore, 1968, p 195

Dausset J, Ivanyl P, Ivanyl D: Histocompatibility Testing, 1965. Munksgaard, Copenhagen, 1965, p 51

Grove JS, Bach FH: Modern Trends in Human Genetics. Emery AEH, Ed. Butterworth & Company, Ltd., London, 1969, p 104

Morris RJ, Kincaid-Smith P, Ting A, Stocker JW, Marshall VC: Prospective leukocyte typing in cadaver renal transplantation. Lancet 2:803, 1968

Patel R, Mickey MR, Terasaki PI: Serotyping for homotransplantation. XVI. Analysis of kidney transplants from unrelated donors. N Engl J Med 279:501, 1968

Shreffler DC: The Third International Congress of Human Genetics. Crow JF, Neel JV, Eds. Johns Hopkins University Press, Baltimore, 1967, p 217

Van Rodd JJ, Eernisse JG: The detection of transplantation antigens in leukocytes. Semin Hematol 5:2:187, 1968

Van Rodd JJ, Vooys WH, Fredericks E, Balner H, Eernisse JG: Histocompatibility Testing, 1965. Munksgaard, Copenhagen, p 37

Chapter 19

Carter CO: Genetics of common disorders. Br Med Bull 25:52, 1969

Carter CO: Multifactorial inheritance revisited. Congenital Malformations: Proceedings Third International Conference. Fraser FC, McKusick VA, Eds. Excerpta Medica, Amsterdam, 1970

Carter CO: Genetics of common malformations. Recent Advances in Paediatrics. Gairdner D, Hull D, Eds. Churchill, London, 1971

Edwards JH: Familial predisposition in man. Br Med Bull 25:58, 1969

Falconer DS: The inheritance of liability to certain diseases estimated from the incidence among relatives. Ann Hum Genet 29:51, 1965

Roberts JAF: Multifactorial inheritance in relation to normal and abnormal human traits. Br Med Bull 17:241, 1961

Chapter 20

McKusick VA: Heritable Disorders of Connective Tissue, 4th Ed. The C. V. Mosby Company, St. Louis, 1972

McKusick VA: Mendelian Inheritance in Man. Catalogues of Autosomal Dominant, Autosomal Recessive and X-linked Phenotypes. 3d Ed. Johns Hopkins University Press, Baltimore, 1971

McKusick VA: The distribution of certain genes in the Old Order Amish. Cold Spring Harbor Symp Quant Biol 29:99, 1964

McKusick VA: Phenotypic diversity in human disease resulting from allelic series. Am J Hum Genet (in press)

Mahlondji M, Teasdall RD, Adamkiewicz JJ, Hartman WH, Lambird PA, McKusick VA: The genetic amyloidoses with particular reference to hereditary neuropathic amyloidoses, Type II (Indiana or Rukavina type). Medicine 48:1, 1969

Chapter 21

Carter CO: An ABC of medical genetics. II. Chromosomes and chromosome mutations. Lancet 1:104, 1969

Carter CO, Roberts JAF, Evans KA, Buck AR. Genetic clinic. A follow-up. Lancet 1:281, 1971

Fraser FC: Genetic counselling and the physician. Can Med Assoc J 99:927, 1968

Fraser FC: Counseling in genetics: its intent and scope. Genetic Counseling. Bergsma D, Ed. Birth Defects: Original Articles Series, Vol VI, No 1, May 1970. The National Foundation – March of Dimes

Gordon H: Genetic counseling. Considerations for talking to parents and prospective parents. JAMA 217:1215, 1971

Murphy EA: The rationale of genetic counselling. J Pediatr 72:121, 1968

Stevenson AC, Davison BCC: Genetic Counseling. William Heinemann, Ltd., London, 1970, and J. B. Lippincott Company, Philadelphia, 1971

Chapter 22

Allan JD, Brown JK: Maternal phenylketonuria and foetal brain damage. An attempt at prevention by dietary control. Some Recent Advances in Inborn Errors of Metabolism. Holt KS, Coffey VP, Eds. E & S Livingstone, Ltd., Edinburgh and London, 1968

Guthrie R: Screening for inborn errors of metabolism in the newborn infant: a multiple test program. Human Genetics. Bergsma D, Ed. Birth Defects: Original Articles Series Vol IV, No 6, 1968. The National Foundation – March of Dimes

Guthrie R: Screening for phenylketonuria. Triangle 9:3:104, 1969

Levy HL: Genetic screening. Advances in Human Genetics, Vol III. Harris H, Hirschhorn K, Eds. Plenum Publishing Corp., New York, 1972

Murphey WH, Patchen L, Guthrie R: Screening tests for argininosuccinic aciduria, orotic aciduria, and other inherited enzyme deficiencies using dried blood specimens. Biochem Genet 6:1, 1972

Proceedings of the International Symposium on Phenylketonuria and Allied Disorders, Tel Aviv, June 15-22, 1969. Cohen BE, Rubin MI, Szeinberg A, Eds. Translators' Pool, Israel, 1971

MacCready RA, Levy HL: The problem of maternal phenylketonuria. Am J Obstet Gynecol 113:121, 1972

Chapter 23

Berman JL, Ford R: Intelligence quotients and intelligence loss in patients with phenylketonuria and some variant states. J Pediatr 77:764, 1970

Holtzman NA, Meek AG, Mellits ED: Effectiveness of neonatal phenylketonuria screening. Pediatr Res 6:358, 1972

Holtzman NA: Dietary treatment of inborn errors of metabolism. Annu Rev Med 21:335, 1970

Yu JS, O'Halloran MT, Stuckey SJ: False negative screening tests in phenylketonuria. Arch Dis Child 46:124, 1971

Levy HL, Karolkewicz V, Haughton SA, MacCready RA: Screening the "normal" population in Massachusetts for phenylketonuria. N Engl J Med 282:1455, 1970

Chapter 24

Early Diagnosis of Human Genetic Defects: Scientific and Ethical Considerations. Harris M, Ed. US Government Printing Office, HE 20. 3710:6, 1970

Antenatal Diagnosis. Dorfman A, Ed. NICHD Mental Retardation Series. The University of Chicago Press, Chicago, 1972

Milunsky A, Littlefield JW, Kanvfer JN, Kolodny EH, Shih VE, Atkins L: Prenatal genetic diagnosis. N Engl J Med 283: 1498, 1970

Nadler HL, Gerbie AB: Role of amniocentesis in the intrauterine detection of genetic disorders. N Engl J Med 282:596, 1970

Chapter 25

Kaback MM, Zeiger RS: Heterozygote detection in Tay-Sachs disease: Prototype community screening program for the prevention of recessive genetic disorders. Sphingolipids, Sphingolipidoses and Allied Disorders, Vol IX Advances in Experimental Medicine and Biology. Volk BW, Aronson M, Eds. Plenum Publishing Corp., New York, 1972, p 613

Kaback MM, Zeiger RS: The John F. Kennedy Institute Tay-Sachs program: Practical and ethical issues in an adult genetic screening program. Ethical Issues in Human Genetics. Genetic Counseling and the Use of Genetic Knowledge. Hilton B, Callahan D, Harris M, Condliffe P, Berkley B, Eds. Plenum Publishing Corp., New York, 1973, p 131

O'Brien JS: Ganglioside storage diseases. Advances in Human Genetics, Vol III. Harris H, Hirschhorn K, Eds. Plenum Publishing Corp., New York, 1972, p 39

Chapter 26

Batchelor JR, French ME, Cameron JS, et al: Immunological enhancement of human kidney graft. Lancet 2:1007, 1970

Clarke CA: Blood group incompatibility between mother and fetus. Clin Sci 43:1, 1972

Clarke CA: Prevention of Rh haemolytic disease. Br Med J 4:7, 1967; 4:484, 1967; 1:57, 1968

Clarke CA: The prevention of "Rhesus" babies. Sci Am 219:46, 1968

Clarke CA, Donohoe WTA, Finn R, Lehane MB, McConnell RB, Sheppard PM, et al: Prevention of Rh haemolytic disease: final results of the "high risk" clinical trial (a combined study for centers in England and Baltimore, Maryland). Br Med J 2:607, 1971

Clarke CA, McConnell RB: Prevention of Rh Haemolytic Disease. American Lecture Living Chemistry Series. Charles C Thomas, Springfield, Ill., 1972, p 122

Finn R, Clarke CA, Donohoe WTA, McConnell RB, Sheppard PM, Lehane D, Kulke W: Experimental studies on the prevention of Rh haemolytic disease. Br Med J 5238:1486, 1961

Freda VJ, Gorman JG: Ante-partum management of Rh hemolytic disease. Bulletin of the Sloane Hospital for Women 8:147, 1962

Mollison PL: Blood Transfusion in Clinical Medicine, 5th Ed. Davis Co., Oxford, 1972

Woodrow JC: Rh Immunization and Its Prevention. Series Haematologica, III. Munksgaard, Cophenhagen, 1970, p 151

Chapter 27

Clow CL, Reade TM, Scriver CR: Management of hereditary metabolic disease. The role of allied health personnel. N Engl J Med 284:1292, 1971

Davis BD: Prospects for genetic intervention in man. Science 170:1279, 1970

DiFerrante N, Nichols BL, Donnelly PV, Neri G, Hrgovcic R, Berglund RK: Induced degradation of glycosaminoglycans in Hurler's and Hunter's syndrome by plasma infusion. Proc Natl Acad Sci USA 68:303, 1971

Hermann RE, Mercer RD: Portacaval shunt in the treatment of glycogen storage disease: report of a case. Surgery 65:499, 1969

Knudson AG Jr, DiFerrante N, Curtis JE: Effect of leukocyte transfusion in a child with type II mucopolysaccharidosis. Proc Natl Acad Sci USA 68:1738, 1971

Mahoney CP, Striker GE, Hickman RO, Manning GB, Marchioro TL: Renal transplantation for childhood cystinosis. N Engl J Med 283:397, 1970

Merril CR, Geier MR, Petricciani JC: Bacterial virus gene expression in human cells. Nature (Lond) 233:398, 1971

Starzl TE, Brown BI, Blanchard H, Brettschneider L: Portal diversion in glycogen storage disease. Surgery 65:504, 1969

Stevenson RE, Howell RR: Some medical and social aspects of the treatment for genetic-metabolic diseases. Annals of the American Academy of Political and Social Science 399:30, 1972

The Metabolic Basis of Inherited Disease, 3rd Ed. Stanbury JB, Wyngaarden JB, Fredrickson DS, Eds. McGraw-Hill, Inc., New York, 1972

Index

* *denotes illustration or table*

C

* *denotes illustration or table*

H

* *denotes illustration or table*

I

J

K

* *denotes illustration or table*

N

O

P

Q

R

S

* *denotes illustration or table*

T

WELLCOME LIBRARY INSTITUTE

Edited by Victor A. McKusick, M.D., Chief, Division of Medical Genetics at Johns Hopkins University, and Robert Claiborne, Senior Editor of **Hospital Practice,** "Medical Genetics" details these gains in understanding of this complex subject, in detecting and in anticipating genetic errors, and in managing the mental and physical disorders they cause.

Experts whose names recur frequently in the literature of genetics (see contents listed on back of jacket) cover every major clinical and research aspect of the field:

- The latest knowledge of the mechanisms of transmission of genetic defects and their resultant diseases;
- Chromosomal anomalies involving specific body systems and biochemical pathways;
- Current applications of genetic knowledge through screening, counseling, therapy, and prevention;
- Techniques that are speeding the advance of knowledge in medical genetics and expanding its clinical availability.

For the clinician, the research scientist, the teacher, the house officer and the student preparing for practice, "Medical Genetics," new and authoritative, will, we hope, be an invaluable aid.

The Editors:

VICTOR A. McKUSICK, M.D., is Professor of Medicine, Epidemiology and Biology at Johns Hopkins University, Baltimore, and Chief of its Division of Medical Genetics. He recently was elected a member of the National Academy of Sciences and earlier received the John Phillips Award of the American College of Physicians. He has been chosen to serve as President of the American Society of Human Genetics for the 1974 term.

ROBERT CLAIBORNE is Senior Editor of **Hospital Practice.** He has been an Associate Editor of *Scientific American,* Managing Editor of *Medical World News,* and Editor of the LIFE Science Library. His books include "Climate, Man and History" and "On Every Side the Sea."