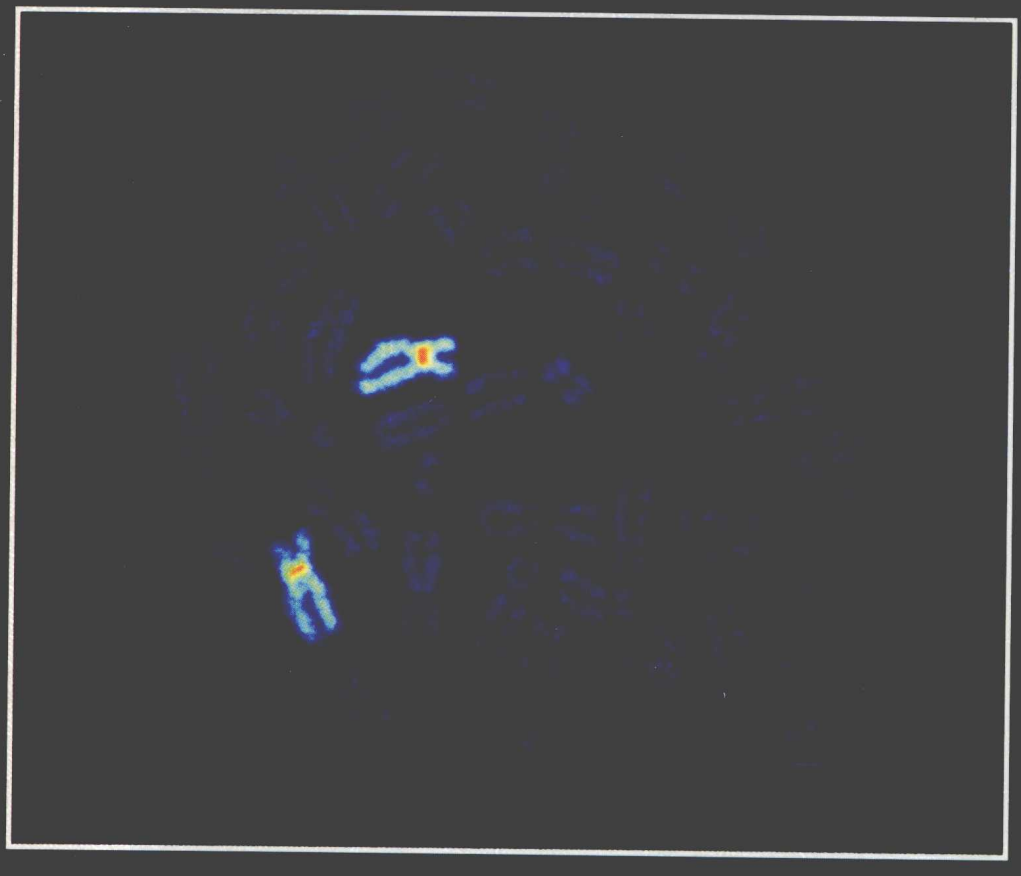
Etiology of Human Disease at the DNA Level

Editors

JAN LINDSTEN ■ ULF PETTERSSON



Nobel Symposium 80

RAVEN PRESS

Etiology of Human Disease at the DNA Level

Editors

Jan Lindsten, M.D., Ph.D.

Department of Clinical Genetics
Karolinska Hospital
Stockholm, Sweden

Ulf Pettersson, M.D., Ph.D.

Department of Medical Genetics
University of Uppsala
Uppsala, Sweden

NOBEL SYMPOSIUM 80

RAVEN PRESS NEW YORK

Raven Press, Ltd., 1185 Avenue of the Americas, New York, New York 10036

Made in the United States of America

Library of Congress Cataloging-in-Publication Data

Etiology of human disease at the DNA level / edited by Jan Lindsten, Ulf Pettersson.
p. cm.
Consists of papers presented at the Nobel Symposium on the "Etiology of Human Disease at the DNA Level," in Björkborn, Karlskoga, Sweden, June 11 and 14, 1990, and sponsored by the Nobel Foundation and the Alfred Nobel's Björkborn Foundation.
Includes bibliographical references and index.
1. Medical genetics—Congresses. I. Lindsten, Jan E. II. Pettersson, Ulf. III. Nobelstiftelsen. IV. Alfred Nobel's Björkborn Foundation. V. Nobel Symposium on the "Etiology of Human Disease at the DNA Level" (1990 : Björkborn, Karlskoga, Sweden)
[DNLM: 1. Chromosome Mapping—congresses. 2. Gene Therapy—congresses. 3. Genetic Markers—congresses. 4. Hereditary Diseases—diagnosis—congresses. 5. Hereditary Diseases—etiology—congresses. 6. Hereditary Diseases—therapy—congresses. QZ 50 E847 1990]
RB155.E828 1991
616'.042—dc20
DNLM/DLC
for Library of Congress 91-8911
ISBN 0-88167-761-2 CIP

The material contained in this volume was submitted as previously unpublished material, except in the instances in which credit has been given to the source from which some of the illustrative material was derived.

Great care has been taken to maintain the accuracy of the information contained in the volume. However, neither Raven Press nor the editors can be held responsible for errors or for any consequences arising from the use of the information contained herein.

Materials appearing in this book prepared by individuals as part of their official duties as U.S. Government employees are not covered by the above-mentioned copyright.

9 8 7 6 5 4 3 2 1

Contributors

Antonio Baldini, M.D. *Department of Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510*

S. Gwyn Ballard, Ph.D. *Department of Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510*

Alan H. Beggs, Ph.D. *Division of Genetics, The Children's Hospital, 300 Longwood Ave., Boston, Massachusetts 02115*

Graeme I. Bell, Ph.D. *Howard Hughes Medical Institute Research Laboratory, Department of Biochemistry and Molecular Biology, The University of Chicago, 5841 S. Maryland Ave., N237A, Box 391, Chicago, Illinois 60637*

John W. Belmont, M.D., Ph.D. *Institute for Molecular Genetics and Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030*

Sir Walter F. Bodmer, M.R.C.P. *Imperial Cancer Research Fund Labs, P.O. Box 123, Lincoln Inn Fields, London WC2A 3PX, U.K.*

David Botstein, Ph.D. *Department of Genetics, Stanford University School of Medicine, Stanford, California 94305-5120*

Frederick M. Boyce, M.D., Ph.D. *Division of Genetics, The Children's Hospital, 300 Longwood Ave., Boston, Massachusetts 02115*

Ann Boyle *Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, 333 Cedar St., New Haven, Connecticut 06510*

Michael S. Brown, M.D. *Department of Molecular Genetics, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, Texas 75235*

C. Brunn *Johns Hopkins Hospital, William Welch Medical Library, 1830 E. Monument St., 3rd Floor, Baltimore, Maryland 21205*

Mario R. Capecchi, Ph.D. *Howard Hughes Medical Institute, Department of Human Genetics and Biology, University of Utah, 337 South Biology Building, Salt Lake City, Utah 84112*

C. Thomas Caskey, M.D. *Institute for Molecular Genetics and Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030*

Denis Cournoyer, M.D. *Department of Medicine, McGill University, Montreal, Canada H3G 1Y6; Division of Hematology, Montreal General Hospital, 1650 Avenue Cedar, Montreal, Quebec, Canada H2L 1H3*

Nancy J. Cox, Ph.D. *Howard Hughes Medical Institute Research Laboratory, Department of Biochemistry and Molecular Biology, The University of Chicago, 5841 S. Maryland Ave., N237A, Box 391, Chicago, Illinois 60637*

Earl W. Davie, Ph.D. *Department of Biochemistry, University of Washington, J405 Health Sciences Building, SJ-70, Seattle, Washington 98195*

Stefan S. Fajans, M.D. *Department of Internal Medicine, Division of Endocrinology and Metabolism, University of Michigan Medical Center, Ann Arbor, Michigan 48109*

Chris A. Feener *Division of Genetics, The Children's Hospital, 300 Longwood Ave., Boston, Massachusetts 02115*

***Frederick A. Fletcher, M.Sc.** *Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030*

Joseph L. Goldstein, M.D. *Department of Molecular Genetics, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, Texas 75235*

Fiona R. Green *Arterial Disease Research Unit, The Charing Cross Sunley Research Centre, 1 Lurgan Ave., Hammersmith, London SW6 6JX, U.K.*

James F. Gusella, M.D. *Molecular Neurogenics Laboratory, Massachusetts General Hospital; Department of Genetics, Harvard University, Boston, Massachusetts 02114*

Stephen Humphries, Ph.D., M.R.C.Path. *Arterial Disease Research Unit, The Charing Cross Sunley Research Centre, 1 Lurgan Ave., Hammersmith, SW6 6JX, U.K.*

Rudolf Jaenisch, M.D. *Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142*

Cecily H. Kelleher, M.D. *Department of Health Promotion, University College Galway, Galway, Ireland*

**Present address:* Department of Experimental Hematology, Immunex R&D Corporation, 51 University Street, Seattle, Washington 98101.

Louis M. Kunkel, Ph.D. *Division of Genetics, The Children's Hospital, 300 Longwood Ave., Boston, Massachusetts 02115*

***Peter Lichter, Ph.D.** *Department of Human Genetics, Yale University School of Medicine, 333 Cedar St., New Haven, Connecticut 06510*

R. Lucier, Ph.D. *Johns Hopkins Hospital, William Welch Medical Library, 1830 E. Monument St., 3rd Floor, Baltimore, Maryland 21205*

Grant R. MacGregor, Ph.D. *Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030*

Hugh O. McDevitt, M.D. *Department of Microbiology and Immunology, Sherman Fairchild Science Building, Stanford University School of Medicine, Stanford, California 94305-5402*

Tom W. Meade, D.M., F.R.C.P., F.F.C.M. *MRC Epidemiology and Medical Care Unit, Northwick Park Hospital, Harrow HA1 3UJ, U.K.*

Joan Menninger *Department of Human Genetics, Yale University School of Medicine, 333 Cedar St., New Haven, Connecticut 06510*

Marilyn Monk, Ph.D. *MRC Mammalian Development Unit, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, U.K.*

Kateri A. Moore, D.V.M. *Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030*

Richard C. Mulligan, Ph.D. *Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, and Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142*

Marsha Newman *Howard Hughes Medical Institute and Departments of Biochemistry and Molecular Biology, The University of Chicago, 5841 S. Maryland Ave., Chicago, Illinois 60637*

Richard J. O'Reilly, M.D. *Department of Pediatrics, Memorial Sloan-Kettering Cancer Center, 1275 York Ave., New York, New York 10021*

P. L. Pearson, Ph.D. *Johns Hopkins Hospital, 1830 E. Monument St., 3rd Floor, Baltimore, Maryland 21205*

****Maurizio Scarpa, M.D.** *Department of Immunology, University College and Middlesex School of Medicine, London, U.K. W1*

**Present address:* Deutsches Krebsforschungszentrum, In Neuenheimer Feld 280, D-6900 Heidelberg, Germany

***Present address:* Department of Pediatrics University of Padova, Via Gustiniani 3, 35100 Padova, Italy.

Mitsuko Seino, R.N. *Howard Hughes Medical Institute Research Laboratory, Department of Biochemistry and Molecular Biology, The University of Chicago, 5841 S. Maryland Ave., N237A, Box 391, Chicago, Illinois 60637*

Susumu Seino, M.D., D.M.Sc. *Howard Hughes Medical Institute Research Laboratory, Department of Biochemistry and Molecular Biology, The University of Chicago, 5841 S. Maryland Ave., N237A, Box 391, Chicago, Illinois 60637*

Oliver Smithies, Ph.D. *Department of Pathology, School of Medicine, University of North Carolina at Chapel Hill, 7525 Brinkhous-Bullitt Building, Chapel Hill, North Carolina 27599-7525*

Judith R. Snyder *Division of Genetics, The Children's Hospital, 300 Longwood Ave., Boston, Massachusetts 02115*

Angela E. Thomas, M.B., M.R.C.P., M.R.C.Path. *The Charing Cross Sunley Research Centre, 1 Lurgan Avenue, Hammersmith, SW6 6JX, U.K.*

David C. Ward, Ph.D. *Departments of Human Genetics and Molecular Biophysics and Biochemistry, Yale University School of Medicine, 333 Cedar St., New Haven, Connecticut 06510*

Robert Williamson, Ph.D., F.R.C.Path. *Department of Biochemistry and Molecular Genetics, St. Mary's Hospital Medical School, Paddington, London W2 1PG, U.K.*

Debra J. Wolgemuth, Ph.D. *Department of Genetics and Development, College of Physicians and Surgeons of Columbia University, New York, New York 10032*

Song-hua Wu, M.D. *Howard Hughes Medical Institute Research Laboratory, Department of Biochemistry and Molecular Biology, The University of Chicago, 5841 S. Maryland Ave., N237A, Box 291, Chicago, Illinois 60637*

Preface

In addition to awarding the Nobel prizes, the Nobel Institution is involved in various scientific activities. Most important of these are the Nobel symposia, which are sponsored by the Nobel Foundation and its Symposium Committee.

The Nobel symposia are traditionally arranged according to certain rules, the most important one being that the topics should be of great current scientific interest. Furthermore, the meetings are small and closed; only 40–50 internationally leading scientists are invited.

The present Nobel symposium on the etiology of human disease at the DNA level is number 80 in this series of symposia. It was held at Alfred Nobel's Björkborn, which is located in the middle of Sweden, between June 11 and 14, 1990. The meeting covered an extremely interesting aspect of biomedical research. The field of human molecular genetics is advancing at an enormous pace. We were fortunate to attract the leading scientists in the field to the symposium; they discussed the molecular basis of human disease from all possible angles. The chapters in this book represent the forefront of human genetics; the book should be of wide interest to students and scientists in the field of human genetics.

Jan Lindsten
Ulf Pettersson

Stockholm and Uppsala, Sweden
October 1990

Acknowledgments

We would like to thank all the participants for their stimulating contributions, which are presented in this book. We would also like to thank the Nobel Foundation and Alfred Nobel's Björkborn Foundation for sponsoring the symposium and Raven Press for publishing these proceedings. Special thanks goes to Gertie Ågren, Anita Lundmark, and Ann-Mari Dumanski for their excellent help with practical arrangements.

Contents

Mapping of Human Disease Traits

Etiology of Human Disease at the DNA Level,
edited by Jan Lindsten and Ulf Pettersson.

1

Linkage Genetics in Humans: Origins and Prospects

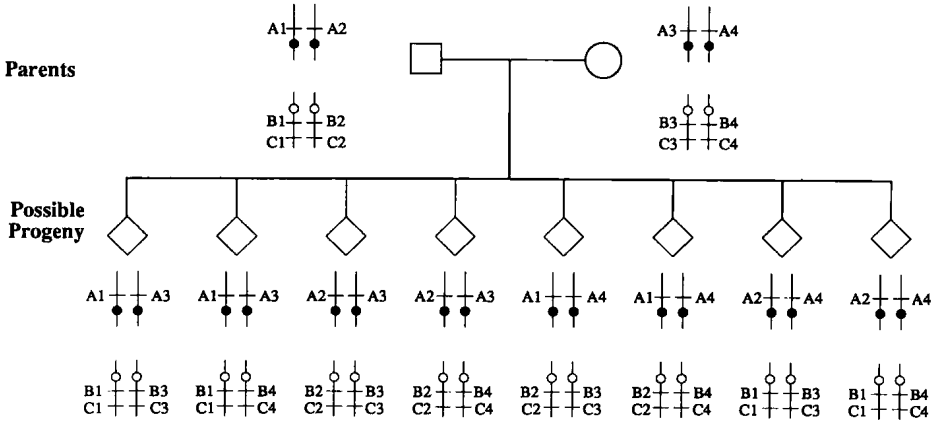
David Botstein

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305

The rediscovery of Mendel in the first years of this century was immediately followed by the recognition that some human traits, indeed some human diseases, are inherited according to Mendel's laws. It soon became unmistakably clear that the mechanisms of heredity in humans are entirely typical of those in all higher eukaryotes. Among the earliest post-Mendelian discoveries was the principle of genetic linkage and the idea of a linkage map (Sturtevant, 1913). Yet linkage mapping in humans was not practiced on a large scale until after 1980, even though the applicability of the principle and, indeed, a number of essential statistical methods for detecting linkage had been in place since the 1930s (see Ott, 1985, for a review). It is the purpose of this short chapter to sketch briefly how linkage mapping was finally applied to inherited human diseases, to place the ideas behind human genetic linkage mapping in their proper historical context, and to give some inkling of the future and limits to linkage mapping in humans.

LINKAGE MAPPING WITH DNA POLYMORPHISMS

The limiting factor that made linkage mapping difficult between the 1930s and the 1980s, even for diseases obviously inherited in a simple Mendelian way, was the supply of adequately polymorphic genes that could serve as markers. Figure 1 illustrates that Mendelian inheritance and, indeed genetic linkage, can be observed readily, given only loci (marked A, B and C) such that the four alleles at each locus in the parents (A1, A2 in the father and A3, A4 in the mother; B1, B2 in the father and B3, B4 in the mother; and so on) can all be distinguished. As shown in Fig. 1, when all the alleles can be distinguished, it is easy to see segregation of the parental alleles (each child gets either A1 or A2 from the father and A3 or A4 from the mother, etc.).



Totals	A1 and B1 = 2	A,B all equal frequencies No linkage	A3 and B3 = 2
	A1 and B2 = 2		A3 and B4 = 2
	A2 and B1 = 2		A4 and B3 = 2
	A2 and B2 = 2		A4 and B4 = 2
	C1 and B1 = 4	C,B unequal frequencies Linkage	C3 and B3 = 4
	C1 and B2 = 0		C3 and B4 = 0
	C2 and B1 = 0		C4 and B3 = 0
	C2 and B2 = 4		C4 and B4 = 4

FIG. 1. Mendelian inheritance, and genetic linkage, can be observed given loci (A, B, C) such that the four alleles at each locus in the parents can be observed (A1, A2 in the father and A3, A4 in the mother, etc.). See text for details.

Independence of inheritance of alleles residing on different chromosomes is equally easily seen (e.g., A1 from the father is accompanied by B2 as frequently as B1; the combinations A3,B3; B3,B4; A4,B3; A4,B4 are equally frequently inherited from the mother; etc.). Linkage is easily observed as well, as the inheritance of C1 is highly correlated with the inheritance of B1; C2 is correlated with B2 (both from the father); on the mother's side we see correlation of C3 with B3 and C4 with B4 in Fig. 1. Note also that we see, from these correlations, not only that loci B and C are linked, but also the so-called linkage phase (i.e., the fact that C3 and B3 lie on the same chromosome in the mother).

Recombinant DNA technology provided a source of polymorphic markers in the form of restriction fragment length polymorphisms (RFLPs). In 1980 my colleagues and I (Botstein et al., 1980) noted that if there is enough variation among the DNA sequences of humans, differences in pattern of digestion by sequence-specific endonucleases (the *restriction enzymes*) would be found from individual to individual. These differences in restriction fragment length could be used as codominant genetic markers, just as we used the

hypothetical alleles A1, etc. in Fig. 1. The method of detection proposed in 1980 (and still the most common) is to use single-copy DNA probes derived from human genomic clones as hybridization probes in gel-transfer experiments by the method of Southern (1985). Genomic DNA is extracted from white blood cells, cut with restriction enzymes, separated by size by gel electrophoresis, "blotted" onto filter paper, hybridized with labeled (radioactive or fluorescent) DNA probe, and analyzed in comparison with other samples from the same family. It should be emphasized that the DNA probe used to elicit the RFLP has a dual nature: it is a *genetic marker* that can be placed on a *genetic map* by *linkage* via polymorphism it reveals; it is also a *physical marker* that can be placed on a *physical map* because it is a single-copy DNA sequence. RFLPs thus bind together the genetic and physical maps of the human genome.

Of course, real RFLP markers are never completely polymorphic, and thus not every mating is "informative," in the sense of allowing the distinction between parental alleles. For a given locus, the inheritance of the marker could be followed in only a fraction of the families under study; in the remainder the marker is homozygous in key individuals, and thus yields no information. For this reason, we emphasized the importance of the "informativeness" of markers for mapping, proposing a measure of usefulness we called the *polymorphism information content* (PIC). The problem of informativeness of markers made us also propose that a map consisting mainly of polymorphic markers be constructed, so that even if a marker near a particular region turned out to be uninformative, the next marker on the map could be used in its stead, albeit with some loss of resolution. We proposed in 1980 that standard likelihood measures ("LOD scores") would serve admirably for disease mapping as well as for the construction of RFLP linkage maps.

DISEASE MAPPING WITH RFLP MARKERS

In 1983, the first autosomal disease gene was mapped using RFLPs: Gusella et al. (1983) found that the gene causing Huntington's disease is linked to a RFLP marker located near the end of the short arm of chromosome IV. Since then, a large number of other disease genes, including notably the gene for the recessive disease cystic fibrosis, have been mapped (Table 1). Because the RFLP markers are useful in physical as well as genetic mapping, they are crucial in the isolation of actual DNA sequences corresponding to the disease gene; the particular recent success in isolating the gene that causes cystic fibrosis was an example of this use of RFLP markers (Rommens et al., 1989; Kerem et al., 1989).

Despite the success in mapping some of the major inherited disease genes by linkage mapping using individual markers, theoretical analysis shows that

TABLE 1. *Diseases mapped using RLFPs*

Disease	Chromosome	Reference
Duchenne's muscular dystrophy	X	Davies et al., 1983
Huntington's disease	4	Gusella et al., 1983
Retinoblastoma	13	Cavenee et al., 1983
Cystic fibrosis	7	Tsui et al., 1985; Knowlton et al., 1985; White et al., 1985
Adult polycystic kidney disease	16	Reeders et al., 1985
Familial colon cancer	5	Bodmer et al., 1987
von Recklinghausen's neurofibromatosis	17	Barker et al., 1987; Seizinger et al., 1987
Bilateral acoustic neurofibromatosis	22	Rouleau et al., 1987
Multiple endocrine neoplasia type 2A	10	Simpson et al., 1987

use of a set of mapped markers whose linkage relationships are known and that span all of the genome will be more efficient and powerful (Lander and Botstein, 1986a). In large part, this is because of the increased power of linkage tests with markers flanking a disease gene as compared with single markers lying to one side. With single markers, single crossover events can completely reverse the relationship between the marker and the disease gene, but with flanking markers, only a double-crossover (an exceedingly rare event) will suffice to switch the marker-disease gene relationship completely.

Several techniques that take advantage of a complete linkage map have been suggested that allow either disease gene mapping without family study (*homozygosity mapping*; Lander and Botstein, 1987), mapping of diseases showing heterogeneity of cause (*simultaneous search*; Lander and Botstein, 1986b) or even mapping of genes contributing to quantitative traits (Lander and Botstein, 1988). This last method, which is applicable only to model systems, has been used to map several traits specified by as many as five contributing quantitative trait loci in tomato (Paterson et al., 1988).

In 1987, the first reasonably complete genetic linkage map of the human genome was published by Donis-Keller et al. This map has been preceded by maps of individual chromosomes and chromosome arms, notably the X chromosome (Drayna and White, 1985). The Donis-Keller et al. map covers, in the sense of showing continuous linkage, about 95% of the genome. However, not all the markers in it are highly informative and routine use as suggested above is not quite yet a reality. Improvement of the map, both with respect to informativeness of markers and density of markers, is a short-term goal of the Human Genome Initiative. As the maps become better,