Program Advisory Committee on the Human Genome

FIRST MEETING

January 3-4, 1989

OFFICE OF EXTRAMURAL RESEARCH
1-3-89
GLOSSARY OF COMMON NIH AWARDS

Grant - a financial assistance award to an institution, for profit or nonprofit, on behalf of a principal investigator to facilitate pursuit of a scientific focus or objective in the area of the investigator's interest or competence. The applicant provides the concept and scientific approach for its study.

Contract - an acquisition award for a research and/or development project to an institution, for profit or nonprofit, for a specific scientific inquiry directed towards a particular area of research and/or development needed by the U.S. Government. The contracting officer provides the specifications in a Request for Proposal and makes the award only after a contract is signed between the Government and the successful offeror.

Cooperative Agreement - a financial assistance award similar to a grant except that a cooperative agreement gives a substantive programmatic role to the Governmental agency making the award. However, the role of the Government is always subordinate to that of the grantee in design and execution of the project.

RO1 - a research project grant which supports a discrete, specified, circumscribed project to be performed by the named investigator(s) in an area representing his/her specific interest and competencies.

Program Project - a research project grant which supports a broadly based, multidisciplinary, often long-term research program which has a specific major objective or basic theme. It generally involves the organized efforts of relatively large groups, members of which are conducting research designed to elucidate the various aspects or components of this objective.

Center - often similar to a program project in that large, multidisciplinary groups are involved in the study. A specialized center is usually developed in response to an announcement of the programmatic needs of an NIH program and receives continuing attention from NIH staff. It may be funded as a grant or as a cooperative agreement. Centers are organized in a number of different patterns. Some Center grants support core research programs; a Comprehensive Center grant supports a multipurpose unit designed to bring together into a common focus divergent but related facilities within a given community.

OFFICE OF EXTRAMURAL RESEARCH
1-3-89
GLOSSARY OF COMMON NIH AWARDS

Grant - a financial assistance award to an institution, for profit or nonprofit, on behalf of a principal investigator to facilitate pursuit of a scientific focus or objective in the area of the investigator's interest or competence. The applicant provides the concept and scientific approach for its study.

Contract - an acquisition award for a research and/or development project to an institution, for profit or nonprofit, for a specific scientific inquiry directed towards a particular area of research and/or development needed by the U.S. Government. The contracting officer provides the specifications in a Request for Proposal and makes the award only after a contract is signed between the Government and the successful offeror.

Cooperative Agreement - a financial assistance award similar to a grant except that a cooperative agreement gives a substantive programmatic role to the Governmental agency making the award. However, the role of the Government is always subordinate to that of the grantee in design and execution of the project.

RO1 - a research project grant which supports a discrete, specified, circumscribed project to be performed by the named investigator(s) in an area representing his/her specific interest and competencies.

Program Project - a research project grant which supports a broadly based, multidisciplinary, often long-term research program which has a specific major objective or basic theme. It generally involves the organized efforts of relatively large groups, members of which are conducting research designed to elucidate the various aspects or components of this objective.

Center - often similar to a program project in that large, multidisciplinary groups are involved in the study. A specialized center is usually developed in response to an announcement of the programmatic needs of an NIH program and receives continuing attention from NIH staff. It may be funded as a grant or as a cooperative agreement. Centers are organized in a number of different patterns. Some Center grants support core research programs; a Comprehensive Center grant supports a multipurpose unit designed to bring together into a common focus divergent but related facilities within a given community.

TABLE OF CONTENTS

AGENDA
PROGRAM ADVISORY COMMITTEE ON THE HUMAN GENOME

January 3 and 4, 1989

Building 31, C Wing, Conference Room 6
National Institutes of Health
Bethesda, MD

Tuesday, January 3, 1989		*Morning Session*
8:30	Opening Remarks and Charge to the Committee—history and definition of NIH genome program	*Dr. Wyngaarden*
8:45	Chairman's Remarks—dates of future meetings	*Dr. Zinder*
9:00	The Human Genome Project at NIH—background, goal, reason for	*Dr. Watson*
9:20	Function of Office of Human Genome Research—organization and coordination with other groups	*Dr. Jordan*
9:30	Overview of Ongoing NIH Genome Activities	
	• National Institute of General Medical Sciences	*Dr. Kirschstein and staff*
	• National Library of Medicine	*Dr. Lindberg Dr. Masys and staff*
10:30	Coffee	
11:00	Key Genome—Related Resources Supported by NIH	
	• GenBank—National Institute of General Medical Sciences	*Dr. Cassatt*
	• DNA Probe Repository and Chromosome Library—National Institute of Child Health and Human Development and Division of Research Resources	*Dr. Dayton*

	• Protein Identification Resource (PIR)—Division of Research Resources	*Dr. Holloway*
	• Bionet—Division of Research Resources	*Dr. Holloway*
	• Cell Bank—National Institute of General Medical Sciences	*Dr. Greenberg*
12:15	General Discussion	
12:30	**Lunch**	

Afternoon Session

1:30	Overview of Genome Activities in Other Agencies	
	• U.S. Department of Energy—(including resources)	*Dr. Barnhardt*
	• Howard Hughes Medical Institute—(CEPH, HGML, OMIM)	*Dr. Cahill*
	• National Science Foundation	*Dr. Wooley*
	• U.S. Department of Agriculture	*Dr. Faust*
3:00	**Coffee**	
3:30	International Activities	
	• Human Genome Organization (HUGO)	*Dr. Watson* *Dr. McKusick*
	• Japan	*Dr. Olson*
	• United Kingdom	*Dr. M. Pearson*
	• European Economic Community (EEC)	*Dr. P. Pearson*
	• Other	*Participants*
4:30	General Discussion	
5:00	Adjourn First Day	

8:30 Strategy for NIH:

A. Scope

- Biological
- Technical
- Division of Labor (other organizations)

B. Program Management

- Centers and R01s
- Grants and Contracts
- Construction

C. Accessibility of Materials and Data

10:00 **Coffee**

10:30 C. Model Program

- Design
- Special Needs

D. Action Items

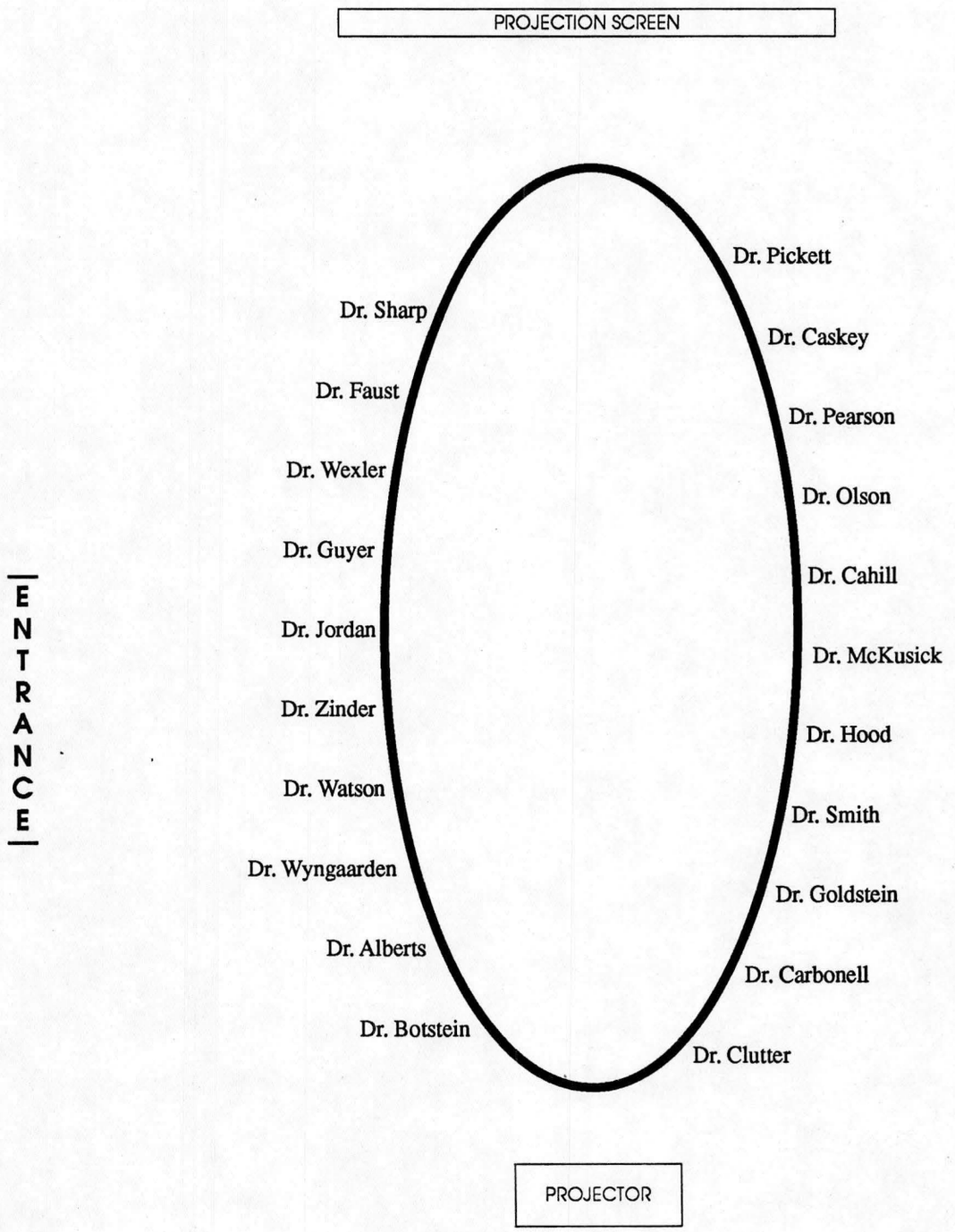
- Initiate Program
- Working Groups

12:00 Summary

12:30 Adjournment *Dr. Zinder*

Program Advisory Committee on the Human Genome

January 3 and 4, 1989
NIH, Building 31, C Wing, Conference Room 6



Roster

PROGRAM ADVISORY COMMITTEE
ON THE HUMAN GENOME

January 3 and 4, 1989

Building 31, C Wing, Conference Room 6
National Institutes of Health
Bethesda, MD

Chairman

Norton D. Zinder, Ph.D.
John D. Rockefeller, Jr. Professor
The Rockefeller University
1230 York Avenue
New York, NY 10021-6399
(212) 570-8644
(212) 570-7974 (FAX)

Executive Secretary

Elke Jordan, Ph.D.
Director
Office of Human Genome Research
National Institutes of Health
Building 1, Room 332
Bethesda, MD 20892
(301) 496-0844
(301) 496-4843 (FAX)

Members

Bruce M. Alberts, Ph.D.
Chairman
Department of Biochemistry
and Biophysics
University of California,
San Francisco
Box 0448
513 Parnassus Avenue, Room S-960
San Francisco, CA 94143
(415) 476-4132
(415) 476-0961 (FAX)

David Botstein, Ph.D.
Vice President-Science
Genentech, Inc.
460 Point San Bruno Boulevard
South San Francisco, CA 94080
(415) 266-2199
(415) 266-2739 (FAX)

Jaime G. Carbonell, Ph.D.
Associate Professor
Computer Science Department
Carnegie-Mellon University
Wean Hall, Room 4212
Pittsburgh, PA 15213
(412) 268-3064
(412) 268-5016 (FAX)

Joseph L. Goldstein, M.D.
Chairman
Department of Molecular Genetics
University of Texas
Southwestern Medical Center
5323 Harry Hines Boulevard
Dallas, TX 75235-9046
(214) 688-2141
(214) 688-8804 (FAX)

Leroy E. Hood, Ph.D.
Chairman
Division of Biology, 156-29
California Institute of Technology
1201 East California Boulevard
Pasadena, CA 91125
(818) 356-4951
(818) 449-0756 (FAX)

Victor A. McKusick, M.D.
University Professor
Division of Medical Genetics
Johns Hopkins Hospital
600 North Wolfe Street, Blalock 1007
Baltimore, MD 21205
(301) 955-6641
(301) 955-4999 (FAX)

Maynard V. Olson, Ph.D.
Professor
Department of Genetics
Washington University
School of Medicine
P.O. Box 8031
4566 Scott Avenue
Saint Louis, MO 63110
(314) 362-2721
(314) 362-4137 (FAX)

Mark L. Pearson, Ph.D.
Director, Molecular Biology
Central Research and
Development Department, E328/251
E.I. du Pont de Nemours
and Company
du Pont Experimental Station
P.O. Box 80328
Wilmington, DE 19880-0328
(302) 695-2140
(302) 695-4864 (FAX)

Cecil B. Pickett, Ph.D.
Executive Director of Research
Merck Frosst Centre for
Therapeutic Research
16711 Trans Canada Highway
Kirkland, PQ H9H 3L1
CANADA
(514) 630-2683
(514) 630-2624 (FAX)

Phillip A. Sharp, Ph.D.
Professor and Director
Center for Cancer Research
Massachusetts Institute
of Technology
40 Ames Street, Room E17-529B
Cambridge, MA 02139
(617) 253-6421
(617) 253-8000 or
(617) 258-8728 (FAX)

Nancy S. Wexler, Ph.D. (pending)
Associate Professor
Clinical Neuropsychology
Department of Neurology
and Psychiatry
College of Physicians and Surgeons
Columbia University
New York State Psychiatric Institute
722 West 168th Street
Box 58
New York, NY 10032
(212) 960-5650
(212) 960-5624 (FAX)

Liaison Members

George F. Cahill, Jr., M.D.
Vice President
Scientific Training and Development
Howard Hughes Medical Institute
6701 Rockledge Drive
Bethesda, MD 20817
(301) 571-0326
(301) 571-0573 (FAX)

C. Thomas Caskey, M.D., F.A.C.P.
Member, National Advisory
General Medical Sciences Council and
Professor and Director
Institute for Molecular Genetics
Baylor College of Medicine
One Baylor Plaza, T809
Houston, TX 77030
(713) 799-4773
(713) 799-6521 (FAX)

Mary E. Clutter, Ph.D.
Acting Assistant Director
Biological, Behavioral, and
Social Sciences
National Science Foundation
Room 506
1800 G Street, N.W.
Washington, DC 20550
(202) 357-9854
(202) 357-7346 (FAX)

Robert M. Faust, Ph.D.
National Program Leader
Crop Protection
National Program Staff
Agricultural Research Service
U.S. Department of Agriculture
Building 005, Room 236
BARC-West
Beltsville, MD 20705
(301) 344-3918
(301) 344-3191 (FAX)

David A. Smith, Ph.D.
Acting Director
Health Effects Research
U.S. Department of Energy
ER-GTN
Washington, DC 20545
(301) 353-3682
(301) 353-3884 (FAX)

Speakers

PROGRAM ADVISORY COMMITTEE ON THE HUMAN GENOME

January 3 and 4, 1989

Building 31, C Wing, Conference Room 6
National Institutes of Health
Bethesda, MD

Benjamin J. Barnhardt, Ph.D.
Biologist
Health Research Division
Office of Health and
Environmental Research
ER-72
U.S. Department of Energy, GTN
Washington, DC 20545
(301) 353-3683

James C. Cassatt, Ph.D.
Deputy Director
Biophysics and Physiological
Sciences Program
National Institute of General
Medical Sciences, NIH
Westwood Building, Room 907
Bethesda, MD 20892
(301) 496-7309

Delbert H. Dayton, M.D.
Chief
Genetics and Teratology Branch
Center for Research for Mothers
and Children
National Institute of Child Health
and Human Development, NIH
Executive Plaza North, Room 643B
6130 Executive Boulevard
Bethesda, MD 20852
(301) 496-5541

Judith Greenberg, Ph.D.
Director
Genetics Program
National Institute of General
Medical Sciences, NIH
Westwood Building, Room 910
Bethesda, MD 20892
(301) 496-7175

Caroline H. Holloway, Ph.D.
Acting Director
Biomedical Research
Technology Program
Division of Research Resources, NIH
Building 31, Room 5B41
Bethesda, MD 20892
(301) 496-5411

Ruth L. Kirschstein, M.D.
Director
National Institute of General
Medical Sciences, NIH
Westwood Building, Room 926
Bethesda, MD 20892
(301) 496-5231

Donald A.B. Lindbergh, M.D.
Director
National Library of Medicine, NIH
Building 38, Room 2E17B
Bethesda, MD 20892
(301) 496-4450

Daniel R. Masys, M.D.
Director
Lister Hill National Center for
Biomedical Communications
National Library of Medicine, NIH
Building 38A, Room 7N707
Bethesda, MD 20892
(301) 496-4441

Peter L. Pearson, Ph.D.
Professor
Department of Human Genetics
Sylvius Laboratoria
Wassenaarsweg 72
2333 AL Leiden
THE NETHERLANDS
071-276081 or -276101

James D. Watson, Ph.D.
Associate Director for
Human Genome Research
National Institutes of Health
Building 1, Room 332
Bethesda, MD 20892
(301) 496-0844

John C. Wooley, Ph.D.
Director
Division of Instrumentation
and Resources
National Science Foundation
1800 G Street, N.W., Room 504
Washington, DC 20550
(202) 357-9880

THE SECRETARY OF HEALTH AND HUMAN SERVICES
WASHINGTON, D.C. 20201

C H A R T E R

PROGRAM ADVISORY COMMITTEE ON THE HUMAN GENOME

Purpose

The Program Advisory Committee on the Human Genome will advise the NIH on all aspects of research in the area of genomic analysis. The Committee will identify opportunities to advance the ability of scientists to analyze the composition and organization of the genetic material of a number of organisms, with the goal of applying this information to the analysis of the human genome. The Committee will recommend initiatives that will promote the development of new technologies that will facilitate the acquisition, interpretation, analysis, and distribution of genetic and physical mapping information and deoxyribonucleic acid (DNA) sequence data. The Committee also will advise on research directions and identify areas of research requiring additional effort. The Committee will address the resource and training needs of the research community, as they pertain to genomic analysis.

Authority

42 U.S.C. 217a (Section 222 of Public Health Service Act as amended). This Committee is governed by provisions of P.L. 92-463, as amended (5 U.S.C. Appendix 2), which sets forth standards for the formation and use of advisory committees.

Function

The Program Advisory Committee on the Human Genome shall advise the Secretary; the Assistant Secretary for Health; the Director, National Institutes of Health; the Associate Director for Human Genome Research, National Institutes of Health; and the NIH Working Group on the Human Genome on long- and short-term planning to meet research needs for genomic analysis. Specifically, the Committee shall identify opportunities to further research on information and database technology and the methodology of genomic analysis and the characterization of the genomes of a variety of organisms, with the goal of applying this knowledge to the analysis of the human genome and ultimately to the prevention, diagnosis, and treatment of human disorders; recommend areas in which research should be stimulated; and suggest conferences, workshops, or other activities that the NIH should support to further the development of this research area.

Structure

The Program Advisory Committee on the Human Genome shall consist of 12 members selected by the Secretary, who shall be authorities knowledgeable in the fields of basic genetics, medical genetics, molecular biology, biochemistry, physical chemistry, information science, and engineering. The chair shall be selected by the Secretary from the membership and shall serve for at least one year and may be reappointed.

Members are invited to serve for overlapping four year terms, except that a member may serve after the expiration of the member's term until a successor has taken office. Terms of more than two years are contingent upon the renewal of the charter of the Committee by appropriate action prior to its expiration.

Management and support services shall be provided by the Office of the Associate Director for Human Genome Research, Office of the Director, NIH.

Meetings

Meetings shall be held at least twice a year at the call of the Chair with the advance approval of a Government official who will also approve the agenda. A Government official shall be present at all meetings. A quorum for the conduct of full committee business shall be seven.

Meetings shall be open to the public except as determined otherwise by the Secretary; notice of all meetings shall be given to the public.

Meetings shall be conducted, and records of the proceedings kept as required by applicable laws and departmental regulations.

Compensation

Members shall be paid at the rate of $200 per day for time spent at meetings, plus per diem and travel expenses as authorized by Section 5703, Title 5, United States Code, for persons in the Government service employed intermittently. Members who are officers or employees of the United States shall not receive compensation for service on the Committee.

Annual Cost Estimate

Estimated annual cost for operating the Committee, including compensation and travel expenses for members but excluding staff support, is $65,944. The estimated annual staff years of support is .45 at an estimated cost of $18,234.

Reports

An annual report shall be submitted to the Secretary; the Assistant Secretary for Health; and the Director, National Institutes of Health, which shall contain, as a minimum, the Committee's functions, a list of members and their business addresses, the dates and places of meetings, and a summary of the Committee's activities and recommendations during the year. A copy of the report shall be provided to the Department Committee Management Officer.

Termination Date

Unless renewed by appropriate action prior to its expiration, the Program Advisory Committee on the Human Genome shall terminate two years from the date of establishment.

APPROVED:

JUL 21 1988
Date

Otis R. Bowen M.D.
Otis R. Bowen, M.D.
Secretary

OFFICE OF HUMAN GENOME RESEARCH

Administrative Plan

In recognition of the high priority placed on mapping and sequencing the human genome, and the overarching planning and resource demands of a systematic targeted effort, it is proposed that an Office of Human Genome Research be established within the Office of the Director, NIH. The function of the Office will be to provide coordination, integration, review of progress, and planning in genomic analysis research. Research goals and long-range plans will be formulated with the guidance of the NIH Program Advisory Committee on the Human Genome and the NIH ~~Working Group on the~~ Human Genome Coordinating Committe

Coordination Function

Given the current broad involvement in research related to the characterization of complex genomes, the essential coordination and integration function of the Office will span four areas:

- o Overall intra-agency NIH coordination;
- o Interagency coordination between NIH and other Federal agencies (DOE and NSF), and other research-funding organizations;
- o Collaboration with industry and academia; and
- o International cooperation.

The Office of Human Genome Research is envisioned as a new entity with a mandate to develop proposals for analysis of complex genomes. This strategy is not intended to supersede ongoing efforts within other NIH components, but to integrate those efforts into a cohesive plan. One goal will be to maximize the efficiency of information exchange regarding new mapping data, improved techniques for storage and handling of biological materiel, and enhanced data processing and analysis. Therefore, centralized coordination will rely heavily on effective interactions with BID programs, as well as with other research funding organizations and the academic research community.

The NIH Program Advisory Committee on the Human Genome

The NIH Program Advisory Committee on the Human Genome will be comprised of non-Federal employees with demonstrated expertise in the scientific disciplines related to genomic analysis. The Committee will advise the NIH on all aspects of genomic analysis. The Committee will identify opportunities to further advance characterization of the genetic material of many organisms. The Committee will also recommend initiatives that should be undertaken to promote the development of new technologies that will lead to a deeper understanding of molecular biology. The Committee will also advise on research directions and identify areas of research requiring additional effort. The Committee will propose administrative solutions to the resource and training needs of the research community, specific to genomic analysis.

Of necessity, the membership of the NIH Program Advisory Committee on the Human Genome will represent a number of diverse research disciplines including, but not limited to, molecular genetics, physical chemistry, bioengineering, mathematics, and computer science.

The Office of Human Genome Research

The Office of Human Genome Research will serve as a focus within NIH and with other components of Public Health Service, reviewing policy questions, and coordinating plans for future research efforts. The Office will play an important role in exchanging information on the scientific activities in the Intramural Research Program. The Office will provide an internal framework for the review and consideration of a number of issues requiring the viewpoint of the biomedical research community.

Leadership for this initiative will be provided by the NIH Associate Director for Human Genome Research. This position will be held by a distinguished scientist who will be expected to remain current in his/her discipline. One means for recruiting such an individual would be to classify the position as part-time, offering the opportunity to maintain an ongoing research program. The Director of the Office of Human Genome Research will be responsible for day-to-day administrative operations in accordance with the guidance of the Associate Director.

The Office of Human Genome Research will develop a plan for a centralized, systematic, targeted effort to create detailed maps of the genomes of several organisms. The precise order and choice of goals would be determined with the advice of the Program Advisory Committee, but examples might include yeast, Drosophila, Caenorhabditis elegans, mouse, and human genomes.

The Office of Human Genome Research will not have a research budget to fund new initiatives. These initiatives will be supported by the BIDs and will pay particular attention to interdisciplinary projects that may not have a traditional locus in one BID. In addition, the coordination function of the Office will facilitate multiple Institute support for suitable proposals.

All grants and contracts funded as part of the genome program will be approved in accordance with traditional NIH peer review procedures.

MEMORANDUM OF UNDERSTANDING

BETWEEN THE

UNITED STATES DEPARTMENT OF ENERGY

AND THE

NATIONAL INSTITUTES OF HEALTH

TO COORDINATE RESEARCH AND TECHNICAL ACTIVITIES

RELATED TO THE HUMAN GENOME

I. Introduction

The National Institutes of Health (NIH), Department of Health and Human Services, and the United States Department of Energy (DOE) agree to foster interagency cooperation that will enhance the human genome research capabilities of both agencies.

DOE and NIH are the Federal Agencies primarily responsible for supporting research relating to the human genome. There has been considerable discussion in the scientific community over the past two years about the need for a coordinated long-term project to map and sequence the human genome. While NIH and DOE have informally coordinated such research efforts, the increasing complexity and scope of the project require a more formal mechanism. The purpose of this Memorandum of Understanding (MOU) is to provide for the formal coordination of the activities of DOE and NIH, and to provide for interfaces with relevant activities both within and outside the United States. The MOU also provides a mechanism by which NIH and DOE can jointly obtain outside advice regarding the human genome project.

II. Definition

For the purposes of this MOU, human genome research encompasses efforts to develop and apply technologies for the large-scale mapping, sequencing and analysis of the human genome. It includes the development of shared centralized facilities such as repositories for cloned DNA fragments, databases, and data centers to collect and distribute the large amounts of information generated on the project.

III. Goals

The goals of the project include: completion of a high-resolution genetic map of the human genome; completion of a series of complementary physical maps of increasing resolution; acquisition of a collection of ordered DNA clones encompassing the entire genome; determination of the complete nucleotide sequence of a reference genome; location of all the genes; and development of the tools to use the above information for a variety of biological and medical applications. Parallel studies in model organisms will be required in order to achieve a full understanding of the human genome.

IV. Management and Program Guidelines

A. Establishment of a joint advisory subcommittee chosen from the members of the DOE Health and Environmental Research Advisory Committee and the NIH Program Advisory Committee on the Human Genome.

The joint subcommittee will receive charges jointly prepared by NIH and DOE and communicated to their appropriate parent advisory committees. The joint subcommittee shall be co-chaired by representatives from the DOE and NIH committees. The joint subcommittee shall meet quarterly in order to advise and review the relevant activities of the two agencies. Subcommittee reports will be delivered through the two parent advisory committees to appropriate senior officials of NIH and DOE.

B. Establishment of an Interagency Working Group (IAWG) on genome research between DOE and NIH. The IAWG will be co-chaired by NIH and DOE and will meet at least on a quarterly basis to explore the need for and the feasibility of initiating a variety of cooperative and complementary programs and projects in order to advance knowledge in human genome research. The IAWG will also provide oversight of activities carried out under this MOU. In addition to the chairpersons, the IAWG will consist of an equal number of full members from DOE and NIH. Additional ad hoc members may be added for temporary assignments by either agency with prior concurrence of the chairpersons.

C. Continued coordination with other Federal agencies, with outside scientific groups, both national and international, and with private organizations involved in the genome project.

D. Continued joint participation and sponsorship of meetings and workshops for the purposes of planning and review of technical progress including an annual symposium to review progress in the science, to identify areas of need, and to address general policy questions.

E. Development of synchronous calendars for the agencies' research award cycles.

F. Concurrent funding and management of selected programs in human genome research that require utilization of unique NIH or DOE facilities.

G. Maintenance of regularly scheduled joint program staff meetings to exchange program information and plans.

H. Promotion of the sharing of technological advances and relevant biological materials (probes, cell lines, etc.) among investigators supported by both agencies. Assurance that relevant data are rapidly placed in appropriate databases and that relevant biological materials are rapidly placed in appropriate repositories.

I. Promotion of coordination and exchange of data with other countries.

J. Advance sharing of public policy statements relevant to human genome research.

V. Administration

A. Public Information Coordination: Subject to the Freedom of Information Act (5 U.S.C. 552), decisions on disclosure of information to the public regarding projects and programs implemented under the Memorandum of Understanding will be made following consultation between DOE and NIH respresentatives.

B. Intellectual Property: Specific provisions concerning the disposition of rights in intellectual property will be included in any interagency agreement under this Memorandum of Understanding.

C. Amendment and Termination: This Memorandum of Understanding may be modified or amended by written agreement between NIH and DOE and terminated by mutual agreement of DOE and NIH or by either party upon 90-day written notice to the other.

D. Effective Date: This Memorandum of Understanding is effective when signed by both parties.

James B. Wyngaarden
James B. Wyngaarden
Director
National Institutes of Health

Sept 30, 1988
Date

Robert O. Hunter, Jr.
Robert O. Hunter, Jr.
Director
Office of Energy Research
U. S. Department of Energy

October 7, 1988
Date

Genome Grants Paid in FY88

GENETIC MAPPING

Human

1 R29 GM41440-01
Boehnke, Michael L., Ph.D.
University of Michigan
Anne Arbor, MI

"Design Issues in Genetic Linkage and Segregation"
$104,207* / 5 years

1 R01 GM 40543-01
Donis-Keller, Helen R., Ph.D.
Collaborative Research, Inc.
Bedford, MA

"A Complete Genetic Linkage Map of the Human Genome"
$381,197 / 5 years

1 R29 GM 41253-01
George, Varghese T., Ph.D.
Louisiana State University
Medical Center
New Orleans, LA

"Estimation and Test of Marker Association in Family Data"
$68,713 / 5 years

1 R01 GM 39245-01A1
Jorde, Lynn B., Ph.D.
University of Utah
Salt Lake City, UT

"Linkage Disequilibrium in DNA Polymorphisms" / 3 years
$134,841

2 R01 GM 32793-04
Keats, Bronya J., Ph.D.
Louisiana State University
Medical Center
New Orleans, LA

"Maximum Likelihood Mapping of the Human Chromosomes"
$118,924 / 5 years

2 R01 GM 15883-24
Pious, Donald A., M.D., Ph.D.
University of Washington
Seattle, WA

"Genetic Structure and Function in Human Cells"
$215,765 / 5 years

1 R01 GM 39812-01A1
Risch, Neil L., Ph.D.
Yale University
New Haven, CT

"Statistical Methods and Applications in Human Genetics"
$288,313 / 5 years

1 R01 GM 40886-01
White, Raymond L., Ph.D.
University of Utah
Salt Lake City, UT

"Tools for High Resolution Linkage Mapping of Human Genes"
$1,499,741 / 5 years

---

*figures represent direct and indirect costs for the first years of support only; most of the grants will run from 3 to 5 years

Genome Grants Paid in FY88

Mammal

1 R01 GM 37550-01A1
Hardies, Stephen C., Ph.D.
University of Texas Health Science Center at San Antonio
San Antonio, TX

"Mammalian Repetitive Sequences as a Genetic Tool"
$98,547 / 3 years

1 R01 GM 39414-01
Nadeau, Joseph H., Ph.D.
Jackson Laboratory
Bar Harbor, ME

"Gene Mapping and Genome Organization in Mammals"
$228,353 / 5 years

9 R01 GM 40399-10
Passmore, Howard C., Ph.D.
Rutgers University
New Brunswick, NJ

"Recombination Analysis of the Mouse H-2 Complex"
$162,091 / 3 years

PHYSICAL MAPPING

Human

1 R01 GM 40873-01
Drabkin, Harry A., M.D.
University of Colorado Health Science Center
Denver, CO

"Construction of a Physical Map for Human Chromosome 3"
$177,098 / 3 years

1 R01 GM 40877-01
Fain, Pamela R., Ph.D.
University of Utah
Salt Lake City, UT

"One CentiMorgan Genetic Maps of Chromosomes X and 17"
$194,251 / 3 years

1 R01 GM 40865-01
Gardiner, Kathleen, Ph.D.
Eleanor Roosevelt Institute for Cancer Research
Denver, CO

"Cloning Large (>200 kb) Fragments of Chromosome 3 DNA"
$99,383 / 5 years

1 R01 GM 40860-01
Gemmil, Robert M., Ph.D.
Southwest Biomedical Research Institute
Scottsdale, AZ

"A Physical Map for Human Genomic Region 3p14.2 to 3p21.1"
$159,834 / 3 years

Genome Grants Paid in FY88

PHYSICAL MAPPING (con't)

Human (con't)

2 R01 GM 27882-09
Housman, David E., Ph.D.
Massachusetts Institute of Technology
Cambridge, MA

"Genetic Mapping and DNA Structure of Human Chromosome 11"
$390,683 / 5 years

1 R43 GM 40829-01
Lovett, Michael, Ph.D.
Genelabs Incorporated
Redwood City, CA

"Study of Linked Growth Factor Genes on Chromosome 5"
$50,000 / 1 year

1 P01 GM 41015-01
McKusick, Victor A., M.D.
Johns Hopkins University
Baltimore, MD

"Mapping the Chromosomes of Man"
$963,420 / 5 years

1 R01 GM 40864-01
Murray, Jeffrey C., M.D.
University of Iowa
Iowa City, IA

"Detailed Mapping and Recombination of Chromosome Four"
$428,127 / 5 years

1 R13 GM 40883-01
Ruddle, Frank H., Ph.D.
Yale University
New Haven, CT

"High Resolution Genetic Analysis of Complex Genomes"
$879,164 / 5 years

1 P01 GM 40606-01
Schlessinger, David, Ph.D.
Washington University
St. Louis, MO

"Human Genome Analysis With YAC Clones"
$723,160 / 3 years

1 R01 GM 40876-01
Shows, Thomas B., Ph.D.
Roswell Park Memorial Institute
Buffalo, NY

"Mapping Human Chromosome 11"
$404,391 / 5 years

1 R01 GM 40882-01
Westbrook, Carol A., M.D., Ph.D.
University of Chicago
Chicago, IL

"Macrorestriction Mapping Applied to Chromosome 5q"
$264,298 / 3 years

1 R01 GM 40878-01
Willard, Huntington F., Ph.D.
University of Toronto
Ontario, Canada

"Mapping Centromeric Regions of Human Chromosomes"
$153,724 / 3 years

Genome Grants Paid in FY88

Invertebrates

1 R01 GM 40869-01
Hartl, Daniel, Ph.D.
Washington University
St. Louis, Mo

"Drosophila Genome Mapping Using YAC Vectors"
$395,322 / 5 years

1 R01 GM 41422-01
Waterston, Robert H., Ph.D.
Washington University
St. Louis, MO

"A Physical Map of the C. Elegans Genome"
$195,659 / 5 years

Microorganisms

1 R01 GM 40594-01
Altschuler, Marsha I., Ph.D.
Williams College
Williamstown, MA

"Cloning and Analysis of Tetrahymena Somatic Chromosomes"
$142,262 / 5 years

2 R01 GM 35682-07
Daniels, Donna L., Ph.D.
University of Wisconsin
Madison, WI

"Complete Dissection of the Escherichia Coli Genome"
$136,542 / 3 years

1 R29 GM 39887-01A1
Hackett, Neil R., Ph.D.
Vanderbilt University
Nashville, TN

"Physical Mapping of an Archaebacterial Genome"
$111,406 / 5 years

1 R01 GM 40889-01
Olson, Maynard V., Ph.D.
Washington University
St. Louis, MO

"Refinement of the Yeast Physical Map"
$152,426 / 3 years

Plants

1 R01 GM 41010-01
Burr, Benjamin, Ph.D.
Brookhaven National Laboratory
Upton, NY

"Mapping and Characterizing Quantitative Trait Loci in Maize"
$135,502 / 5 years

SEQUENCE DETERMINATION

1 R01 GM 39598-01A1
Blattner, Frederick R., Ph.D.
University of Wisconsin
Madison, WI

"Determination of the Complete Sequence of E. coli"
$540,410 / 5 years

Genome Grants Paid in FY88

SEQUENCE DETERMINATION (con't)

1 R01 GM 41387-01
Dausset, Jean M.D.
Centre d'Etude du Polymorphisme Humain
Paris, France

"Primary Structure of the Human MHC"
$384,150 / 2 years

1 R01 GM 40867-01
Hood, Leroy E., M.D., Ph.D.
California Institute of Technology
Pasadena, CA

"Characterization of the Genomes of Humans and Model Organisms"
$501,934 / 5 years

TECHNOLOGY DEVELOPMENT/INSTRUMENTATION

2 R44 GM 37456-02
Anderson, Norman, Ph.D.
Large Scales Biology
Rockville, MD

"Camera/Scanner for Genomic Mapping and Protein Analysis"
$252,266 / 2 years

2 R01 GM 34960-04
Collins, Francis, M.D., Ph.D.
University of Michigan
Ann Arbor, MI

"Chromosome Mapping by a DNA Circularization Technique"
$99,629 / 5 years

2 R37 GM 21891-14
Davis, Ronald W., Ph.D.
Stanford University
Stanford, CA

"Cell Regulation: Biochemically Isolated DNA Segments"
$513,684 / 5 years

2 R01 GM 29848-07
Deininger, Prescott L., Ph.D.
Louisiana State University Science Center
New Orleans, LA

"Human Interspersed Repeated DNA Sequences"
$108,188 / 5 years

1 R01 GM 40880-01
Efstradiatis, Argiris, M.D., Ph.D.
Columbia University
New York, NY

"cDNA Probes for Exon Mapping of the Human Genome"
$322,887 / 5 years

1 R43 GM 40828-01
Goustin, Anton S., Ph.D.
CBR Laboratories, Inc.
Boston, MA

"Novel HLA-Linked DNA in a Yeast Artificial Chromosome"
$50,000 / 1 year

1 R43 GM 40794-01
Helentjaris, Timothy G., Ph.D.
National Plants Sciences
Salt Lake City, UT

"Novel Rapid RFLP Analysis With Fluorescence Detection"
$49,808 / 1 year

Genome Grants Paid in FY88

TECHNOLOGY DEVELOPMENT/INSTRUMENTATION (con't)

2 R44 GM 38941-02
Huse, William D., M.D., Ph.D.
Stratagene, Inc.
La Jolla, CA

"Novel Vectors for Gene Mapping and Sequencing"
$277,637 / 1 year

1 R01 GM 40554-01
Kolodner, Richard D., Ph.D.
Dana Farber Cancer Institute
Boston, MA

"Enzymology of Mismatch Repair in Yeast"
$190,637 / 5 years

2 R37 GM 35095-04
Lerman, Leonard S., Ph.D.
Massachusetts Institute of Technology
Cambridge, MA

"Strategy for the Characterization of the Human Genome"
$328,567 / 5 years

1 R43 GM 40887-01
Lovett, Michael, Ph.D.
Genelabs Incorporated
Redwood City, CA

"New Methods for Producing Chromosome-Specific Libraries"
$252,054 / 1 year

2 R01 GM 35635-09
McDevitt, Hugh O., M.D.
Stanford University
Stanford, CA

"Genetics of HLA-D and T-Cell Receptor Haplotypes"
$159,372 / 4 years

1 R43 GM 41559-01
McGann, William, Ph.D.
Radiation Monitoring Devices, Inc.
Watertown, MA

"Improved RFLP Analysis for Genetic Linkage Mapping"
$50,000 / 1 year

1 R01 GM 40699-01
Moir, Donald T., Ph.D.
Collaborative Research, Inc.
Bedford, MA

"Molecular Cloning and Analysis of Human Genomic Regions"
$272,376 / 3 years

1 R01 GM 40719-01
Orias,Eduardo, Ph.D.
University of California
Santa Barbara, CA

"Molecular Genetics of a Cell Cell Recognition System"
$157,115 / 3 years

1 R01 GM 40896-01
Rinchik, Eugene M., Ph.D.
Martin Marietta Energy Systems, Inc.
Oak Ridge, TN

"Saturation Mutagenesis of Mouse Genomic Regions"
$183,538 / 5 years

1 R01 GM 40537-01
Roberts, Richard J., Ph.D.
Cold Spring Harbor Laboratory
Cold Spring Harbor, NY

"A Search for New Restriction Endonucleases"
$213,026 / 4 years

Genome Grants Paid in FY88

TECHNOLOGY DEVELOPMENT/INSTRUMENTATION (con't)

1 R01 GM 40881-01
Sorge, Joseph A., M.D.
Stratagene, Inc.
La Jolla, CA

"Technologies Aimed at a Physical Map and Sequence of the Human Genome"
$269,720 / 3 years

1 R01 GM 40936-01
Sutherland, John C., Ph.D.
Brookhaven National Laboratory
Upton, NY

"New Approaches to DNA Mapping: Electronic Imaging"
$155,635 / 5 years

1 R01 GM 40885-01
Warburton, Dorothy P., Ph.D.
Columbia University
New York, NY

"Monochromosomal Hybrids by Retroviral Marker Transfers"
$137,846 / 3 years

1 R01 GM 40633-01
Ward, David C., M.D.
Yale University
New Haven, CT

"Affinity Purification of Large Fragments of Human DNA"
$222,945 / 5 years

1 R43 GM 40256-01
Watkins, Paul C.
Intergrated Genetics, Inc.
Franmingham, MA

"Automated Large Scale Analysis of Human Cosmid Libraries"
$50,000 / 1 year

2 R01 GM 30140-07
Weaver, Steven G., Ph.D.
University of Illinois at Chicago
Chicago, IL

"The Maintenance of Sequence Homology by Gene Conversion"
$224,770 / 5 years

INFORMATICS/COMPUTATIONAL

1 R01 GM 39907-01
Karlin, Samuel, Ph.D.
Stanford University
Stanford, CA

"Analysis of Molecular Sequence Data"
$367,851 / 5 years

1 R01 GM 40789-01
Lapedes, Alan S., Ph.D.
Los Alamos National Laboratory
Los Alamos, NM

"Genetic Databases: Applications for Machine Learning"
$350,770 / 3 years

2 R44 GM 36180-02
Waleh, Ahmad, Ph.D.
Applied Sciences Consultants
San Jose, CA

"Computer Folding of RNA Using Monte Carlo Methods"
$135,934 / 2 years

Genome Grants Paid in FY88

CONFERENCES

1 R13 GM 39692-01
Fox, Fred C., Ph.D.
University of California
Los Angeles, CA

"Gene Transfer in Animals"
$22,000 / 1 year

1 R13 GM 41447-01
Ruddle, Frank H., Ph.D.
Yale University
New Haven, CT

"Tenth International Workshop on Human Gene Mapping"
$879,164 / 1 years

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 8, 1988

From Huber Warner, Ph.D., Deputy Associate Director, BRCM, NIA

HRW

Subject NIA Activities Related to the Human Genome Project

To Elke Jordan, Ph.D, Director, Office of Human Genome Research

I have attached a partial list of NIA-supported projects related to the Human Genome Project. Only the first two are directly related to this Project, but neither has been funded by funds allocated for the Project. NIGMS has recently made an award to Dr. Davis to fund this research, so the NIA support is now focussed on application of these techniques to specific genes. I have also attached a copy of a response to a Congressional question about NIA involvement in this Project.

Attachments

N. Robakis - R01 AG08200 "Structure and Expression of the beta-Protein Precursor Gene" Determination of the sequence of the gene for the precursor of human beta-amyloid protein.

G. Burmer - R29 AG07359 "Cloning of Werner's Syndrome Defect' Identification and characterization of gene(s) able to overcome the defect in Werner's syndrome.

M. Weksler - P01 AG00541 "Immunobiology of Aging" Cloning and sequencing of human genes for prothymosin alpha and rat gene for parathymosin.

J. Smith - P01 AG07123 "Molecular Approaches to the Study of Cellular Aging" Cloning and sequencing the gene(s) for an inhibitor of cell proliferation.

I. Statement of Issue

How is the NIA involved in the initiative to sequence the human genome?

II. Discussion of the Issue

The Congress of the United States has recently recognized the scientific potential afforded by the knowledge of the complete sequence of the human genome, and is appropriating funds specifically for this purpose; most of these funds will be spent by DOE and NIH, with the majority going to the NIH. The staff of the National Institute of General Medical Sciences has recently published a program announcement on "Mapping and DNA Sequence Determination of the Genome of the Human and Model Organisms". Responsive proposals will focus on developing the technology to expand genetic and physical maps of the human genome and improve techniques for determining the sequence of the DNA; research on the application of this genetic information to the diagnosis, prevention or treatment of specific genetic disorders is not within the scope of this project. While the NIA will therefore not play a major role in the human genome initiative, the NIA does support research highly relevant to this initiative. The NIA supports two projects to develop technology for isolation of large pieces of DNA, which is crucial for physical mapping studies. The NIA provides extensive support to map and sequence chromosome 21, because of its potential relevance to Alzheimer's disease. The NIA also supports research on the DNA sequences coding for various immunoglobulin polypeptide chains, mouse myogenic determination factor, statin (a protein which appears in quiescent and senescent cells), and the precursor to the beta-amyloid protein.

III. Alternatives and/or Suggested Position

The NIA will continue to support research on sequencing mammalian genomes where relevant to the mandate of the NIA, and when appropriate will request that this research be supported by the funds allocated by Congress to support the human genome initiative. Currently, the NIA receives none of these funds, as most of these have been allocated for proposals in response to RFAs on physical mapping of genomes and technology development.

Genome-Related Projects Supported by the National Institute of Allergy and Infectious Diseases

The NIAID supports investigator initiated research projects designed to determine the sequence of specific segments as well as the entire genome of many different bacterial, viral, fungal and parasitic infectious organisms. These organisms include: HIV, vaccinia virus, rabies virus, hepatitis A virus, hepatitis delta virus, Treponema pallidum, Bordetella pertussis, Mycobacterium tuberculosis, and Neisseria gonorrhea. The cloning and sequencing of the genomes of pathogenic agents should provide valuable information leading to the development of improved and more sensitive clinical diagnostic procedures and aid in the production of more effective and safe vaccines.

Support is also provided for basic studies on the molecular genetics of the immune system. Investigators supported by the NIAID are involved in the cloning and sequencing of both rodent and human genes encoding many immune related proteins including: the major histocompatibility antigens (HLA and H-2), immunoglobulins, the antigen specific T cell receptors, the interleukins and their cell surface receptors, the lymphocyte adhesion molecules and components of the complement cascade. These studies should provide important insights into the molecular events controlling susceptibility to infections and autoimmune diseases and should also aid in the development of immunotherapeutic modalities for the modulation and manipulation of the human immune system in health and in disease.

Selected Genome Mapping Activities Supported by the National Cancer Institute

The National Cancer Institute (NCI) is involved in a variety of activities related to mapping of the human and other complex genomes. These activities are being carried out by intramural scientists as well as being supported through the extramural scientific programs. The following brief review is not all inclusive but rather has been developed to present the range and types of projects currently supported by the NCI.

The studies supported in general range from relatively limited projects that seek to locate specific genes involved in malignant transformation to physical and genetic mapping of larger regions. The NCI supports many projects involving cloning and sequencing of particular genes but in general, these have not been included in this summary since they deal with genes that have already been mapped and therefore do not directly contribute new information useful for development of a physical or genetic map. Descriptions of some of the mapping projects specifically related to the human genome follow:

Linkage studies of multiple endocrine neoplasia types 1 and 2 (MEN 1, MEN 2) have resulted in mapping these diseases to chromosomes 11 and 10, respectively. Continuing efforts are aimed at cloning the specific genes.

Efforts to understand the relationship between dysplastic nevi and malignant melanoma have led to mapping of a gene involved in a familial form of melanoma to chromosome 1.

Many studies have analyzed families of neurofibromatosis patients and have resulted in mapping genes for at least two forms of this disease.

Considerable effort is being devoted to locating the gene for neuroblastoma and some data support a chromosome 1 location. There is a suggestion that a suppressor gene may be involved.

Chromosome 3 is the subject of major mapping efforts because it appears to be involved in a number of cancers. In addition, there is a common fragile site on 3p that may be involved in translocations associated with specific cancers.

Genes associated with leukemias and lymphomas are continuing to receive attention. Analyses of rearrangements are yielding information about the gene structure and function of immunoglobulin genes and immune cell receptor genes.

Specific translocations identified in particular solid tumors are providing starting points for mapping efforts on the chromosomes involved in the translocations, e.g., t(11;22) in Ewing's sarcoma, t(3;8) in renal cell carcinoma. Other chromosomal rearrangements are being studied as well.

Major efforts are underway to map the major histocompatibility complex in humans.

The Li-Fraumeni syndrome (sarcoma-breast cancer syndrome) is being studied with samples being collected from family members for gene linkage analysis.

Further studies on the frequency of cancers secondary to retinoblastoma are proceeding in order to gain a better understanding of the carcinogenetic mechanisms. Specimens for these secondary tumors are being collected so that the gene structure around the retinoblastoma gene can be evaluated.

The studies described represent the work of many groups, both intramural and extramural. About 80 grants are involved and several intramural laboratories.

National Institute of Child Health and Human Development

Research Activities Relating to Mapping the Human Genome

Mission

The mission of the National Institute of Child Health and Human Development is to help families have healthy children at the time they are wanted, to prevent disease and disability among children, to foster normal development early in life, and to ensure that every child has the opportunity to fulfill his or her potential for a healthy and productive adulthood. In pursuit of its mission, the NICHD supports programs focused on the reproductive, developmental, and behavioral processes that determine the health of children, adults, families, and populations. By increasing our knowledge in these areas, the Institute is contributing to a healthier, more productive life for all. Because of its interest in human development and abnormalities (such as mental retardation) that interfere with and inhibit the full development of human potential, understanding basic genetic mechanisms is an important part of all NICHD research programs. The Institute's interest and emphasis on research related to mapping the human genome follows from this fact.

Activity

NICHD support of research related to mapping the human genome exceeded $15.0 million in fiscal year 1988. Research projects were supported in the intramural and extramural programs of the Institute. Examples of extramural research supported by the Institute included:

1. Mapping the long arm of chromosome 21 in order to better understand the etiology of Down syndrome.

2. Mapping genes associated with reproduction.

3. Mapping genes associated with sex differentiation.

4. Mapping sex-specific genes.

In addition to specific extramural research projects, the NICHD established a repository of human DNA probes and chromosome-specific libraries in 1985. It is a centralized national and international resource providing researchers with a reliable and efficient means to exchange cloned human DNA. In its first three years, 619 probes have been received, expanded, verified and stored in multiple samples. Two hundred forty-three probes have been characterized, approved and authenticated by the contributor and are ready for distribution. Other probes are nearing the requirements needed to allow distribution. Sixty-seven chromosome-specific libraries are available to scientists upon request. Quality control is maintained and the probes deposited in this facility emphasize relevancy to human genetic disease. Representative chromosome-specific genes/probes are being acquired to span each individual chromosome and the probes represent

important genes, polymorphisms, diseases, and significant chromosomal locations for genetic linkage analysis. Since 1985, 1,473 clones and 732 libraries have been distributed. To date, there are 1,813 registered users of this service. Each human DNA probe and library is accompanied by a computerized data base allowing multi-user access to descriptive and bibliographic information. Information stored on genomic and cDNA libraries includes history of construction, biochemical descriptions, and comments recorded by those who have worked with the libraries. Information on cloned genes or arbitrary fragments include map location, notes on sequence content, polymorphism, linkage, haplotype descriptions, cross-references to other data bases and literature references.

The NICHD intramural program houses many investigators actively involved in mapping human genes and identifying their function. In 1990, a Human Genome Mapping Section will be developed in the NICHD Laboratory of Molecular Genetics.

National Institute of Dental Research

NIDR genome-related activities fall into several categories. One concerns the synthesis and expression of extracellular matrix proteins such as those found in basement membrane and in bone, cartilage, and tooth matrices. These matrix molecules are essential to the formation of normal bone, cartilage, tooth enamel and dentin, and it is assumed that defects in one or more of the genes controlling the production of these matrix molecules lead to such heritable conditions as osteogenesis imperfecta and dentinogenesis imperfecta. In the case of basement membrane, key proteins are involved in embryogenesis, cell migration and differentiation, tissue repair and regeneration, and in cell attachment and metastatic proceses.

A second area of gene research concerns the identification of plasmid bacterial genes that can facilitate development of useful recombinants.

A third area involves identification of the genes controlling the production of virulence factors, attachment proteins, receptors, and other elements of bacteria and viruses which contribute to pathologic processes. This research permits the construction of mutant forms of bacteria and the development of recombinant vaccines using genes to express protein immunogens as well as pure synthetic peptide vaccines (e.g., experimental herpesvirus vaccines).

Other disease-related genetic research has led to the isolation of genes coding for autoantigens in autoimmune diseases and to research on the gene processes involved in retroviral-induced expression of cellular oncogenes.

The following table provides a sample of the kinds of genome-related activities conducted and supported by NIDR:

NIDR Genome-related Activities--Representative studies

	Status
Matrix Proteins	
Basement Membrane	
Laminin	Gene cloned and sequenced in humans and mice.
Laminin receptor	Gene cloned and sequenced in humans and mice.
Type IV collagen	Gene cloned and sequenced; mapped to chromosome 13.
Heparan sulfate proteoglycan	Gene cloned and sequenced in mice.
Nidogen	Gene cloned and sequenced in humans and mice.
Cartilage	
Cartilage Proteoglycan	Gene cloned and sequenced in human and in rat.
Bone	
Bone proteoglycans: Decorin, Biglycanin; Bone sialoproteins: Osteopontin, bone sialoprotein	cDNAs constructed and sequenced; genes mapped to particular chromosomes.
Osteonectin	Genomic sequencing in progress (human).
Tooth	
Amelogenin	Genomic sequencing in progress in human; bovine.
Enamelin	cDNAs constructed in bovine.
Bacterial Genes	
Lactobacillus casei	The 3 genes for the lactose P.T.S. operon and the gene for beta-galactosidase have been cloned and sequenced. These genes are on plasmids.

Actinomyces vicosus	Genes for the type 1 and type 2 fimbriae have been cloned and sequenced.
Actinomyces naeslundi	Type 2 fimbriae cloned and sequenced.
Streptococcus mutans	Genes for virulence factors have been cloned and sequenced.
Streptococcus sanguis) Streptococcus mutans)	A gene coding for a receptor found on the cell surface of these two species of bacteria has been cloned. The receptor is responsible for binding to the agglutinin found in human saliva.

Autoantigens

A cDNA clone has been sequenced for a 70 kD novel thyroid auto-antigen associated with Graves disease. The active gene has been localized to chromosome 22 and genomic sequencing is under way.

An almost full-length cDNA clone has been identified in an existing library. It codes for an autoantigen associated with insulin-dependent diabetes mellitus.

GENOME-RELATED RESEARCH OF THE

NATIONAL INSTITUTE OF DIABETES AND DIGESTIVE AND KIDNEY DISEASES

The National Institute of Diabetes and Digestive and Kidney Disease (NIDDK) supports a wide range of genome-related research projects. The following are highlights of recent advances emanating from NIDDK extramural and intramural research:

Cystic Fibrosis: NIDDK grantees are taking part in the international quest for the cystic fibrosis gene, identifying new markers that are extremely close to the gene. They have also shown that there is a defect in the regulation of membrane channels for chloride, an ion that controls many cellular activities by changing the electrical potential of the cell membrane. NIDDK-supported researchers have sequenced a chloride ion transport protein, now used to study chloride ion transport. Simultaneous pursuit of both avenues of research should enable scientists to target therapies to the defect and identify similar genes that may also be candidates for a cystic fibrosis gene.

Gene Map: The availability of a complete linkage map of the human genome would greatly amplify the power of molecular biologists to find critical genes. NIDDK researchers recently reported the construction of one such human genome map. Using DNA from 21 three-generation families, they found 403 chromosomal locations that were recognized by 393 enzymes that cut DNA. With this system, they reassembled the DNA fragments to generate a complete map. This map is linked to about 95 percent of the human genome, representing progress toward developing a more highly refined map of the human genome.

Human Protein Production in Animal Milk: A long-standing collaboration between NIDDK, the National Institute of Child Health and Human Development and private industry resulted in a new method to produce a human protein, tissue plasminogen activator (tPA), in mouse milk. A blood-borne protein, tPA is also used therapeutically to treat heart and blood disorders. Researchers created transgenic animals that contained the tPA gene in their DNA. Because the gene was regulated by the promoter for a milk protein gene, the animals secreted tPA during normal milk production. In a recent collaboration with the United States Department of Agriculture, these researchers have succeeded in producing transgenic pigs to see if newly introduced genes can alter the composition of their milk. Eventually, larger animals may be used to generate an inexpensive source of large quantities of therapeutic proteins, which can be extracted easily from the animals' milk.

Extreme Insulin Resistance: NIDDK intramural researchers have identified multiple receptor defects in patients with genetic forms of extreme insulin resistance. This research is yielding important insights into the normal activity of the insulin receptor and about defects that could cause diabetes. Using a clone of the insulin receptor gene, these investigators found three receptor defects that curtail receptor activity. Because these mutations lead to insulin resistance, a characteristic of noninsulin-dependent diabetes, these studies are being extended to see whether mutations in the insulin receptor gene underlie that disease. NIDDK scientists also used

state-of-the-art mapping techniques of inbred families to find the mutations and are now searching for other genes that may play roles in extreme insulin resistance.

Insulin-dependent Diabetes: Much insight into genetic susceptibility to insulin-dependent diabetes has been gleaned as a result of NIDDK-supported development of two animal models: the non-obese diabetic mouse and the Bio-Breeding rat. Data from animals, and data now emerging from patients, is shedding light on the multi-gene defect that is believed to lead to the development of this disease. The genetic flaw culminates in a destructive autoimmune process that destroys insulin-producing pancreatic cells. These studies showed that 95 per cent of people who develop insulin-dependent diabetes have specific genes associated with the immune response complex. Recently, they learned that other parts of this genetic complex may either coincide with or confer resistance to insulin-dependent diabetes.

Tay-Sachs Disease: A lethal metabolic disorder, Tay-Sachs disease results from the lack of an enzyme that leads to the accumulation of metabolic waste products in cellular compartments, called lysosomes. Research into the defective enzyme in Tay-Sachs disease has become a model for understanding lysosomal enzymes. Much information about the Tay-Sachs mutation stems from studies by intramural researchers, who cloned genes for the missing enzyme. The isolation of the clone for the gene whose defect causes Tay-Sachs disease has enabled these scientists to pinpoint the genetic defects underlying the disease. Using the clones, they showed different genetic defects in the two populations--Ashkenazi Jews and French Canadians--whose carrier rate is 10 times greater than that of the general population. Whereas French Canadian patients had a large gene deletion, very recent studies imply that many smaller mutations underlie the disease in the Ashkenazi Jewish population.

Metabolic Diseases: NIDDK extramural and intramural research is revealing the genetic mutations underlying such metabolic diseases as Lesch-Nyhan disease, Fabry's disease, and isovaleric acidemia. These devastating diseases stem from a genetic defect, leading either to the production of a faulty enzyme or to its total absence. Over the past twenty years, Institute-supported scientists have obtained a wealth of information about isovaleric acidemia. They developed a method to detect the acid that accumulates in the blood of affected infants and characterized the enzyme that is defective in this disease. By cloning the gene for the enzyme, these researchers discovered many mutations that bring about this disease. They also cloned the gene for methylmalonic acidemia, a related disease. These studies are enhancing the potential for better diagnosis and treatment of such deadly metabolic diseases.

Polycystic Kidney Disease: Polycystic kidney disease is actually a group of genetic diseases, one group inherited via a dominant mode, and the other by a recessive mode. Institute-supported scientists have recently found the approximate location of the gene for the dominant form on chromosome 16. Prior to very recent NIDDK-sponsored studies, however, the location of the recessive gene was unknown. Using linkage analysis, scientists discovered that the gene responsible for the recessive form is located in the same area of chromosome 16, but appears to arise from a mutation in another gene. These studies will enable researchers to diagnose these diseases more effectively and develop treatments once the genes are identified.

Congenital Adrenal Hyperplasia: Congenital adrenal hyperplasia, a fairly common genetic defect, leads to defective production of cortisol, a vital stress hormone. Patients produce excessive amounts of the male hormone androgen, resulting in ambiguous or abnormal genitalia in affected females. Two-thirds of these patients are also unable to synthesize another hormone and, if untreated, die shortly after birth. Using a gene clone for one of the enzymes needed for cortisol synthesis, scientists discerned multiple mutations in patients, including deletions in the gene and the substitution of another, nonfunctional gene for the normal one. Very recent studies disclosed a genetic marker that seems to be universal for the presence of these mutations. This marker should greatly improve the diagnosis of congenital adrenal hyperplasia early in fetal development, of pivotal importance for prenatal treatment of this disease.

HHS NEWS

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOR RELEASE
Thursday, July 21, 1988

PUBLIC HEALTH SERVICE
Ann Dieffenbach
(301) 496-7301

The first research grants awarded under a special gene mapping initiative of the National Institute of General Medical Sciences started on July 1. This initiative is supported by a $17.2 million Congressional appropriation to NIGMS in fiscal year 1988.

Gene mapping, the process of pinpointing the specific locations of genes on chromosomes, enables scientists to learn more about genes involved in inherited disorders and may lead to new means of diagnosing, treating and preventing such disorders. Knowing the locations of genes also provides a wealth of information on the genetic makeup of all humans.

The new awards are a component of the National Institutes of Health's efforts to characterize the genomes (the complete genetic endowment) of humans and model organisms such as yeast, fruit flies and mice. In addition to gene mapping, this endeavor involves the development of new tools for and approaches to genome analysis, as well as the determination of the sequence, or order, of subunits of the genetic material DNA.

The current efforts to characterize complex genomes are an outgrowth of studies in the underlying fields of molecular genetics and gene expression that have been supported by NIH, and in particular by NIGMS, for more than 20 years. This research has already led to greatly improved strategies for studying human genetics and inherited disorders. While most research done in the past has focused on locating specific genes of interest, the new

(More)

initiative will support a more systematic approach that involves mapping all of an organism's genes.

The gene mapping initiative will be facilitated by several NIH-supported research resources. These include a genetic sequence data bank; a repository of cells from people with genetic diseases; repositories of specific segments of DNA; and computer hardware and software programs that enhance communication and data exchange among biomedical researchers.

A list of the grants awarded on July 1 is attached.

# # # #

NIGMS GENE MAPPING GRANTS AWARDED JULY 1, 1988

Frederick R. Blattner, Ph.D. University of Wisconsin, Madison Madison, WI	"Determination of the Complete Sequence of E. Coli" $540,410*
Helen R. Donis-Keller, Ph.D. Collaborative Research, Inc. Bedford, MA	"A Complete Genetic Linkage Map of the Human Genome" $381,197
David E. Housman, Ph.D. Massachusetts Institute of Technology Cambridge, MA	"Genetic Mapping and DNA Structure of Human Chromosome 11" $390,683
Richard D. Kolodner, Ph.D. Dana-Farber Cancer Institute Boston, MA	"Enzymology of Mismatch Repair in Yeast" $190,637
Alan S. Lapedes, Ph.D. Los Alamos National Laboratory Los Alamos, NM	"Genetic Databases: Applications for Machine Learning" $350,770
Richard J. Roberts, Ph.D. Cold Spring Harbor Laboratory Cold Spring Harbor, NY	"A Search for New Restriction Endonucleases" $213,026
David Schlessinger, Ph.D. Washington University St. Louis, MO	"Human Genome Analysis With YAC Clones" $723,160
David C. Ward, M.D. Yale University New Haven, CT	"Affinity Purification of Large Fragments of Human DNA" $222,945

*figures represent direct and indirect costs for the first year of support only; the grants will run from 3 to 5 years

# # # #

HHS NEWS

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOR RELEASE
Friday, October 21, 1988

Public Health Service
Ann Dieffenbach
(301) 496-7301

The National Institute of General Medical Sciences recently awarded 55 research grants under a gene mapping and genome analysis initiative begun in fiscal year 1988 with special funds from Congress. The first-year costs of these grants exceed $13.5 million, bringing the total spent on this initiative in fiscal year 1988 to $17.2 million.

Gene mapping, the process of determining the locations of genes on chromosomes, helps scientists understand inherited disorders and may lead to new ways to diagnose, treat and prevent such disorders. Genome analysis involves the development of new tools for and approaches to the study of the genomes--the complete genetic endowments--of humans and model organisms such as yeast, fruit flies and mice. In addition, some scientists supported under this initiative will be determining the sequence, or order, of subunits of the genetic material DNA.

The awards are part of the National Institutes of Health's efforts to characterize the genomes of humans and other complex organisms. What separates grants supported under the NIGMS initiative from similar research supported over the past several decades is the systematic approach scientists will take in mapping all of an organism's genes and analyzing complete genomes, rather than searching for and studying specific genes of interest.

A list of the newly awarded grants is attached.

# # #

NIGMS GENE MAPPING/GENOME ANALYSIS GRANTS

AWARDED IN AUGUST AND SEPTEMBER 1988

Marsha I. Altschuler, Ph.D. Williams College Williamstown, MA	"Cloning and Analysis of Tetrahymena Somatic Chromosomes" $142,262*
Norman G. Anderson, Ph.D. LSB Corporation Rockville, MD	"Camera/Scanner for Genomic Mapping and Protein Analysis" $252,266
Michael L. Boehnke, Ph.D. University of Michigan Ann Arbor, MI	"Design Issues in Genetic Linkage and Segregation Studies" $104,207
Benjamin Burr, Ph.D. Brookhaven National Laboratory Upton, NY	"Mapping and Characterizing Quantitative Trait Loci in Maize" $135,502
Aravinda Chakravarti, Ph.D. University of Pittsburgh Pittsburgh, PA	"Genetic Heterogeneity in Human Disease" $69,477
Francis S. Collins, M.D., Ph.D. University of Michigan Ann Arbor, MI	"Chromosome Mapping by a DNA Circularization Technique" $99,629
Donna L. Daniels, Ph.D. University of Wisconsin Madison, WI	"Complete Dissection of the Escherichia Coli Genome" $136,542
Jean Dausset, M.D. Centre d'Etude du Polymorphisme Humain Paris, France	"Primary Structure of the Human MHC" $384,150
Ronald W. Davis, Ph.D. Stanford University Stanford, CA	"Cell Regulation: Biochemically Isolated DNA Segments" $513,684
Prescott L. Deininger, Ph.D. Louisiana State University Medical Center New Orleans, LA	"Human Interspersed Repeated DNA Sequences" $108,188
Harry A. Drabkin, M.D. University of Colorado Health Science Center Denver, CO	"Construction of a Physical Map for Human Chromosome 3" $177,098

*figures represent direct and indirect costs for the first year of support only; most of the grants will run from 3 to 5 years

(More)

Argiris Efstratiadis, M.D., Ph.D. Columbia University New York, NY	"cDNA Probes for Exon Mapping of the Human Genome" $322,887
Pamela R. Fain, Ph.D. University of Utah Salt Lake City, UT	"One CentiMorgan Genetic Maps of Chromosomes X and 17" $194,251
Katheleen Gardiner, Ph.D. Eleanor Roosevelt Institute for Cancer Research Denver, CO	"Cloning Large (>200 Kb) Fragments of Chromosome 3 DNA" $99,383
Robert M. Gemmill, Ph.D. Southwest Biomedical Research Institute Scottsdale, AZ	"A Physical Map for Human Genomic Region 3p14.2 to 3p21.1" $159,834
Varghese T. George, Ph.D. Louisiana State University Medical Center New Orleans, LA	"Estimation and Test of Marker Association in Family Data" $68,713
Anton S. Goustin, Ph.D. CBR Laboratories, Inc. Boston, MA	"Novel HLA-Linked DNA in a Yeast Artificial Chromosome" $50,000
Neil R. Hackett, Ph.D. Vanderbilt University Nashville, TN	"Physical Mapping of an Archaebacterial Genome" $111,406
Stephen C. Hardies, Ph.D. University of Texas Health Science Center at San Antonio San Antonio, TX	"Mammalian Repetitive Sequences as a Genetic Tool" $98,547
Daniel L. Hartl, Ph.D. Washington University St. Louis, MO	"Drosophila Genome Mapping Using YAC Vectors" $395,322
Timothy G. Helentjaris, Ph.D. NPI Salt Lake City, UT	"Novel Rapid RFLP Analysis With Fluorescence Detection" $49,808
Leroy E. Hood, M.D., Ph.D. California Institute of Technology Pasadena, CA	"Characterization of the Genomes of Humans and Model Organisms" $501,934
William D. Huse, M.D., Ph.D. Stratagene, Inc. La Jolla, CA	"Novel Vectors for Gene Mapping and Sequencing" $277,637

(More)

Lynn B. Jorde, Ph.D. University of Utah Salt Lake City, UT	"Linkage Disequilibrium in DNA Polymorphisms" $134,841
Samuel Karlin, Ph.D. Stanford University Stanford, CA	"Analysis of Molecular Sequence Data" $367,851
Bronya J. Keats, Ph.D. Louisiana State University Medical Center New Orleans, LA	"Maximum Likelihood Mapping of the Human Chromosomes" $118,924
Leonard S. Lerman, Ph.D. Massachusetts Institute of Technology Cambridge, MA	"Strategy for the Characterization of the Human Genome" $328,567
Terry J. Lerner, Ph.D. Integrated Genetics, Inc. Framingham, MA	"Chromosome Signposts for Gene Mapping" $50,000
Michael Lovett, Ph.D. Genelabs Incorporated Redwood City, CA	"New Methods for Producing Chromosome-Specific Libraries" $252,054
Michael Lovett, Ph.D. Genelabs Incorporated Redwood City, CA	"Study of Linked Growth Factor Genes on Chromosome 5" $50,000
Hugh O. McDevitt, M.D. Stanford University Stanford, CA	"Genetics of HLA-D and T-Cell Receptor Haplotypes" $159,372
William McGann, Ph.D. Radiation Monitoring Devices, Inc. Watertown, MA	"Improved RFLP Analysis for Genetic Linkage Mapping" $50,000
Victor A. McKusick, M.D. Johns Hopkins University Baltimore, MD	"Mapping the Chromosomes of Man" $963,420
Orlando J. Miller, M.D. Wayne State University Detroit, MI	"Function of CpG-Rich Islands in Human DNA" $177,599
Donald T. Moir, Ph.D. Collaborative Research, Inc. Bedford, MA	"Molecular Cloning and Analysis of Human Genomic Regions" $272,376
Jeffrey C. Murray, M.D. University of Iowa Iowa City, IA	"Detailed Mapping and Recombination of Chromosome Four" $428,127

(More)

Joseph H. Nadeau, Ph.D. Jackson Laboratory Bar Harbor, ME	"Gene Mapping and Genome Organization in Mammals" $228,353
Maynard V. Olson, Ph.D. Washington University St. Louis, MO	"Refinement of the Yeast Physical Map" $152,426
Eduardo Orias, Ph.D. University of California, Santa Barbara Santa Barbara, CA	"Molecular Genetics of a Cell-Cell Recognition System" $157,115
Howard C. Passmore, Ph.D. Rutgers University New Brunswick, NJ	"Recombination Analysis of the Mouse H-2 Complex" $162,091
Donald A. Pious, M.D., Ph.D. University of Washington Seattle, WA	"Genetic Structure and Function in Human Cells" $215,765
Eugene M. Rinchik, Ph.D. Martin Marietta Energy Systems, Inc. Oak Ridge, TN	"Saturation Mutagenesis of Mouse Genomic Regions" $183,538
Neil J. Risch, Ph.D. Yale University New Haven, CT	"Statistical Methods and Applications in Human Genetics" $228,313
Frank H. Ruddle, Ph.D. Yale University New Haven, CT	"High Resolution Genetic Analysis of Complex Genomes" $879,164
Thomas B. Shows, Ph.D. Roswell Park Memorial Institute Buffalo, NY	"Mapping Human Chromosome 11" $404,391
Joseph A. Sorge, M.D. Stratagene, Inc. La Jolla, CA	"Technologies Aimed at a Physical Map and Sequence of the Human Genome" $269,720
John C. Sutherland, Ph.D. Brookhaven National Laboratory Upton, NY	"New Approaches to DNA Mapping: Electronic Imaging" $155,635
Ahmad Waleh, Ph.D. Applied Sciences Consultants San Jose, CA	"Computer Folding of RNA Using Monte Carlo Methods" $135,934
Dorothy P. Warburton, Ph.D. Columbia University New York, NY	"Monochromosomal Hybrids by Retroviral Marker Transfer" $137,846

(More)

Robert H. Waterston, Ph.D. Washington University St. Louis, MO	"A Physical Map of the C. Elegans Genome" $195,659
Paul C. Watkins Integrated Genetics, Inc. Framingham, MA	"Automated Large Scale Analysis of Human Cosmid Libraries" $50,000
Steven G. Weaver, Ph.D. University of Illinois at Chicago Chicago, IL	"The Maintenance of Sequence Homology by Gene Conversion" $224,770
Carol A. Westbrook, M.D., Ph.D. University of Chicago Chicago, IL	"Macrorestriction Mapping Applied to Chromosome 5q" $264,298
Raymond L. White, Ph.D. University of Utah Salt Lake City, UT	"Tools for High Resolution Linkage Mapping of Human Genes" $1,499,741
Huntington F. Willard, Ph.D. University of Toronto Toronto, Ontario, Canada	"Mapping Centromeric Regions of Human Chromosomes" $153,724

# # #

DEPARTMENT OF HEALTH & HUMAN SERVICES

National Institute of Neurological Disorders and Stroke

Memorandum

Date December 14, 1988

From Director, Basic Neurosciences Program, DIR, NINDS

Subject NINDS Activities Related to the Human Genome Project

To Elke Jordan, Ph.D., Director, Office of Human Genome Research, OD
Building 1, Room 332

The NINDS has a significant interest in the human genome project because many neurological diseases or the disposition to develop neurological disorders of stroke are genetically inherited. In addition, diseases primarily affecting other organisms have neurological consequences.

I attach a list of some of the disorders whose chromosome locations have been identified by NINDS grantees so that a detailed genetic analysis of the relevant parts of these chromosomes would help to determine the exact location of the involved gene and could then be used for exact diagnosis and genetic counseling. If one wanted to support work that is not only of neurological but also of general genetic interest, one could study the DNA sequence in the centromere region of chromosome 17 in order to identify the properties of the centromere; by using patients with neurofibromatosis (NF-2), the gene of which is closely linked to the centromere, one might then be able to locate this gene.

The NINDS Extramural program also supports the mapping of mitrocondrial genes responsible for physiological functions and for maternally inherited diseases. It supports the mapping and sequencing of homeobox-, proto- and oncogenes. Furthermore, it supports the identification and sequencing of genes that encode glial specific proteins and various growth factors important for the growth or movement of neurons or glial cells, and the characterization of genes expressed in neural crest cells.

The Intramural Research Program examines the properties, sequence, and chromosome location of genes controlling the production of certain receptors (adrenergic, muscarinic acetylcholine, glutamate receptors) and of genes concerned with the metabolism of glutamate (glutamine synthetase, glutamate dehydrogenase, glutaminase) and the mechanisms controlling the expression of S-100, glial fibrillary acidic, and calcium transport proteins.

Ernst Freese, Ph.D.

Attachment

cc: Dr. Roger Porter
Ms. Mary Miers

Table 1

Disorder	Heredity	Location	
Adrenoleukodystrophy	XL	Xq27-28	White matter
Alport syndrome	XL	Xq	acoustic nerves & other nerves
Amyloid polyneuropathy	AD	18p	Neurons - CNS & PNS
Becker muscular dystrophy	XL	Xp21	In muscle
Beta amyloid gene (Alzheimer, Down)	AD	21q	Neurons C
Bilateral acoustic neurinoma (NS-2)	AD	22q11	Neurons, tumor in brain
Duchenne muscular dystrophy	XL	Xp21	
Friedreich ataxia	AR	9	Neurons: pons, spinal cord
Hereditary motor and sensory neuropathy	XL	Xq13-21	Neurons: C & P
Huntington chorea	AD	4p16.3ter	Neurons: C
Myotonic dystrophy	AD	19p	In muscle & spinal neuron
Spinocerebellar ataxia	AD	6q	Neurons: Cerebellum
Tay-Sachs Hexosaminidase A	AR	5	
Tay-Sachs Hexosaminidase B	AR	15	
Tourette syndrome	AD	18q22?	Neurons: C
Tuberous sclerosis	AD	9q	Brain tumors
Von Hippel-Lindau disease	AD	3p	Brain tumors
Neurofibromatosis, Von Recklinghausen (NF-1)	AD	17cen	Peripheral nerve & skin tumors

XL = X-linked

AD = Autoneural dominant

AR = Autoneural recessive

cen = close to centromere

ter = close of terminal

DEPARTMENT OF HEALTH & HUMAN SERVICES

National Institute of Neurological and
Communicative Disorders and Stroke

Memorandum

Date December 6, 1988

From Director DDADD, NINDS

Subject DDADD Material for Meeting of NIH Advisory Committee on the Human Genome

To Associate Director for Extramural Activities, NINDS

The genome related activities of this Division are concerned with research on the genetics of neurodegenerative disorders of adult life. The major diseases for which high interest in genetic factors is relevant and in which exciting recent findings have been made are Alzheimer's and Huntington's diseases. The activities in each of these areas is briefly described below:

-Alzheimer's disease. Four large families have been described in which Alzheimer's disease is caused by an autosomal dominant gene defect. This defective gene has been linked to markers on chromosome 21. The findings that the genes on chromosome 21 may be the proximate cause of Alzheimer's disease are being pursued and extended. Additional families will be studied to confirm the autosomal dominant pattern of inheritance in familial Alzheimer's disease. Additional markers have been identified on chromosome 21 which are linked to the gene for familial Alzheimer's disease and it is clear that other markers will soon be discovered. A better definition of the location of the familial Alzheimer's disease defect will be accomplished with these markers and they should facilitate the isolation of the gene itself.

-Huntington's disease. Thirteen new markers have been identified since the discovery that the Huntington's disease gene is located on chromosome 4. Research is actively continuing on the precise localization and possible cloning of this gene.

-There is also significant interest in identifying genetic factors which predispose to multiple sclerosis.

-Recent findings have reported localization of the hereditary ataxia gene to chromosome 9.

I hope this information is helpful.

Carl M. Leventhal, M.D.

DEPARTMENT OF HEALTH & HUMAN SERVICES

National Institute of Neurological and Communicative Disorders and Stroke

Memorandum

Date December 7, 1988

From Director, DCDND, NINDS

Subject Your request for Information Concerning the Breadth of Genome Related Activities Ongoing at NIH

To Associate Director, DEA, NINDS

Our response is very elliptic in form, since we felt that the scientific members of the Committee would not need a detailed explanation of gene research. Our response lists nine areas of research activity supported by NINDS and identifies 16 neurogenetic diseases in which localization of the abnormal gene has been accomplished by NINDS supported research..

1. Mapping of chromosomal genes responsible for single gene disorders.

2. Mapping of mitochondrial genes responsible for physiolgical functions and for maternally inherited disease.

3. Identification, mapping and sequencing of homeobox genes (which control regulation).

4. Identification and mapping of protoonco- and oncogenes.

5. Identification of variety of mutations that occur within a gene or gene complex (genetic heterogeneity).

6. Identification and sequencing of genes that encode glial specific proteins.

7. Identification, mapping and sequencing of genes for the various growth factors.

8. Identification and mapping of genes that control behavioral traits (eg birdsong).

9. Isolation and sequencing of genes expressed in neural crest cells.

Table 1 lists sixteen neurogenetic diseases in which chromosome assignments have been made.

F. J. Brinley, Jr., M.D., Ph.D.

Table 1

	Disorder	Heredity	Location
008	Adrenoleukodystrophy	XL	Xq27-28
014	Alport syndrome	XL	Xq
017	Amyloid polyneuropathy	AD	18p
020	Becker muscular dystrophy	XL	Xp21
V 023 024	Beta amyloid gene (Alzheimer, Down)	AD	21q
V 030 031	Bilateral acoustic neurinoma (NS-2)	AD	22q11
042 043	Duchenne muscular dystrophy	XL	Xp21
046	Friedreich ataxia	AR	9
V 049 050	Hereditary motor and sensory neuropathy	XL	Xq13-21
053	Huntington chorea	AD	4p16.3ter
056	Myotonic dytrophy	AD	19p
069	Spinocerebellar ataxia	AD	6q
072	Tourette syndrome	AD	18q22?
075	Tuberous sclerosis	AD	9q
078	Von Hippel-Lindau disease	AD	3p
C 081 082	Von Recklinghausen neurofibromatosis (NF-1)	AD	17cen

Division of Research Resources

Resources Supporting Mapping and Sequencing

Capabilities For Complex Genomes

The Division of Research Resources (DRR) of the National Institutes of health is a medium-sized extramural granting program (Fiscal Year [FY] 1988 = $368 Million) whose mission is carried out though five subactivities. Its purpose is to provide access to necessary specialized resources, environments, and facilities for biomedical research. The DRR provides small but significant research resources for complex genome research by assisting scientists in the design, analysis, and interpretation of experiments in genetics and gene expression. Attached are brief descriptions of these resources.

U41RR01685 BIONET

Michael J. Kelly, Ph.D.
IntelliGenetics, Inc.

BIONET, a National Computer Resource for Molecular Biology, is supported by a cooperative agreement through DRR's Biomedical Research Technology Program. The BIONET resource supplies computer software, hardware, databases and consultation services to molecular biologists. The Resources also serves as a focus for development and sharing of software and information among a national community of scientists.

The Core library includes nine programs from IntelliGenetics (of IntelliCorp, Palto Alto, California) for representing and manipulating nucleic acid and protein sequence data. The Database library contains existing databases of nucleic acid and protein sequences, including NIH's DNA sequence library: GenBank; the European Molecular Biology Laboratory's nucleotide sequence data library; IntelliGenetics' Vector Bank of cloning sequences; and the restriction enzyme database from Cold Springs Harbor. The System and Programming Support Library provides tools for program development. Staff support includes on-line consultations, regional training sessions, and documentation of the core and contributed programs.

BIONET was also designed to enhance communication among the research community. This is accomplished primarily through electronic mail and eight bulletin boards. Together, electronic mail and bulletin boards provide rapid, community-wide access to the latest information on topics in molecular biology.

P41RR01315 National Flow Cytometry and Sorting Resource (NFCR)

L. Scott Cram, Ph.D.
University of California, Los Alamos National Laboratory

The Division of Research Resources, through the Biomedical Research Technology Program, supports the National Flow Cytometry and Sorting Resource (NFCR) at Los Alamos National Laboratory. The significance of the Resource is in its unique ability to solve forefront biomedical problems such a karyotype analysis and chromosome sorting using flow cytometry. Over 180 man years of collective flow cytometry experience is represented by the staff at the NFCR.

The chromosome High Resolution Imaging Sorter (CHRIS) developed at the NFCR fills a crucial need arising from the molecular genetics community. In the first year of operation and continually since then, the NFCR has received a large number of requests for sorted chromosomes. To avoid sorting the same chromosome repeatedly, DNA libraries of the flow sorted chromosomes were constructed to amplify the material. Many investigators desire probes directly from sorted chromosomes, other investigators have requested chromosomes from a particular cell strain they are studying to determine sites of viral integration or genetic rearrangement.

The research to be developed at NFCR for chromosome analysis and sorting will offer the molecular and cellular geneticist advanced techniques that are available in only a few laboratories. The availability of flow sorted chromosomes can be a critical step in the search for restriction fragment length polymorphisms, disease locii and ordering DNA sequences, because the genome is divided into its natural units of heredity. The purpose of this research is to extend the means of resolution of normal and abnormal chromosomes, to improve methods for sorting chromosomes, and to advance flow cytometric instrumentation. These developments are essential o meet the increasing needs of the genetics community.

P41RR01821 Protein Identification Resource (PIR)

Robert S. Ledley, D.D.S.
Winona C. Barker, Ph.D.
Georgetown University Medical Center

The Protein Identification Resource provides on-line public access to a comprehensive protein sequence knowledge base and retrieval system, performs sequence identification, and (through funding from NSF) publishes the Atlas of Protein Sequence and Structure.

Research efforts at the resource are directed towards development of computer programs and expert systems for protein identification and knowledge-based information analyses. Examples of the latter include predictions of secondary structure, antigen recognition sites, best nucleic acid sequence probes and enzyme cut sites of coding regions.

U41RR52101 PROPHET National Computer Research Resource

Charlotte Hollister, Ph.D.
Bolt Beranek and Newman Laboratories, Inc.
Cambridge, Massachusetts

PROPHET is a research resource developed specifically for use by pharmacologists, medicinal chemists, and other biomedical scientists studying the interactins of chemical and biolcical systems. PROPHET features tools for data management, data analysis, molecular modeling, and sequence analysis. Utilizing a network of high performance graphics workstations and ocmmunicating with PROPHET via higher level, English-like commands, scientists can create, modify, display and manipulate tables, graphs, molecular structures and amino and nucleic acid sequences.

Currently, more than 800 researchers at more than 40 academic, government and commercial research institutions throughout the United States are using PROPHET.

A second generation of PROPHET is currently being developed and is due to replace the existing system in FY 1988. PROPHET II is designed as a fully distributed system, primarily used on local workstations. The central facility will be used to provide access to the latest versions of scientific databases, to facilitate information sharing, software development and distribution, to enable electronic mail interchange, and to provide resources for specialized computer applications.

P41RR02188 A Resource for Analysis of 2-D Protein Gels

James I. Garrels, Ph.D.
Cold Spring Harbor Laboratory

The regional Resource for Analysis of 2D Protein Gels at Cold Spring Harbor provides protein mixture separation and identification using two-dimensional gel electrophoresis, numerical databases and interactive computer graphics. The resource is based on the laboratory and computer facility for gel analysis established by Dr. James Garrels. The resource's research focus is directed towards establishment of public domain protein databases and dissemination of 2D gel electrophoretic analytic technology more widely throughout the biomedical research community.

A spin-off commercial venture, Protein Databases Incorporated, has been organized to develop and provide state-of-the-art graphic workstations and software for performing 2D protein gel analysis.

R24RR02581 Molecular Biology Information Resource (MBIR)

Charles B. Lawrence, Ph.D.
Baylor College of Medicine

The Molecular Biology Information Resource (MBIR) supported through DRR's Biomedical Research Technology Program is a regional resource in the Houston area, serving the needs of molecular biologists at the Baylor College of Medicine, Rice University, the University of Texas Health Center and the M.D. Anderson Tumor Institute.

The MBIR is conducting research and development related to computtational support of molecular biolgo research. A major goal of the MBIR is to put advanced computer tools in the hands of working molecular biologists by providing: 1) access to state-of-the-art hardware; 2) access to specialized software; 3) training on the use of computers in support of molecular biology research; and 4) a center of expertise. Present research is focused on the development of methods for performing rapid and sensitive sequence and similarity searches of the gene and protein sequence data banks.

P41RR02275 Molecular Biology Computer Resource and Research Program (MBCRR)

Temple F. Smith
Dana-Farber Cancer Institute

The Molecular Biology Computer Resource and Research Program provides DNA, RNA and protein sequence analysis computer programs to investigators in the Northeastern United States, on-line, and to a larger research community on tape and diskette. This service component is coupled with a research program directed at development of statistical methods and microcomputer software for pattern recognition in molecular genetics.

R01AI25616 Amino Acid and Nucleotide Sequences of Proteins of Immunological Interest

Elvin Abraham Kabat, Ph.D.
Columbia University

Since 1975, the Division of Research Resources, with NCI, NIAID, NIDDK and NIGMS, has funded the Amino Acid and Nucleotide Sequences of Proteins of Immunological Interest project at Columbia University. This research project involves computer-assisted tabulation and analysis of amino acid and nucleotide sequences. The database contains the most complete set available of protein and nucleic acid sequences in gene coding regions of immunoglobulins and related proteins. The project is critically dependent on the capabilites of the PROPHET system resource supported by DRR.

The information compiled in this extensive, thoroughly annotated database has been published in 1976, 1979, 1983 and 1987, with the 5th edition scheduled for late 1990. The first two editions contained amino acid sequences only. Nucleotide data was added in the later two editions.

The format of the published data provides classified and aligned sequences of amino acids and nucleotides so that rapid examination and comparisons of data are possible.

The 3rd edition of 2,000 copies was completely exhausted and the 4th edition printing was over twice as large. Additionally, the 4th edition is available on tape and diskette.

The Caenorhabditis Genetics Center
Columbia, Missouri

Dr. Donald L. Riddle, P. I.
Dr. Mark L Edgley, Curator

The Biological Models and Materials Section supports the Caenorhabditis Genetics Center which is a repository and distribution center for mutants of the invertebrate nematode, C. elegans. The small size of C. elegans and its defined embryonic lineage provide an important model and approach to research in genetics. Presently, over 1200 strains of the nematode are available to researchers. The Center also updates and distributes the C. elegans genetic map which allows researchers to have immediate access to invaluable information on the genome.

The MIT Cell Culture Center
Cambridge, Massachusetts

Dr. Phillip Sharp, P. I.
Mr. Donald Giard, Director

The Section supports the Center to provide a customized service for research investigators needing extremely large quantities of cells in culture, cell products, or viruses in their research. The primary mission of the Center is to produce these biological materials on a large scale to allow scientists to conduct novel and important experiments in basic biology that could not be accomplished with the facilities and resources available in their own laboratories. Many types of cell lines have been grown in the Center. Investigators in a wide range of areas, including genetics, and cellular and molecular biology, use the Center's services.

The Repository of Human DNA Probes and Libraries
Rockville, Maryland

Dr. William C. Nierman
Dr. Donna R. Maglott

The Division of Research Resources, through the Biological Models and Materials Resources Section, supports the human chromosome-specific libraries of the Repository. The Repository, located at the American Type Culture Collection (ATCC) in Rockville, Maryland, collects, maintains and distributes cloned genes, human DNA probes, and the DNA libraries developed at the Los Alamos and Lawrence Livermore Laboratories. The Repository now has a complete set of HindIII and EcoR1 limit digest libraries for each flow-sorted human chromosome in Charon 21A vectors. The Repository provides these invaluable tools to research scientists interested in many areas of human molecular gentics, including mapping the human genome and understanding the control of gene expression. Many genetic diseases can now be analyzed by specific probes and this will result in the isolation of the disease-associated genes and the development of diagnostic methods and new gene therapies. Thus far, the Repository has distributed over 700 DNA libraries in a two-year period to researchers in 28 countries.

Unisys
Paoli, Pennsylvania

Dr. Christopher Overton, P.I.

One of the objectives of the Biological Models and Materials Resources Section is to explore and develop nonbiological models for biomedical research. The Section supports the development of the "matrix of biological knowledge" as a potential tool for the biomedical community. The matrix is defined as the organization, or complete database, of all published biological and clinical data, and the theories and analyses of these data.

To pursue the matrix concept, the Section supports a pilot project with Unisys to examine and develop a means to systematically connect biological databases such that researchers can access the matrix of biological knowledge in a meaningful way. The project will build a testbed for the matrix concept on top of a restricted area of knowledge, the existing nucleic acid sequence library of Genbank. This library has become an invaluable tool in investigations of the human genome; however, many types of complex queries cannot be posed without special-purpose, dedicated computer programs and computer assistance. This project should provide researchers in human molecular genetics with greater access to the existing nucleic acid sequence information and will test the feasibility of the matrix.

Yeast Genetic Stock Center
Berkeley, California

Dr. Robert Mortimer, P.I.
Dr. Rebecca Contopoulou, Curator

The Biological Models and Materials Section, Division of Research Resources, supports the Yeast Genetic Stock Center which collects and distributes approximately 900 genetically-defined strains of the yeast, Saccharomyces cerevisiae, to biomedical researchers. This organism provides an important model and approach for research in genetics. The Center also publishes an updated catalogue of strains, and distributes and updates the S. cerevisiae gentic map.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date December 9, 1988

From Director, FIC

Subject Meeting of NIH Program Advisory Committee on the Human Genome

To Director, Office of Human Genome Research

This is in response to your memorandum of November 29 requesting summaries of genome-related activities to be used as background information for members of the NIH Program Advisory Committee on the Human Genome. Information on a single fellowship being supported by the Fogarty International Center (FIC) is attached.

I would also like to take this opportunity to affirm the FIC's interest in the overall international effort to sequence and map the human genome and to offer our assistance in coordinating any international efforts in which the NIH may become involved. It is clear that, in an activity of this magnitude and complexity, careful attention must be given on an international level to issues such as cost sharing, reciprocal access to R&D facilities, comparability of and access to data, and protection of intellectual property rights, and various commercial implications of the technology and data involved in the effort. The FIC staff have extensive experience dealing with such matters and are prepared to work with your office in any way you deem necessary and beneficial.

I would be pleased to discuss this further at your convenience.

Philip E. Schambra, Ph.D.

Attachment

Human Genome Research Supported by the Fogarty International Center

An International Research Fellow from Spain, Dr. Juan C. Sabala, is studying the structure and function of the genes encoding the microtubule-associated proteins, MAP1 and MAP2, with a U.S. researcher at New York University Medical Center.
(Grant Number 1 FO5 4098-01)

NATIONAL LIBRARY of MEDICINE NEWS

Bethesda, MD 20894
Robert Mehnert
(301)496-6308

September 26, 1988

Research Grants for Informatics Aspects of Biotechnology

The first group of seven grants to investigate the computer science aspects of molecular biology data management and analysis has been awarded by the National Library of Medicine in fiscal year 1988. They are funded under a special appropriation by the Congress to the NLM.

The objective of this new grant program is to encourage research into advanced computing methods for understanding the molecular machinery of life. New experimental methods have increased the rate at which laboratories are producing data about nucleic acids, proteins, and other biologically important molecules in order to diagnose and treat human disorders. The computer databases that hold this information, currently numbered in millions of nucleotide base pairs and thousands of amino acids, are expected to grow at a great rate to encompass sequences totaling *billions* of nucleotides. Current methods for structuring, searching, and analyzing such databases need to be greatly improved to handle this information.

In seeking grant applications, the Library posited as sample research topics:

- Design of databases and how data are represented in them
- Algorithms to improve the efficiency of retrieval
- Algorithms to retrieve from multiple, related databases
- Algorithms to predict structure/function
- Computers systems to recognize patterns and compare sequences

The amounts listed on the attached sheet are for the first year of multi-year grants and include both direct and indirect costs.

NLM RESEARCH GRANTEES IN BIOTECHNOLOGY

Brutlag, Douglas L., Ph.D. $146,913
Associate Professor of Biochemistry
Stanford University
Stanford, CA 94305
"Symbolic Simulation of DNA Metabolism"

Markley, John L., Ph.D. $232,095
Professor of Biochemistry
National Magnetic Resonance Facility at Madison
University of Wisconsin
Madison, WI 53706
"Creation and Analysis of Archival Protein NMR Database"

Marr, Thomas G., Ph.D. $140,914
Staff Member
Theoretical Biology & Biophysics
Los Alamos National Laboratory
Los Alamos, NM 87545
"Computer Representation and Reduction of Chromosome Mapping Data"

Merriam, John R., Ph.D. $ 63,894
Associate Professor
Department of Biology
University of California, Los Angeles
Los Angeles, CA 90024-1606
"Cloned DNA by Chromosome Location"

Myers, Eugene Wimberly, Ph.D. $141,069
Associate Professor
Department of Computer Science
University of Arizona
Tucson, AZ 85721
"Efficient Software for the Analysis of Biosequences"

Pearson, William R., Ph.D. $105,138
Assistant Professor
Department of Biochemistry
University of Virginia
Charlottesville, VA 22908
"Comparison of Protein Sequences and Structures"

Roberts, Richard J., Ph.D. $159,719
Assistant Director for Research
Nucleic Acid Chemistry
Cold Spring Harbor Laboratory
Cold Spring Harbor, NY 11724
"Functional Motifs and Errors in Biological Sequences"

National Biotechnology Information Center of the National Library of Medicine

NBIC Center Programs: Current and Future

December, 1988

Current activities

On November 4, 1988, President Reagan signed Public Law 100-607 creating a National Center for Biotechnology Information at the National Library of Medicine, funded at the level of $8 Million in FY 1989. The Center represents a strengthening of ongoing NLM projects to improve the content and useability of molecular biology (Biotechnology) factual databases. Center activities may be functionally divided into those which assist in the building and maintenance of factual databases for genetics and molecular biology, those which will facilitate retrieval of information from those databases, and support for intramural and extramural informatics research.

A. Database Building and Maintenance

1. MEDLINE indexing of nucleic acid and protein sequence data

   Since October, 1987 the index terms MOLECULAR SEQUENCE DATA (+ NUCLEIC ACIDS or PROTEIN) have been applied to new MEDLINE entries which either contain molecular sequence data or include an accession number for a new entry into the GenBank, EMBL, or PIR databanks. This process provides systematic surveillance for the appearance of new data in over 4000 biomedical journals. The new index terms are useful for online searches. As well, they are used to produce monthly subsets of literature citations with abstracts, which are provided to molecular biology databank builders.

2. GenBank-MeSH linkage

   The NLM has "mapped" the keywords used in the GenBank nucleotide sequence database to relevant Medical Subject (MeSH®) headings, so that once a researcher has found a set of GenBank records of interest, he may easily and automatically retrieve relevant literature indexed by the same concepts. In addition, the MEDLINE database has been modified to insert "pointers" to GenBank and other factual databases in a "Secondary Source ID" field of the MEDLINE record, so that a user who retrieves literature will know if the data described in an article is also available from a computer database. As well, at the request of the GenBank Advisory Board, scientists in NLM's Biotechnology Information Center are assisting in monitoring the currency and completeness of GenBank and other molecular biology databases as compared to the biomedical literature.

3. Full text and image databases in genetics

The Online Reference Works research project of the Lister Hill National Center for Biomedical Communications maintains an online, full text version of *Mendelian Inheritance in Man*, the premier medical textbook describing over 4000 human genetic diseases. Current research involves enhancement of the access to this database. Retrieval of text by natural language query (e.g. "Show me the genetic syndromes which cause ear deformities") is being improved using an expanded thesaurus of synonyms and related terms, and a visual library of clinical images (photographs and x-rays) on videodisc is being linked to the text.

4. A Database of Biotechnology Information Resources

There are many databanks nationally and internationally which contain biotechnology information. A contract to build a "database of databases" containing directory information for genetics and molecular biology information is underway under the direction of the Library's Specialized Information Services Division. This database will serve as a pointer to other databases, and may become part of the Library's DIRLINE (Directory of Information Resources onLINE) file. Public availability is expected in early 1989.

5. Future database design

The Library actively sponsors and participates in scientific meetings related to information and computer science in biology and medicine, and the implications for future database design. Examples include a March 1988 Workshop on Algorithms for Molecular Genetics held at the Lister Hill Center, and numerous biotechnology and genomic database planning meetings involving other NIH Institutes, other Federal Agencies such as the Department of Energy, and private foundations such as the Howard Hughes Medical Institute.

In 1989 the NLM will sponsor the second "Macromolecules, Genes and Computers Workshop" in New Hampshire, and provide support for the Tenth Human Gene Mapping Workshop in New Haven.

B. Information Retrieval from databases

1. IRX multi-database prototype

Collaborating with NIH intramural molecular biologists, the developers of the Lister Hill Center's generalized information retrieval experiment (IRX) research software are adapting the system to produce an advanced laboratory workstation for molecular biology. As shown in the accompanying diagram (Fig. 1), this prototype system provides access to multiple factual databases using natural language. The test community for this effort will be NIH molecular biologists, who will be provided custom-written terminal software for PCs. Development work on the system is being guided by advice from biologists in the Laboratory of Biochemical

Pharmacology of NIDDK, and NCI's Laboratory for Mathematical Biology in Frederick, where researchers use the high resolution graphical computer workstations for which the IRX system is best suited (see sample computer screen and query, Figure 2). After the initial development stage is complete in 1988, the software will be distributed to all interested NIH molecular biology labs, and later made available publicly.

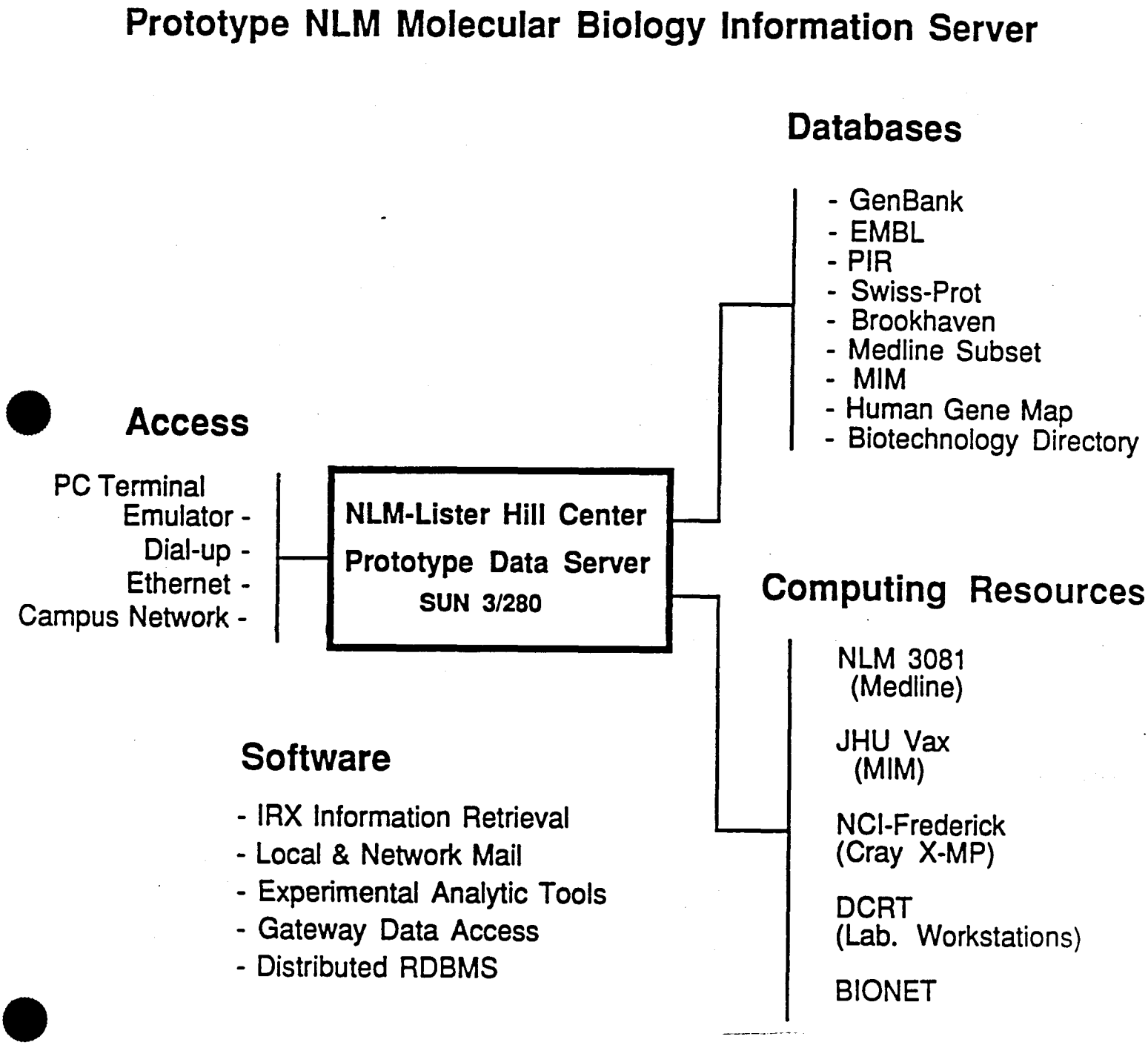


Figure 1

Sample IRX computer screen, showing the query "What is known about the chromosome location of growth hormone deficiency?" answered with records from *Mendelian Inheritance in Man*, GenBank, MEDLINE, and chromosome map databases.

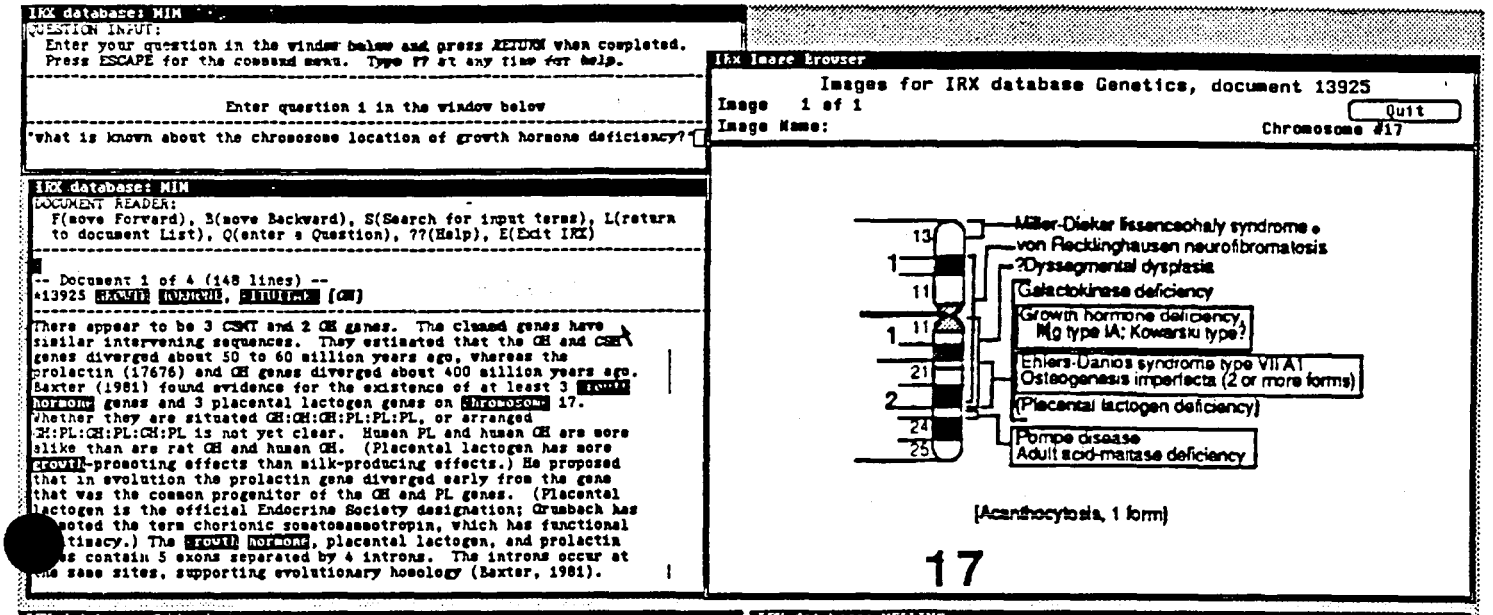


Figure 2

2. NIH Campus digital internet

As part of the collaboration with NIH laboratories, a high-speed digital link has been installed between the Lister Hill Center's Molecular Biology server computer, the Cray supercomputer at NCI's Frederick Research Facility, and the Division of Computer Research and Technology's Ethernet environment. This linkage provides DCRT and NCI access to the DOD Arpanet-MILNET computer network through the Lister Hill Center, and will provide users of the prototype molecular biology server computer access to the Cray for computationally intensive analysis tasks.

3. Biotechnology Environmental Release Database

The release of genetically modified organisms into environments outside of laboratories is an area of substantial scientific and regulatory interest in many countries. Beginning with an international workshop in March 1987, the NLM has undertaken a collaborative effort with other Federal and international organizations to develop a database system tailored to answering the experimental design, policy, and regulatory questions which are part of this fast-changing domain of biological science.

2. NLM molecular biology lecture series (SIS)

Since 1987, a number of eminent basic and clinical research scientists have presented seminars at the NLM, on topics related to laboratory methods in biotechnology, genetic engineering, and prospects for diagnosing and treating human disease using new technologies. Each of these lectures has been videotaped and made a part of the NLM audiovisual lending collection.

Toward the Future: The Biomedical Research Laboratory in Transition

Within ten years the tools necessary to conduct research in molecular biology will have changed substantially nationwide. Current manual methods of analyzing macromolecules (e.g., DNA and proteins) will have been replaced by benchtop auto-analyzers whose output includes computer data files as well as paper printouts. The retrieval of such data, "browsing" by researchers, and hypothesis testing will need to be carried out on powerful computer workstations with high-speed communications capabilities and high resolution graphical displays; such machines will have replaced the current generation of personal computers, at approximately the same cost.

High capacity information storage devices such as optical disk will be used to keep large volumes of relevant scientific data quickly available within the researcher's own computer, however these data sources will need to be supplemented by online access to new data and changes to existing data which are recorded between releases of entire databases. "Transaction-oriented" systems will become the norm, wherein the computer which calls to make a query of an online system will tell the central computer how much information it has locally; the central computer will send only that information which has changed since the local computer's last update. Relatively high communications costs (compared to storage media costs) and the tremendous volume of recorded data will be the driving forces for such system designs.

The recording and annotation of scientific data will need to be a distributed activity. Researchers who publish findings in the area of molecular biology will be encouraged and in most cases required to submit the actual data to central databanks. Computerized editing and annotation programs will allow the investigator to express the findings of his research in a machine understandable "molecular algebra", which will accompany the transmission of the data to central repositories. At such central databanks, the information will be validated by subject specialists who are scientists in their own right, conducting research to discover patterns, trends, and biologic principles evident in the database which may not be apparent from any single entry or group of entries. The capacity for online conferencing and peer review of databank entries will speed the entry and validation processes.

The National Library of Medicine in the Service of Future Biomedical Research

As the major national site for coordination and dissemination of biomedical research information, the NLM will have a multi-division program. There will be two complementary themes for NLM activities: the first will be information resource programs, serving as "library of record" for factual research data much the same way that the library maintains the biomedical literature; the second will be a focus of support for scientific discovery using information databanks as a substrate for experimentation.

A. Information resource programs

1. In the area of information resources, the National Center for Biotechnology Information staff will use contract and cooperative agreements to support key molecular biology databanks located around the country.

2. The contents of these databanks will be available through the NLM via gateways (i.e., translators which allow different types of computers to speak a common digital language) and networks, and will also be available on various physical media to researchers and value-added resellers.

3. The Center will develop standards for data interchange and computerized tools for distributed data entry and annotation by investigators.

4. The Center will sponsor extramurally (via grants and contracts) research to develop new methods of information representation and retrieval from factual databases of biological information. A parallel effort will be conducted in the intramural laboratories of the Lister Hill National Center for Biomedical Communications.

B. Information Science Research for Biotechnology

The design of such databanks and their resulting utility will, of necessity, need to follow scientific trends in research. The stimulus to maintain state-of-the-art systems will come from an intramural and extramural program supporting scientific discovery.

1. A core group of resident investigators and postgraduate fellows will develop new algorithms for analysis and hypothesis testing within the oceans of molecular data resident in the Center's databanks.

2. This will be complemented by a visiting scientist program, and an active extramural grant program. Fertile areas for such research in the next decade will include molecular sequence-to-structure prediction and determination of biologic similarity by expert system techniques.

3. The Center will sponsor informatics workshops and short courses, to instruct molecular biologists and medical researchers in the use of advanced computerized methods of data analysis.