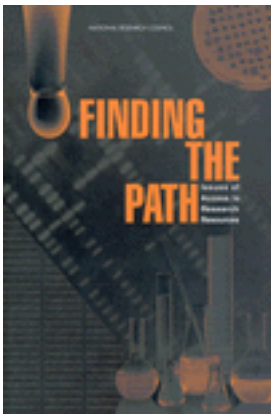

Finding the Path: Issues of Access to Research Resources

Committee on Federal Policy for Access to Research Resources, National Research Council

ISBN: 0-309-56318-6, 68 pages, 6 x 9, (1999)

FINDING THE PATH

Issues of Access to Research Resources

Summary of a Conference Held at the
National Academy of Sciences
January 27–28, 1999

Commission on Life Sciences
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C.

NATIONAL ACADEMY PRESS 2101 Constitution Avenue, N.W. Washington, D.C. 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. William A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

This study was supported by Grant No. DBI-9819600 between the National Academy of Sciences and the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number 0-309-06625-5

Additional copies of this report are available from the Commission on Life Sciences, Rm. 353, 2101 Constitution Avenue, N.W., Washington, D.C. 20418; (202) 334-1238; fax (202) 334-1687

Printed in the United States of America

CONFERENCE ADVISORY COMMITTEE

DAVID J. GALAS, (Chairman), Keck Graduate Institute of Applied Life Science, Claremont, CA
PAUL BERG, Stanford University, Stanford, CA
DAVID V. GOEDDEL, Tularik, Inc., South San Francisco, CA
ELLIOT M. MEYEROWITZ, California Institute of Technology, Pasadena, CA
RONALD R. SEDEROFF, North Carolina State University, Raleigh, NC
SHIRLEY M. TILGHMAN, Princeton University, Princeton, NJ

Science Writer

ROBERT POOL, Arlington, Virginia

NRC Staff

ROBERTA SCHOEN, Project Officer
JENNIFER KUZMA, Staff Officer
MIMI ANDERSON, Project Assistant
ANNE POND, Program Assistant
MARK RIVERA, Intern
NORMAN GROSSBLATT, Editor

COMMISSION ON LIFE SCIENCES

Staff

Priority Issues of Access to Research Resources

A Letter Report from the
Commission on Life Sciences

Advancement of scientific research in the life sciences is possible only with access to physical, biologic, and information resources. Such resources include plant and animal tissues, microbial cultures, monoclonal antibodies, reagents, animal models, combinatorial chemistry and DNA libraries, drug targets, clones and cloning tools, methods, laboratory equipment, databases, and software. Nearly every field of biology is experiencing problems in the transfer of research resources among members of its research community. While science continues to bring forth research resources of great potential, their dissemination often gets bogged down in issues of ownership, equity, availability, cost, appropriate use, value, and maintenance.

Many of those issues were aired on January 27-28, 1999, at the National Research Council's conference “Finding the Path: Issues of Access to Research Resources”. Sponsored by the Subcommittee on Biotechnology of the National Science and Technology Council's Science Committee, the conference convened over 300 participants from academe, government, and industry to discuss research-resource issues that affect numerous scientific disciplines. The purpose of the conference was to identify common issues and to place the challenge of access to research resources in a larger frame of reference—the entire scientific enterprise, but not to reach consensus on solutions to these challenges. A summary of the conference is published in this volume.

In March 1999, the Commission on Life Sciences met to discuss the issues further. We observed that many of the problems raised at the conference are important to the health and future of the scientific enterprise and the effective application of science. Some of the problems are not fundamentally difficult to overcome but will require the collective thought, organization, and

consensus of members of the scientific disciplines affected. Others are much more difficult and will require new approaches.

At the request of Dr. Mary Clutter, Chairman of the Subcommittee on Biotechnology, the Commission has identified priorities from among the issues raised in the conference summary. Based on that document and the collective experience of the Commission in the life sciences, we believe the following issues are particularly important and require attention by the federal government, and in some cases, by various sectors of the scientific community.

- Policies on the patenting of biological materials.
- Material transfer agreements and licensing.
- International material transfer.
- Database development and use.
- Access to data in the private sector.

This is not necessarily a comprehensive list of all the important issues of access to research resources. Indeed, all the issues raised at the January conference were important. We believe these issues are priorities because they affect research across the full spectrum of subdisciplines in the life sciences, and because they impact scientists in academe, government, and industry. International material transfer is included as a subclass of material transfer agreements with slightly different dimensions that warrant a separate discussion.

In the near term, the issues of patent policies, material transfer agreements, and access to privately held data are the most time-critical, and should be addressed sooner rather than later, because proposing and adopting solutions to them now is likely to have the greatest chance of success. Stakeholders involved in these problems are beginning to take actions—defensive patenting, excessive demands in exchange for access, increasing use of trade secrets—that will be difficult to reverse and that will have lasting effects on scientific progress.

The life sciences are in a revolutionary period of discovery, so identifying research resources and barriers to their development and dissemination should be a continuing part of the management of our scientific enterprise in the long term. The variety of barriers—in such forms as the high cost of a single piece of equipment, a bottleneck in software distribution, and competitive secrecy—requires constant monitoring and creative response. This effort must be the shared responsibility of the federal government, the academic scientific community, and corporations.

POLICIES ON THE PATENTING OF BIOLOGIC MATERIALS

In the relatively new, rapidly unfolding field of biotechnology, scientists and companies have envisioned future products of gene research for human health, agriculture, and many other fields. The realization of these products will depend, in part, on the accumulation of knowledge about the

functioning of a genome as a whole; this collective effort is proceeding quickly in the public and private sectors. As the sequencing of the human genome and the genomes of plants and other organisms are completed, there is a danger that the intellectual property rights afforded to new genetic constructs will be so broadly drawn that future scientific investigation and commercial development will be inhibited.

Since the Supreme Court opened the door to the patenting of genetically modified organisms in 1980, patenting has accelerated commercial development and complemented the progress of basic research in genetics. Recently, however, the award of broad proprietary rights to a new category of DNA sequences has had a dampening effect on academia and industry. In 1999, the first patent on an expressed sequence tag (EST) was issued. ESTs are small pieces of DNA that are part of complete, but as yet uncharacterized genes. Such gene fragments are potentially valuable research tools: they are used as probes and markers in the genomes of humans and other organisms. There is concern that the scope of the patents will be so broad as to interfere with basic research on the function of genes that are associated with a patented EST. In addition, the number of ESTs that might be eligible for patenting is potentially in the hundreds of thousands. When companies began to identify huge numbers of ESTs mechanically and to apply for patents on them, the US Patent and Trademark Office found it necessary to issue a policy to limit the number of ESTs per application to 10.

The award of the first EST patent is fueling speculation about the possibility that patents will be sought on other types of genetic information, such as single-nucleotide polymorphisms (SNPs), which are variations in DNA that provide insight into the genetic basis of disease, among other things. Many research scientists, particularly those in the academic community, consider SNPs to be research tools; like ESTs, SNPs are being identified rapidly and methodically in the genome.

The question of what scope of intellectual property rights protection best balances the public interest in creating, stimulating, and rewarding invention with the needs of the scientific community for access to research resources is urgent and important. When scientific material or information qualifies as the “door” through which all research must pass, its encumbrance by intellectual property rights, such as patents, has the potential to inhibit advances in a field. Moreover, applications to patent ever-smaller pieces of the genomes of a wide variety of plants and animals are pouring into the Patent and Trademark Office.

If rapid progress in basic science and commercial development is to be fostered, protection of intellectual property should be carefully applied. A balance of interests is necessary between the stimulation of research and innovation through the open exchange of research resources and the promotion of innovation and commercialization of new technologies through patenting.

Such a balance in the life-science context might or might not be consistent with the legal interpretation and implementation of existing patent law.

Recommendation: An analysis of the potential effects of different types of patent protection and the breadth of patent rights on basic research and commercial interests in the life sciences should be undertaken, taking into account societal goals in granting intellectual property rights. Priority should be placed on examining options for protecting inventions that contain nucleic acid sequences, before forthcoming decisions on patents of biologics set precedents that make consideration of alternatives difficult.

MATERIAL TRANSFER AGREEMENTS AND LICENSING

Research scientists have a long tradition of sharing research findings and experimental materials with one another in the interests of collegiality and furthering the scientific research enterprise. However, since the 1980 enactment of the Bayh-Dole Act to foster technology transfer, nonprofit organizations like universities have been obliged to promote the utilization, commercialization, and public availability of inventions that arise from their federally funded research. As a financial incentive, they are permitted to seek and hold rights to the intellectual property embodied in inventions made with public funding.

Research resources, including those described in the scientific literature, are disseminated to interested investigators or organizations through direct transfer or via a third party, usually a licensee that produces and sells the resources to others. These transfers are typically accompanied by material transfer agreements or licensing agreements that are negotiated by the technology transfer offices of the transferring and receiving institutions. As a result, what was formerly a free, open, and rapid exchange of research resources has become an often uncertain, restricted exchange that is subject to protracted negotiations.

Most research resources are innovations of value for scientific investigation, and some have the potential for commercial uses beyond research. By attempting to protect an institution's future financial and other interests in a biologic research resource, the above negotiations can hinder the pace at which the resources are available for use in research. University research faculty's attempts to acquire materials from other academic institutions are hindered by the material transfer process, which requires agreement to assurances that are difficult to monitor in any case, and university officials have not yet resolved these issues.

Faculty attempts to acquire access to resources owned by business organizations can be especially thorny, because commercial operations are bound to the interests of investors, not to the public good. Time-consuming negotiations over access to proprietary research resources are detrimental to academic research, but a private firm cannot usually accept a no-strings agreement. If industry and academe have a compelling interest in sharing their

resources, a genuine effort to develop a mutually recognized and accepted set of minimum requirements could help to expedite future negotiations on research resources.

Tensions that arise during the transfer of research resources can be attributed, in part, to the financial incentives provided to universities by the Bayh-Dole Act. By exercising their right to patent and license their inventions, including research resources, universities generate income for themselves and their researchers. As potential sellers of innovations in the marketplace, however, universities can be viewed as commercial competitors by the business sector—the same business sector whose research resources are sought "without strings" by federally funded, university investigators. The image of the university as a player in the commercial world, and therefore one with which private resources cannot be freely shared, is strengthened by the increasing number of university partnerships with individual companies that often compete with each other.

Finally, as recognized in the National Institutes of Health (NIH) proposed principles and guidelines for sharing biomedical research resources, the financial incentive provided by the Bayh-Dole Act can work against its own objectives and inhibit the dissemination of research resources when universities inappropriately capitalize on the value of a resource. The guidelines note that "restrictive licensing, especially when coupled with indiscriminate use of the patent system, can be antithetical to the goals of the Bayh-Dole Act, such as where these are employed primarily for financial gain" and add that such practices "are likely to thwart, rather than promote utilization and public availability of the invention."

The principles and guidelines proposed by NIH seem to be a constructive step in the right direction. The principles emphasize academic freedom and publication, the appropriate treatment of research tools under the Bayh-Dole Act, and the need to minimize administrative impediments to the transfer of research resources. They also exhort institutions to be mindful of potential conflicts between their obligations to NIH and to other parties that provide research resources, and to establish clear and unyielding policies on acceptable conditions for importing research resources. The guidelines provide specific examples of appropriate language for agreements that accompany the transfer of research materials into and out of universities. If implemented, they could speed the development of material transfer agreements and add certainty to the outcome of such agreements. A copy of the proposed principles and guidelines can be found at http://www.nih.gov/od/ott/RTguide.htm.

Recommendations:

(1) All federal agencies should examine the proposed NIH principles and guidelines and participate in the development of strong and consistent policies across the federal government on acceptable terms for transferring and accepting research resources.

(2) University recipients of federal funds should develop, with input from scientific faculty and university leadership, policies for the identification, valuation, and dissemination of research resources.

(3) Business concerns should recognize their long-term interests in supporting scientific progress and work with universities to determine basic terms of agreement for sharing resources.

(4) An independent and balanced review of the extent to which the financial incentives created by the Bayh-Dole Act affect, favorably and unfavorably, the technology transfer process and the conduct of science should be carried out, taking into consideration the purpose of the Act and the different values and interests of stakeholders involved in and affected by the process.

INTERNATIONAL MATERIAL TRANSFER

At the international level, an issue of concern to scientists who study different aspects of the life sciences is the increasing difficulty of gaining access to wild materials, especially from the tropics, where most of the world's biologic resources exist. The Convention on Biological Diversity, to which many developing countries in the tropics are signatories, recognizes the rights of nations to control access to and to participate in the use of biodiversity resources, particularly the commercial exploitation of native germplasm or local knowledge. Restrictions on exploration of, collection of, and access to information on wild resources have become common, and they affect not only the field work of US scientists, but also the work of local scientists and research institutions. In many nations, there is no clear differentiation between the collection of biologic materials and information for academic purposes and for commercial applications. As a result, every research project (ecologic, systematic, ethnobiologic) is treated as a potential “bioprospecting” agreement.

The US government has sponsored research aimed at involving biodiversity-rich countries in the development of commercial applications derived from native resources, but its ability to negotiate access to biodiversity resources for academic research is inhibited by the fact that the United States is not a signatory to the convention. The creation of joint, basic research programs in which resources can be shared through material transfer agreements that appropriately restrict their distribution or the scope of their application is one approach to this problem. The joint development of mechanisms to document germplasm and other information so that its appropriate and legitimate use can be traced is another.

Recommendation: The federal government should seek discussion with other countries' science agencies to find appropriate terms, which could be applied generally, for the transfer of biodiversity materials for academic research.

DATABASE DEVELOPMENT AND USE

Databases are increasingly critical as research resources, not only for geneticists and molecular biologists, but also for computational and structural biologists, chemists, ecologists, anthropologists, zoologists, botanists, crystallographers, social scientists, and people in many other disciplines. The contents of such databases are as varied as information about rare resources (such as museum and biodiversity specimens and culture collections), DNA sequences, and sensitive identifiers of human subjects. Computer-accessible databases are in increasing demand by researchers of all types.

New scientific discoveries are often based on previously published findings, but data in many fields can be generated so quickly that data "mining" and reanalysis are often as important for the advancement of scientific understanding as data collection in the next experiment. In pharmacology and ecology and in academe, government, and business, the pace of advancement in the life sciences will depend in many ways on access to existing databases as much as the generation of new data.

Indeed, a new and exciting field of scientific inquiry has developed: bioinformatics—the use of computers to manipulate biologic information. With software that permits investigators to query databases in flexible and creative ways, bioinformatics facilitates the rapid and expansive analysis of data. The development of bioinformatics will be a key to using databases fully in the future.

If databases constitute a major leap forward in how scientific information can be viewed and analyzed, new strategies to embrace and take advantage of this power are warranted. Establishing and getting the most out of databases will require investment in the following:

- Their conceptual and physical development
- Data-quality assurance
- Data acquisition and maintenance
- The software needed to operate them.

If a database is to be of maximal value, its potential uses must be reflected in its design. For example, databases need structure and consistency in the variables to be used for sorting or compiling. Inconsistencies in nomenclature (such as species name, symptom, and pathologic condition) make it difficult to analyze large databases that use such characteristics. Even numbers, dates, and geographic locations must be consistent in a database if its utility is to be maximized. Such structure and consistency must be incorporated into the database from the start.

The data quality and quality assurance needed for a database depend on the nature of the analyses being performed. For example, trying to match billions of fragments of DNA to elucidate the human genome requires that there be very few errors in the data. But a few misidentifications of individual

members of fairly common species in a regional database are likely not to affect studies of ecosystem health.

To determine the most useful structure and format of a database, and to establish mechanisms for quality assurance or peer review, potential data contributors and users should be involved in database design. Often it takes a dedicated effort to obtain, via experimentation or monitoring, data with the structure and consistency required by the design of the database, so the community should determine how credit should be assigned to data contributors and whether they have any rights to the use of the data, once deposited. Similarly, rules for the appropriate use of data need to be established, especially to protect sensitive information and personal privacy related to data on human subjects.

The value of a database will also be determined by the nature and extent of the data that it contains. Once a database is created, considerable effort is often needed to maintain it through curation and the addition of data, provision of user support, and the development of software updates.

Finally, being able to glean knowledge from databases requires analytic software. Indeed, the field of bioinformatics involves the development of many sophisticated analytic software tools.

Recommendations:

(1) Databases and the bioinformatics tools needed to analyze them offer an opportunity to gain new insights in the life-sciences and should be considered for increased government and private support.

(2) Before a database is established, data acquisition and maintenance, user support, quality assurance, and analytic software development needs should be carefully considered.

ACCESS TO DATA IN THE PRIVATE SECTOR

Many important databases are being developed by private organizations, especially in the business sector. These databases are often held confidentially and are not available to other scientists except through individual arrangements, some of which restrict investigators' ability to share the results of later work freely; confidentiality is intended to keep competing commercial interests from exploiting investments made in creating the data.

Because of the strong tradition of federal support of the collection and distribution of basic scientific information, access to proprietary data has not been a major issue for publicly funded scientists in the past. But in fields in which scientific and commercial interests overlap, the relevance and importance of the data to new breakthroughs is increasing. Thus, there is a tension between the scientific researchers' need for access to databases and the private database owners' need for confidentiality. That tension is evident in the development of DNA-sequence databases, in which both the nonprofit and business sectors have invested. For example, in agricultural research, the private sector is far ahead of

the public sector in collecting DNA-sequence information about important field crops. Databases of that information could be a great asset to academic researchers studying plant physiology, growth, and resistance to disease and pests.

Although the federal government generally seeks to make public the results of research that it sponsors, it faces a daunting decision between expensive duplication of the efforts that went into obtaining the valuable information that is already in privately held databases and paying substantial costs and acceding to the terms of access to the private data. Public funding in some fields of research is tightly constrained, so it is often questionable whether scarce funds should be used to recreate private-sector databases. Arguments to support an independent, federal database effort or to rely on private data providers should be developed for a variety of data types; scientific considerations and access must be weighed against cost and other factors, such as the effect of intellectual property rights on the material that underlies the data. Options include the outright purchase of access to the data, perhaps leveraged through a public-private database effort; the creation of public-private consortia to develop bioinformatics tools; and the establishment of incentives to share private data with the public. A mixture of approaches that depends on the values and tradeoffs identified might be proposed. Thus, the issue of how best to balance the needs for access to scientific databases and for recognition of the proprietary value of the investments that created them is important and challenging.

Recommendation: Continuing discussion between the various scientific, public, and private interests on the subject of access to and use of scientific databases should be established to promote agreement on approaches that represent the best balance of interests. A candidate for an early topic for such a discussion is access to agricultural genomic data.

Preface

In the fall of 1998, staff of the National Research Council met with the National Science and Technology Council's Subcommittee on Biotechnology[1] to discuss concerns about the scientific community's ability to obtain and share "research resources". Research resources, broadly defined, include plant and animal tissues, reagents, animal models, combinatorial chemistry libraries, drugs and drug targets, clones and cloning tools, methods, laboratory equipment and machines, databases, and computer software. Despite the efforts of federal agencies to promote the efficient dissemination of research tools, problems in accessing resources are increasing in many different fields of science.

Some of the problems arise from the fact that many desirable resources are in private hands, have proprietary rights attached to them, or are perceived as having commercial potential even if they are not yet developed; public and private holders of such resources are unlikely to share them freely. Other resources, like databases, are the collective products of individual scientific contributors, who have a personal stake in how their data are used by others; these resources also face social, cultural, logistical, and financial obstacles that dampen their potential to be used widely as scientific tools.

Earlier in 1998, the National Institutes of Health (NIH) Working Group on Research Tools, chaired by Professor Rebecca Eisenberg, released its report describing the difficulties of biomedical scientists in obtaining research resources. Included in its recommendations were the establishment of a research tools forum, and the development of guidelines for recipients of NIH funds as to

[1]Office of Science and Technology Policy, National Science and Technology Council, Science Committee, Subcommittee on Biotechnology

reasonable terms in licensing and material transfer agreements. These draft guidelines were posted on the NIH web site for public comment in June 1999.

Scientists in disciplines other than biomedicine, with research funding from other federal agencies, are equally affected by issues of access to research resources, especially in obtaining information held in databases of the private sector, and in some cases, by their colleagues in academia. Each scientific field seems to exhibit its own unique hurdles to resources.

Recognizing the need to approach the question of access more broadly, the Subcommittee on Biotechnology, chaired by Dr. Mary Clutter, encouraged the National Research Council (NRC) to hold a public meeting on the spectrum of issues affecting several different research fields.

On January 27-28, 1999, the NRC Commission on Life Sciences organized "Finding the Path: Issues of Access to Research Resources", a conference to explore the breadth of problems and opportunities related to obtaining and transferring research resources. Scientists, entrepreneurs, corporate representatives, university administrators, and government officials attended the conference, which was organized around three panel discussions:

- Issues in Biotechnology and Genomics.
- Issues at the Interface of University, Industry, and Government Policy.
- Issues of Access to Research Resources Across the Disciplines.

The following summary of the 2-day meeting lays out the problems concerning access to research resources as discussed by 2 dozen speakers and members of the audience. Some topics on the conference agenda prompted more discussion than others. The first section of the summary is entitled Material Transfer Agreements because this subject dominated discussion in Panel 2, Issues at the Interface of University, Industry, and Government Policy. The frustrations of bench scientists and industry representatives with university technology transfer offices were voiced strongly by individuals like Harry Klee and Tony Hugli. A group of university representatives collectively articulated how different universities view their institutional responsibilities regarding the transfer of innovations, and Joan Leonard explained federal efforts to study and address the problems. But the long-term implications of federal programs and policies on the multiple roles of universities as educators, research institutions, and users, producers, and commercializers of innovations begged for further discussion.

The second section of the summary is entitled Patents, based on the underlying theme of Panel 1, Issues in Biotechnology and Genomics. Craig Venter, Tom Caskey, and Steve Holtzman engaged in a lively debate on the appropriateness of recent Patent and Trademark Office decisions and their effects on the behavior of industry stakeholders. But Mike Synder made a plea for attention to the issues affecting the "little guy", such as the financial cost of access to critical proprietary and public technologies. Finally, the subject of

access to large genomic databases owned by companies like Celera and DuPont was raised. The discussion was somewhat tentative, because the commercial strategies of these information holders are uncertain, and the impact on scientists in academia and industry, who will most certainly want access to this data in the future, is unclear at this time. Clarification of the stakes of information ownership and access will be critical to the future of biology.

The final major section of the summary, Data Collection and Informatics, captures the individual presentations of scientists from a variety of fields. Many subjects discussed in the context of these fields are novel and reflect recent changes in the organization of how scientists work and how they hope to capitalize on the information and resources they collectively produce. The resolution of at least some of the issues raised in this session hinge on the ability to develop consensus within the respective scientific communities.

In all of the sessions, some speakers offered suggestions for improvement, and their suggestions are included. But if any common theme emerged, it was only that access to resources is indeed a major problem that will take much time, creative thought, and effort to ameliorate. The elucidation of these issues at the conference represents an initial step on the path of identifying first, the challenges, and later, the mechanisms and policies that can produce the rapid and equitable dissemination of research resources on which scientific progress depends.

Contents

Introduction

Science, like any other living, growing thing, depends on and is limited by resources. To do their work, scientists must have access to a wide array of items —computers, basic lab equipment, cloned genes and such expensive machines as synchrotrons and particle accelerators. The pace of discovery depends on the magnitude of resources available for basic research. Decreasing the number of obstacles that scientists face in getting access to these resources speeds up scientific discovery; increasing the barriers slows things down.

Today science is perhaps more successful than at any time in history. Over the past decade a revolution in molecular biology has opened the door to deciphering the genetic code of humans and other organisms—a development that promises world-shaking changes in medicine, agriculture, and many other areas. At the same time the breakneck development of the computer and of information technology has made it possible to gather, store, and analyze huge amounts of data, profoundly transforming research in numerous fields.

Yet with these successes have appeared a number of developments that threaten the very access to resources that has made the successes possible. It is not surprising, of course, that in a time of such rapid change there should be things to learn and adapt to, but a number of the obstacles to resources have proven quite difficult to resolve. If they are not overcome, they threaten to slow the pace of research significantly.

The major concerns that scientists identify fall roughly into two areas. The first set arises from the application of today's incredible computing power to science. How is it possible, for instance, to accumulate vast databases of information about people—their DNA, for instance, or details about their health

and personal habits—without compromising their anonymity and privacy? What can be done to assure that various areas of science will have the proper software to take advantage of the today's powerful computers and digital storage technologies? Can disciplines such as ecology or psychology, which have always rewarded the individual data-taking researcher, find a place in their hierarchies for scientists who collect no data themselves but instead use computers to assemble and analyze multiple data sets accumulated by others, testing broad hypotheses and looking for general patterns? Unless such issues are resolved, the path to new data-rich resources will be winding and slow.

The second set of concerns centers on the increasing commercialization of certain parts of science and the resultant blurring of the distinction between basic research in universities and technology development in industry. Nowhere is this more obvious than in the field of molecular biology, where giant pharmaceutical companies and tiny start-ups alike are pouring money into genetic research in the expectation that it will pay off in new drugs, in faster and more accurate medical diagnostics, in novel medical treatments, and in genetically improved crops and livestock. But a similar thing is happening in a number of other areas, such as microelectronics and materials science, where research findings get translated quickly into commercial products.

As recently as twenty years ago, universities played very little role in commercializing their own research. Typically the government retained ownership of patentable inventions made with federal funds but rarely exercised its rights. Because scientists were, as they are today, free to publish the results of their research, advances in science quickly entered the public domain. This helped keep access to research resources open, in two complementary ways. First, any tools or materials that one university scientist developed were made available to all other scientists. There were, of course, always researchers who would keep the fruits of their work to themselves in order to maintain a competitive edge over their peers, but this was frowned upon. Sharing was the norm. Second, as long as university researchers were making their findings freely available to everyone, industry was generally willing to provide these basic scientists with research resources at little or no cost, on the theory that the results of the research would ultimately benefit industry.

But the Bayh-Dole and Stevenson-Wydler acts of 1980 and the Technology Transfer Act of 1986 changed the rules, allowing universities to hold the rights to patents on innovations developed using federal funds. Universities have since plunged into the commercial world, licensing the research of their scientists to private companies. The goal of Congress in passing the acts was to encourage industry to exploit the research coming out of universities, and in this the acts have been successful, but they have also intermingled the interests of academia and industry to an unprecedented degree. Because of this intermingling, scientists are finding that access to research resources is often much more complex and frustrating than it has been in the past.

To discuss those various concerns, the National Academy of Sciences and the National Research Council held a meeting, “Finding the Path: Issues of Access to Research Resources”, on January 27-28, 1999. As conference chair David Galas described it, the meeting was intended “not to come to any sort of consensus or do an in-depth analysis of these issues, but rather to get the range of issues, as they exist today, on the table, to try to focus some of these issues, gather the opinion of the speakers, and be informed by the speakers about them, and then attempt to sharpen the focus by our discussion and, where it's possible, to interrelate the issues.”

This summary of the conference has been divided into three chapters on material transfer agreements, on patents, and on data collection and informatics. These represent focal points to which are attached many important issues about the practices of stakeholders of research resources and the policies that shape their motivations.

1

Material Transfer Agreements

Researchers today, particularly those in molecular biology, face a world quite different from that of 20 or 30 years ago. “I was one of the early molecular biologists,” noted Bruce Alberts, president of the National Academy of Sciences, in his introductory remarks. “I worked for 30 years with the bacteriophage T4 that Max Delbrück introduced as a model organism. The spirit that he promulgated in my field was one of complete sharing of ideas and resources, and at that time there were frequent Cold Spring Harbor meetings where everybody laid out their latest data and emptied their notebooks, with no idea that anybody would ever think to steal an idea or claim credit for something they didn't deserve. This was before the biotechnology revolution, before there was any idea that you could become wealthy or start a company and that there could be any major commercial value to what we were doing. We thought about it in terms of new developments for medicine and doing good for people.”

But commercial concerns, Alberts commented, have chased away much of the openness of that earlier era. Now researchers and their employers, whether in private companies or in universities, must constantly be careful about their intellectual-property rights and so are often wary of passing on their research materials without securing some sort of protection. The result, said Joan Leonard, vice president and general counsel of the Howard Hughes Medical Institute (HHMI), is a tremendous loss of time and money in the resulting legal byplay.

“The paradigm transaction,” she said, “is that our investigator wants materials from a company. The company says, ‘Fine, we'd love to give them to you; just have this agreement signed and we'll be happy to send them.' And then the agreement comes to me or one of the people who work for me. And maybe it

has provisions that limit publication, or provisions that seem to have overbroad confidentiality requirements, or provisions that say, 'Oh, by the way, we own everything you ever do that gets anywhere near the reagent we're sending you,' and so on.

"And so discussions ensue. This is not appealing to us or acceptable to us on a first reading, and maybe we work it out and maybe we don't. If we do, the material comes to the investigator—but after a great deal of time is lost and a great deal of high-priced talent has been used to look over and wrangle over some of these issues."

Today, whenever any sort of research material—reagents, cloned bits of DNA, genetically modified mice—changes hands between institutions, it is common to sign a material-transfer agreement (MTA). An MTA governs how a researcher can use materials and what obligations attend their use. The obligations can range from promising not to pass the materials on without permission to signing over all rights to commercial development of any discoveries made with the materials. MTAs have become perhaps the largest obstacle that molecular biologists face in gaining access to research resources, and they are playing an increasing role in other fields of science.

A number of the researchers at the conference expressed their exasperation with MTAs and with the technology licensing offices whose job it is to oversee and implement them. "The first agreement that I tried to do through the university technology licensing office took 6 months and it essentially, in the end, went through signed pretty much as it came to us," said Harry Klee, of the University of Florida. I had been told that this had low priority and they'd get to it when they could get to it. That is not unusual, he said. "I think that the technology licensing offices in most universities are woefully ignorant of the system and are woefully underfunded and understaffed."

Nor was it just the researchers who were unhappy with MTAs. Even Leonard, who is responsible for negotiating the agreements at HHMI agreed that they pose a problem. "It's a drain in terms of time and administrative funds," she said. "There is also tremendous delay. For the scientists, this can be critical, particularly for postdoctoral fellows who are under pressure to get things done and start their careers."

Some MTAs cause more difficulties than others. Lita Nelson, director of MIT's technology licensing office, noted that agreements between universities are often painless. Most major universities and many of the smaller ones, she said, use NIH's uniform biologic materials MTA or its equivalent when sending materials among themselves. It is a simple one-page agreement; if both universities have agreed to use it, Nelson said, "we will accept the signature of the professor—no strings." Transferring materials from university to industry is also straightforward most of the time, she said. In most cases, the university has decided to license the material nonexclusively for a fee. "Usually, it goes very quickly."

The trickiest problems arise when the transfer is from industry to academe; private companies often make demands that researchers—or their technology licensing offices—balk at. A company might, for instance, ask researchers to hold off in publishing their results to give it a head start in applying the results. Or it might insist on rights to an exclusive license on any invention or discovery made using its materials.

Klee, who worked at Monsanto before moving to the University of Florida, said that the company's MTA required four things of scientists: “First, we asked that you not transfer the material or a derivative thereof to third parties without written approval. Second, you had to use it for research and not for profit. Third, of course, Monsanto was not legally liable for anything that you did with the material. Fourth, we wanted the right of first refusal; we wanted to be able to negotiate with you in good faith that you would offer us a license to anything that you invented with it.”

The Monsanto agreement is relatively simple as industry MTAs go—many companies include much more in their agreements, such as restrictions on researchers' publication of their results—but even the simplest proposed agreements often result in long delays as the university technology licensing office negotiates the details. “There is a tremendous administrative burden on both sides,” Klee said. “I saw many of these agreements get bogged down, most often on the university side, because people disagreed with some words in the claims. Those claims usually had to do with liability. Many times, the universities would not sign the agreements at all, and I had to tell people that they could not receive my materials, because their university would not sign an agreement.”

From the university's point of view, the greatest stumbling block is often the request for “reach-through rights” on inventions that come about through use of a company's materials. Like Monsanto, many companies request the right of first refusal to license any discoveries or inventions. That has become common only recently, said Candace Voelker, of the Office of Technology Transfer at the University of California. “It's only within the last 5 years or so that companies have been tapping rights to their materials. Before, they would grant the materials to the university faculty without such strings attached, and it didn't come up that often.” The insistence on exclusivity poses a concern for universities, Leonard said. “If you're granting an option to a company in exchange for a research tool, you can sell that particular horse only once; if the research project is going to require another tool that requires an exclusive license, you have a serious problem on your hands.”

The obvious solution would seem to be to negotiate shared rights, with two or more companies providing research resources and each being able to commercialize technologies arising from the research. But, Voelker said, industry has little interest in that. “I have a company, a licensee, that is willing to share rights with other collaborators of my inventors, but we've never found another company that would be willing to share with that licensee. We went to

the company and said, 'Even though you are funding and you have these rights, this research is at a standstill unless we can get to collaborate with another company, so would you be willing to share co-exclusive rights?' The licensee said yes. We asked three companies to work under those circumstances, and all three said no."

The refusal is understandable, said Thomas Caskey, of Merck Research Laboratories. From industry's point of view, such rights-sharing agreements complicate matters tremendously in that there is no single owner of the resulting technology. "I'm not saying that you can't do it. It's just that it has to be looked at much more carefully."

If academic researchers want a particular resource from industry, they are often faced with the choice between signing an exclusive-rights agreement and going without. "It's unfortunate, but that's the way it is," Klee said. "We do want to distribute the material, but if you can't offer us, the company, anything in return, you shouldn't expect to get the material."

Obtaining resources from universities—particularly items with a strong commercial potential—has also become more of a problem in recent years, although it is still generally simpler than working with industry. Universities are under pressure on a number of fronts to get some return from their research activities. The Bayh-Dole Act, passed by Congress 20 years ago in an effort to see that valuable research is developed, allows universities to patent the inventions of their scientists and to offer exclusive and non-exclusive licensing agreements to companies; since its passage, universities have become increasingly aggressive in seeing that their research is commercialized. Many state governments are pushing universities to turn their research into economic development that will benefit the states. And, more generally, the attitude toward commercialization is changing at universities: many are coming to value the return on investment that their research can bring. Dennis Stone, vice president for technology development at the University of Texas Southwestern Medical Center, spoke of the need to obtain "appropriate value for university-developed technologies." The money earned from these technologies serves many purposes, he said. "It allows investigators to work to make certain that inventions are developed as quickly as possible. It provides income to the university that can be used to do more of the same type of research or a different type."

And so it is that universities, in attempting to protect their interests, sometimes end up going down the same path as private industry, demanding restrictive MTAs on their most valuable technology. Or they sell or assign rights to the product of their research to a company—sometimes one started by the university scientists who performed the work; and that restricts access to the product.

The micro-array developed at Stanford University is perhaps the most talked-about example. It is a device for testing the activity of various genes in a sample taken from an organism. Because it can test thousands of different genes

simultaneously and offer a measure of which ones are more active and which less, it has many applications. “This is a very powerful technology,” said Michael Snyder, of Yale University. “It would be of enormous use to the entire scientific community. Virtually everyone I know is champing at the bit for this technology.” But it is not broadly available, because Stanford patented it and sold an exclusive license to one company, which operates the micro-array technology as a service rather than providing it to others.

Ultimately, many of the obstacles to obtaining research resources have their roots in the difference in goals and attitudes between the suppliers and the users of materials. These differences can make it very hard for the suppliers and the users to agree on what is fair in an agreement to supply research materials.

When an NIH Working Group on Research Tools studied the issue, it found three overlapping ways in which perspectives on the problem diverged, said Leonard, who served on that working group. The first was the split between user and supplier: One person's tool is another person's product. The dilemma is particularly acute for a small company that was built for a single technology, Leonard noted. It is fine in theory for a researcher to argue that he should have access to a small company's technology because, in the long run, the free dissemination of research tools will benefit everyone, but to the company, that research tool is its only way of making money, of surviving. If the company is to extract the value of the technology, Leonard said, it must get something from researchers—“either pay me now, in a large up-front payment, or pay me later through reach-through rights.” Then again, from the investigator's point of view, the technology is just “one of a number of resources that is going to go into a large and complex course of research,” so it is hard to justify mortgaging the entire research project for one tool, no matter how useful. “That polarity is very hard to reconcile and to find common ground on,” she said.

A second related issue is the difference between providers and users in valuing a product. Consistently, Leonard said, providers tend to overvalue their product, and users undervalue it. It can be hard to meet in the middle, particularly because assigning a value to one piece of an entire project, whose ultimate value might not be known for years, is inherently subjective. “There are very few conventions out there for valuing these resources.”

Finally, there is the difference in fundamental missions between suppliers and users. “Universities are engaged in the creation and dissemination of new knowledge,” Leonard noted. “That is their legal and traditional obligation and mission. Companies are in the business of raising money to develop useful products and to extract the value of the products for their shareholders. That is also a legal and traditional obligation. When those two things collide, it is difficult to find common ground.”

To complicate matters, the perspectives can switch, depending on roles. A university attempting to commercialize its intellectual property might, when it comes time to provide materials to researchers elsewhere, view things more from the perspective of industry. And, Leonard said, “when you look at industry,

there are an awful lot of fundamental scientists doing basic research, and they don't see any reason why they should be treated differently from their academic colleagues."

None of those issues will be easy to overcome, but the conference participants offered suggestions for improving the transmission of research resources among universities and private companies.

"We have learned a lot of lessons," said Maria Freire, director of NIH's Office of Technology Transfer. "Bayh-Dole is 20 years old, and perhaps some of the deals that we cut earlier we would not cut now." At the recommendation of the Working Group on Research Tools, the NIH is distilling those lessons into a set of draft guidelines designed to help universities and NIH employees determine what is best practice in negotiating MTAs. Simply circulating the guidelines and getting the people responsible for technology licensing at universities to read them should help smooth out the dissemination of research resources. Many institutions are still new at licensing technology and are making the same mistakes that others have learned, through their own painful experience, to avoid. In particular, the guidelines advise universities to make sure that they do not compromise their researchers' ability to publish their results and advise them to avoid reach-through agreements whenever possible.

A second suggestion was that universities avoid legal agreements altogether when the materials in question are unlikely to have any commercial value. Perhaps 95% of researchers' complaints about MTAs concern "transferring materials that have nothing to do with patents and licensing," said Tony Hugli, of Scripps Research Institute. Offices of technology licensing, he said, should listen to researchers' opinions of the commercial value of their materials and insist on agreements only when there is some chance of payoff. The NIH Working Group came to a similar conclusion, Leonard noted. "There is little to be gained and much to be lost in efficiency by going through the process of having agreements, even if they are relatively simple to negotiate," she said. "It's a burden to the system that we don't need."

Another suggestion was that universities try to commercialize their technologies in ways that ensure access for researchers. In particular, universities should try to avoid repeating what Stanford did with the micro arrays. But that is not always possible. Stanford, for example, could find no one to develop the technology on a nonexclusive basis, because it demanded a great deal of work to bring it to market. So Stanford was forced to go with the exclusive license and with a company that did not make the tool easily accessible to researchers.

Industry, too, should be able to improve how it deals with academic researchers, said Steve Holtzman of Millennium Pharmaceuticals. "For those who have been in this for 15 years, it gets a little disheartening sometimes because the same issues are playing over and over again. You have to recognize that there is a basic standard form that you can use for 95% of the cases." Out of responsibility to their shareholders, private companies must demand some things

when sending out their proprietary materials, he said. "You need a royalty-free license to improvements to what you gave them, and you need a royalty-free right to practice commercially any new inventions—you can't enable someone to block you from exploiting your own technology. Universities should be able to accept those provisions." Conversely, industry has to understand that academic researchers have their own imperatives. "It's the mission of the university to disseminate knowledge. Part of the price you pay to work with academe is that it is going to publish—that is not debatable. That people are still negotiating publication rights is nuts."

Indeed, Holtzman said, only one issue in industry-university negotiations should cause difficulties, and that one is unavoidable: "Where the rubber hits the road is in the rights to new substances created with the material in question." The company wants to protect its investments by getting exclusive rights to future inventions based on its research resources. Universities want to maintain their freedom of future action by not granting exclusivity. Each must ask itself how much flexibility it can afford and work from there. "You can't come up with a general guideline for how you're going to deal with new inventions that use research material," he said. "It ineluctably involves judgment." In other words, there will not be any easy answers.

BOX 1

A MODEL FOR UNIVERSITY INDUSTRY COLLABORATION?

In the midst of the conference's generally gloomy assessments of resource sharing, one bright spot appeared in the description of the agreement between DuPont and the NIH over the use of cre-lox technology in mice.

Cre-lox is a method for creating mice and other animals or plants in which a stretch of DNA is removed from particular cells. Researchers use the technique mainly for studying gene function; they remove one or more genes and observe the results. Cre-lox has become a key element of the molecular-biology toolkit.

Several years ago, however, DuPont decided that it should put conditions on the use of the technology. Until then, researchers had been disseminating and using cre-lox without asking permission, even though DuPont held a valid patent on the technique. The company's first idea, said Maria Freire, director of NIH's Office of Technology Transfer, was to ask universities for a cash payment if their researchers were going to use cre-lox. DuPont was seeking far less than the $100,000 that some commercial outfits pay for cre-lox, but the universities balked. "At most universities," she said, "the first reaction is, 'Oh, my god, not $5,000 a year; that's a lot of money.'" So the company decided instead to ask for reach-through rights, to get a share of any products developed at the universities by using the cre-lox technique.

Several dozen institutions signed such agreements with DuPont, but NIH did not. "When DuPont came to us with that scenario, we were not very happy," Freire said. NIH Director Harold Varmus worried that such an agreement

would impede basic research, and indeed one major problem was already apparent. The Jackson Laboratory, the world's largest supplier of mice for use in research, was refusing to stock or distribute cre-lox mice as long as DuPont insisted on reach-through rights. So NIH set out to negotiate a better understanding with DuPont.

"It was a very difficult negotiation," Freire said, but at the end, NIH had persuaded DuPont to let researchers continue using cre-lox technology in mice and other animals with essentially no strings attached. NIH researchers can use cre-lox at no cost, as long as it is for research purposes only, and can transfer the materials to other researchers with the standard NIH material-transfer agreement. They can also transfer cre-lox materials to researchers in industry, but if they do, they must tell DuPont and they must apprise their industry partners of DuPont's intellectual-property rights to the cre-lox technology. There are no limits on publication of results of research done with the cre-lox technology. Most important, DuPont has no reach-through rights on any discoveries or inventions made at NIH with cre-lox technology. Furthermore, the agreement applies to both researchers at the NIH and those in academia.

In return, NIH agreed to make its cre-lox research materials available to DuPont when requested. Otherwise, DuPont gets little from the agreement other than the knowledge that it is helping move science forward. The company does retain patent rights on commercial uses of the cre-lox technology, and the research should help make the technology more valuable; but DuPont gets no direct return from the use of cre-lox in basic research on animals.

The agreement does, however, prohibit the use of cre-lox to make libraries of mouse embryonic stem cells. And the agreement is not available to plant researchers, noted Barbara Mazur, of the DuPont Agricultural Products Enterprise. The reason is that the genes used in cre-lox technology are included in some of its patented agricultural products, so the agricultural side of DuPont sees cre-lox as a product, as well as a tool, and is consequently less willing to allow its use even for basic research without demanding compensation.

Several speakers suggested that cre-lox might offer a model for university-industry cooperation on guaranteeing access to such basic research tools as cre-lox. Indeed, at the time of the conference, NIH was negotiating with DuPont in an effort to sign a similar accord regarding use of the Harvard oncomouse, a mouse valuable in cancer research. But it is not clear in how many other cases agreements like this can be signed. A giant company like DuPont can afford to allow research use of its material without payment, but a smaller company might not be able to; and even DuPont drew the line at allowing free use of cre-lox in agricultural research, where its interests were more directly affected. But at least in limited cases, such as the cre-lox mouse and the oncomouse, it might be possible to ensure researchers free access to these basic tools.

2

Patents

Few things have greater potential to influence access to research resources than patents and the policies that determine when they are awarded. Patents are intended to encourage innovation by guaranteeing that innovators are appropriately compensated for their work. Patents are also intended to encourage openness. Without patent protection, inventions and discoveries with commercial potential would often be kept as trade secrets. The patent system offers the inventor legal ownership of an innovation in exchange for putting a description of it into the public record.

Patenting an invention keeps people from using it without paying for it, or at least asking permission, but it does not mean that the invention cannot be used at all. “There is a big difference between being available free and being available,” Craig Venter of the Celera Corporation noted. Indeed, if an invention demands expensive development before it is useful to anyone, patenting the invention might be the only way it ever becomes accessible.

Nonetheless, patenting can slow or stop access to some innovations, particularly basic discoveries and inventions that are of value to researchers on the leading edges of their fields, so scientists are especially sensitive to patent policy regarding this sort of fundamental work. Perhaps the best example is a debate that has been roiling the molecular biology research community for some 7 years. The resolution of the debate will come not from the scientific community, but from the US Patent and Trademark Office (PTO), patent attorneys, courts, and perhaps Congress. However, it will be scientists who are most affected by whatever decision is reached. And not just molecular

biologists; the decision of where to set the threshold for patenting could have repercussions for other fields, such as synthetic chemistry.

The debate commenced in 1992 when PTO rejected a patent application from NIH for thousands of expressed-sequence tags (ESTs). ESTs are short stretches of DNA that can be isolated rapidly and in great number and used to identify genes. They are, in effect, gene fragments that are long enough to be characteristic of particular genes but short enough to be manipulated and cloned easily with the standard tools available to molecular biologists. PTO refused to approve the EST patent application on the grounds that the function of the DNA was not known. But the matter did not end there. Private genome companies have since filed patent applications for hundreds of thousands of ESTs. One goal is simply to establish priority, to establish that a particular company was the first to isolate a particular gene, even if only a very small part of it, and even if the company had no idea what gene it was when it filed the application. But another goal is to gain patents and thus establish an interest in whatever else is done later with the genes identified by the ESTs.

More recently, companies have been rushing to patent single-nucleotide polymorphisms. Differences in the human population are produced by variation in the nucleotide sequence of their DNA, which is composed of four types of bases —adenine, guanine, cytosine, and thymine—A, G, C, and T. If the DNA sequences of two individuals are compared, or if the maternally and paternally derived chromosome pairs of one individual are compared to each other, there will be differences. If at a particular position on the DNA sequence the maternal chromosome contains a ‘G' while the paternal chromosome contains a ‘T', each surrounded by otherwise identical sequence, that difference is called a single nucleotide polymorphism—a SNP.

SNPs (pronounced “snips”) can be found on average every few hundred nucleotides across the entire human genome (of a few billion nucleotides) so there are literally millions of SNPs that can collectively distinguish each of the 42 human chromosomes carried by different individuals. Those differences in sum produce a fingerprint, a ‘SNP Map', which can be used to trace patterns of inheritance of disease predisposition genes that contain, or are closely linked to, particular SNPs. It is the potential for mapping complex characteristics by variation among individuals in their genome-wide SNP maps that is engendering considerable excitement, and commensurate concern about access to this powerful research material. Although much effort will be needed in the future to understand the role of SNPs in relation to disease, SNPs, like ESTs, can be identified now with relative ease using the proper tools, with little effort beyond the initial set-up.

It is precisely the latter characteristic that raises questions about whether ESTs and SNPs should be patentable, said Steve Holtzman. The power of modern molecular biology has automated this particular type of discovery to the point where a laboratory with the right equipment can, for example, uncover thousands of ESTs a day. Traditionally, however, a patent has been issued for

something that demanded creative and original work; there, Holtzman noted, is the rub. “Do you feel a cognitive dissonance there—‘automated discovery and invention'? What does it mean to have 10,000 creative, original, useful ideas a day? I think there's a problem here, and it has to do with conceptual impossibility.” Certainly, the originators of our patent laws never conceived that they might be applied in this way, but the patent system has evolved in such a way that granting patents to such automated discoveries is almost inevitable.

Venter, Holtzman said, “came up with the brilliant idea of ESTs. I submit that that was the invention, the method.” The individual ESTs churned out by the thousands are not the invention, he said. “But that's not consistent with patent law. Patent law says that the composition of matter is the embodiment of the idea, and you get the grant on the embodiment of the idea, on the composition of matter. Furthermore, case law has made it very clear—and this is intuitively true—that the method of invention doesn't compromise the invention.” In other words, it does not—and it should not—matter whether you worked for years to identify a single EST or pulled it out of a machine with thousands of others.

But if such automated invention is patentable, it will be possible for a few companies with a lot of money and a lot of machines to lay claim to huge areas of knowledge before these areas are ever explored more than cursorily. It is not just ESTs and SNPs that are in play, Venter argued, but whole genes as well. Nor does the problem stop with molecular biology. There are perhaps 100,000 small molecules—with various combinations of carbon, hydrogen, oxygen, nitrogen, and a few other atoms—that could be useful as drugs, and combinatorial chemists are rapidly—and automatedly—trying out all those possible combinations. “People are getting patents granted daily right now that claim thousands and thousands and thousands of those molecules.”

This sets the stage for so-called submarine patents—patents granted on the composition of ESTs, SNPs, genes, or small molecules—that will one day surface to exert their ownership rights when discoveries are made by others about the function of the these biological and chemical entities.

Not only does that seem unfair and contrary to the original intent of the patent system, but it also makes researchers and companies less willing to invest the effort in pursuing discoveries that could be waylaid by submarine claims. So far the PTO has awarded one EST patent. It is a broad patent that claims rights to any gene that contains the EST, even though the gene is not yet known. Venter argued that the patent is probably not worth the paper it is printed on, because “the patent had very low accuracy data and lousy informatics and claimed things that don't really relate to the sequence.” Holtzman countered, “It might be not worth the paper it's written on, but it will cost you several hundred thousand dollars to litigate. So it's worth something.” It is, if nothing else, a disincentive for anyone besides the patent holder to pursue the gene that contains the EST.

Thomas Caskey agreed. “I can tell you that this is a major issue right now, because broad claims are being filed. The patents are being awarded; therefore, using that gene in the commercial world for any type of utility gets greatly complicated because the corporation lawyers will tell you that there is great jeopardy in proceeding with the use of that gene if you do not have clearance.”

Nonetheless, under existing patent law, the ESTs and similar “discoveries” are patentable. The most relevant issue affecting their patentability is their utility—PTO demands that any invention or discoveries have some potential use if it is to be patented. But even if the biologic function of a gene, SNP, or small molecule is unknown, it can still have some utility.

Consider the SNP, for example. Celera is planning to sequence five human genomes, Holtzman noted. “By definition, they will have 80% of the SNPs with a 20% prevalence in the population. Do they have utility? Yes. They are mapping reagents.” That is, they offer a way to distinguish one person's DNA from another's, which can be important, for instance, in tracking down the genes that cause a disease. “Any utility is sufficient for a patentability claim, said Holtzman, adding sarcastically, “All of us have thousands of SNPs. We could file on them tomorrow.”

The situation for small molecules is similar. When researchers look for drugs that will attach to a particular protein in the body, they often start with one drug that works to some degree and comb through others with similar structures, looking for the one with the best performance. Thus, it is useful to have libraries of small molecules available for searching. “They have utility,” Holtzman said. “Someone will buy those libraries.”

The best way to proceed, Caskey suggested, might be to finetune how PTO awards patents. “We're not going to reverse patent law, so my simplest solution to this would be to ratchet up the specificity and the demonstration of the specificity of utility,” that is, require the patent applicant to demonstrate explicitly the utility of the invention and then grant the patent for that utility alone. If the only known use for a particular small molecule is as part of a library of similar small molecules, grant a patent for that use; if someone later discovers that the molecule is a useful drug for schizophrenia, allow another patent for that specific use. Then, Caskey said, “as you move down the pathway, there is an opportunity for protection with new discovery to be able to go forward to the utility.”

Holtzman did not think that remedy would work. “In an ideal world,” he allowed, “the grants of the composition-of-matter patent would only go along with ‘real utility.’” But, he said, he has little faith in the ability of PTO, Congress, or other institutions to get it right. “PTO was operating that way and basically saying with respect to drug molecules, ‘Do a Phase III study and get it registered by the Food and Drug Administration, and then you will have shown its utility.' It backed off from that 2 years ago.” It had set the bar so high for a patent that it was creating more problems than it had solved.

A different approach that many have suggested for the genome field would be for the federal government and other entities to pay for putting as many data into the public domain as possible, thereby making the data unavailable for patenting. In 1994, for instance, spurred by worries that much of the human genome sequence data would be patented, Merck set up a project to sequence human DNA and deposit it in Genbank, the public database of genes run by the National Library of Medicine. And the NIH National Center for Human Genome Research encourages the researchers who receive its grants to deposit their DNA sequences into a public database as quickly as possible and not to seek patents on the sequence information itself.

That does not mean, however, that discoveries made about the meaning and usefulness of raw genetic data would not be patentable. “We don't think by publishing the human chromosome sequence itself we're blocking others from making the kind of key discoveries that have been talked about,” Venter said. “My understanding is that publishing the human chromosome sequence itself will have no impact on cDNA or protein patents.” For example, Venter has filed for a patent on a human gene that codes for a new serotonin receptor; he found that gene by searching through a human chromosome sequence that had been deposited in Genbank. “The best patent-attorney advice we can get is that those should be valid claims.”

Furthermore, Caskey noted, each step in the process of moving from a gene or protein to a commercial product offers its own opportunities for patent protection. “If we want to try to keep the roadway open for discovery,” he said, the information that underlies everything else—the sequence data—should be in the public domain, and patents should be reserved for discoveries and inventions that go beyond that basic information.

Most recently, Celera, Venter's company, has announced plans to determine all the 3 billion base-pairs of the human genome and put that sequence information on the Internet, making it freely available to anyone who wants it. If all goes as hoped, Celera will have the entire human genome finished by the end of 2001. As a warm-up, the company plans to sequence the smaller genome of Drosophila melanogaster, the fly used by many biologists as a model organism, and make that sequence information available by the end of 1999.

Once the entire human genome—and several other genomes as well—is available, those discoveries and innovations should explode, and Celera plans to profit by serving prospectors who are looking to mine the various genomes. Although it will place no restrictions on the use of its data—anyone can download unlimited human sequence information free, freely distribute it, and never pay royalties on any inventions stemming from it—Celera will have developed a resource that no one is likely to duplicate. “If somebody were going to try to duplicate the data center that we're building, it would cost around $60 million just to build the housing for the hardware, and then they would have to pay for data generation and support.” Thus, people who want to work with the DNA data will come to Celera, Venter said.

Venter views the commercial strategy of on-line information providers like Lexis-Nexis, (which maintains large databases of government publications, legal case histories, and more), as the model for Celera. Once such a vast amount of data is compiled, users will want convenient access to this comprehensive body of information, and be willing to pay for it.

"Pharmaceutical companies are paying pharmaceutical prices for unique early access," he said. "That's at one extreme. The other extreme is that the data will be on our Web site and freely available to academic researchers. In between are the people who want the added value of all the comparative data and all the other information; we're thinking of subscription prices of around $5,000-20,000 a year for research laboratories, which would be compatible with what academic scientists are paying for other research tools and software systems."

Not everything Celera produces will be available, however. In particular, although the company does not believe that sequence data should be patented, it is taking a wait-and-see approach to its SNPs database. "We're eager to see what happens with intellectual-property protection on polymorphic variations," Venter said. "I think they're very important for screening and for a wide variety of tools. What's driving it, obviously, is the pharmaceutical industry, which wants to save billions of dollars off the cost of developing drugs." With billions of dollars in play, Celera will, at least initially, keep its SNPs as trade secrets, Venter said. "But because the basic data are accessible to subscribers at reasonably competitive rates, we think that they will actually be broadly available to a wide array of scientists."

That, to Venter, is the bottom line—not whether something is patented, but whether it is accessible. In discussing patents, he said, many people forget that the patent system makes possible commercial development of many of the research resources that scientists depend on. "If you look at science in this country versus in, for example, the former Soviet Union, the biggest difference is that we have tremendous industry support for what we're doing"—and this industry support is available because the patent system guarantees that companies can profit from developing such things as restriction enzymes and other tools that researchers use, as well as drugs and other products. "If composition-of-matter patents are denied on DNA, it would certainly affect the Amgens of the world and the Genentechs that make incredibly important drugs that have saved millions of lives in this country."

Instead of getting hung up in an emotional debate about whether it is right for someone to "own" our genes, Venter said, the better approach is to ask how well the patent system is working to promote access to scientific discoveries. There might not be agreement on the answer, but there should at least be agreement on the question.

BOX 2

THE HIGH COST OF HIGH TECHNOLOGY

In addition to intellectual-property issues, a major obstacle to research resources for many scientists is cost. Research tools can be divided roughly into three groups, observed Michael Snyder. “There are cheap technologies, such as restriction enzymes and the polymerase chain reaction (PCR), that spread throughout the community like wildfire. They get invented, and even if you try to stop their unlicensed use, you can't. “On the other end of the spectrum are technologies, such as linear accelerators and space stations, that are so expensive that we can afford to build only one or a few of them, and everyone must share. “And then there are the technologies in between, such as confocal microscopes, automatic sequencers, and micro-array technology. These are fairly expensive. Individual laboratories—certainly individual academic laboratories—can't go out and buy these easily.”

The technologies in the middle group are extremely powerful, allowing researchers to do things that would otherwise be impossible, so it is important to make them as widely accessible as possible. Yet, Snyder said, “some of these technologies aren't being distributed well.” For instance, he looked at 28 recent papers by US scientists describing research that used micro-array technology. “Of those 28, 24 came from very large research centers, companies, or a very well funded laboratory. Only four came from what I saw as academic laboratories. So we have a situation of haves and have-nots. A few laboratories are using the technology or can collaborate with people who can do it, and there are a lot of people who want access but cannot get it.” For such situations, Snyder said, it would make sense to set up “minicenters” around the country that make a technology accessible to a much broader range of researchers. “Twenty small centers could blanket the country,” he suggested. “I don't think it would be very expensive to set up this particular kind of technology. For example, $200,000 would certainly cover the cost of one center for micro-array technology, and with matching funds, the center could be even more productive. If you had 20 centers scattered in various geographic locations, you could get all this technology out there for relatively modest cost. Four million dollars is pretty modest when you think about the impact that this technology has on things and how much it cost to invent it in the first place.”

The same argument applies to a number of research tools, Snyder said. “If you think of the nature of science now, a lot of these technologies are coming out of big laboratories or big centers, which devise expensive technologies that individual investigators can't afford.” To make the most of these technologies, “we'll need groups of people to have access, and I suggest that minicenters would be wonderful avenues for dispersing useful information and technology.”

3

Data Collection and Informatics

The computer revolution has given researchers new tools and capabilities. One of the most important is the ability to collect huge amounts of information and manipulate and analyze it quickly and in great detail. This data-handling power has speeded up many of the tasks of the scientist, from data acquisition and analysis to communicating with other scientists. More important, it has allowed researchers to generate hypotheses, perform experiments, and analyze mountains of data in ways that would not even be conceivable without computers. Entire new lines of research have opened up as a result.

Consider the development of the Protein Data Bank (PDB), which contains detailed information about the structures of proteins. Since 1971, when it opened, the PDB has grown from an initial seven protein structures to more than 9,000, said Helen Berman, the data bank's director; in the process, it has evolved into far more than just a way for protein crystallographers to make their structures available to other researchers. “When a large data set became available,” she said, “people began to do comparative and integrative analyses; as a result, they developed a new field of protein-structure prediction, which, in turn, has led to the field of structural genomics, which is giving much more work to the structural biologists to determine new structures.”

In fields as varied as genomics, psychology, chemistry, and archaeology, researchers are coming to see the value of such large databases and, in many cases, finding that they cannot do their jobs without them. But dealing with these huge collections of information is not easy, or inexpensive, and researchers face a variety of obstacles in assembling and using them. Cost is a constant concern, particularly in fields in which databases are not traditional

tools and the funding agencies do not yet value them as highly as they do other research tools; but the other hurdles can be even more vexing than the lack of support. The issues, as identified in the forum, touch on every aspect of databases, including collecting the data, working with them, and disseminating them.

Each field of research that works with databases has its own unique issues, but, as was clear from the presentations at the conference, some issues are common to many fields. The need for software is one such common theme. Researchers often find that little commercial software is suitable for their needs, but there is little funding and little professional reward for scientists who take time out from their own research to write the complex programs needed for work with databases. A second recurrent theme is the challenge of transforming decades of existing data, usually collected in a wide variety of noncompatible formats, including print, into a form that can be deposited into a single database. A third concern centers on getting permission from relevant parties to put data into a collection and then regulating its use and exploitation, scientifically or commercially. Until those various complications can be settled, researchers will not be able to benefit fully from the vast potential of their databases.

PROTEIN CRYSTALLOGRAPHY

To understand how proteins function, which is crucial, for example, for rational drug design and investigating the etiology of various diseases, researchers must learn what the proteins' structures are—how the molecules' carbon, oxygen, nitrogen, hydrogen, and other atoms arrange themselves. To perform this mapping, researchers must crystallize a protein, expose the crystals to intense radiation, and measure the diffraction pattern formed when the radiation passes through the protein crystals. Analysis of the diffraction pattern provides information about the positions of the atoms—information that researchers can combine with other data, such as the sequence of amino acids that make up the protein, to infer the protein's three-dimensional structure.

The rate-limiting step—that is, the step that determines how fast the entire process can proceed—is the accumulation of diffraction data on the protein crystals. It demands large, expensive machines that can supply an intense, focused beam of radiation. Of these machines, synchrotrons are the most expensive and the most desirable because they offer the most intense radiation. According to Vladek Minor, a protein crystallographer at the University of Virginia, 70% of the protein structures published between June 1997 and May 1998 depended on synchrotron radiation.

As might be expected, there is intense competition for access to the machines. Most of them are government-funded, and time on them has traditionally been allotted according to the results of peer review of researchers' proposals. Recently, though, consortia of users and sometimes even individual

users have been buying time on the machines and on new commercial synchrotrons, and this has reshaped the access to this important resource. Researchers are now faced with several choices. They can get access to a government-funded machine via peer review, but they might find that the wait is 12 months or more, and, as Minor pointed out, "their competitors would not necessarily wait a year." Or the researchers can pay to move up in line. Minor said that at least one location, the European Synchrotron Radiation Facility, offers faster access to researchers who pay extra. Finally, researchers can buy dedicated time on a synchrotron—if they have the money. In one case that Minor described, 6 days' access cost $250,000.

The simplest way to increase access, Minor noted, might seem to be to build more beam lines—the individual sources of radiation in a synchrotron or other device—but that would be expensive. Instead, he said, it might make more sense to increase the productivity of the machines. And, he said, "what limits the productivity of each beam line is software."

"On some beam lines, to do a simple experiment, you have to use four computers—you have to jump from one computer to the other and use four different programs." The problem is that the various computers at the synchrotron have never been integrated, and this slows down an experiment considerably. Handling the vast amount of data generated by the beam line is another difficulty. "You are producing 6.5 billion bits of raw data for 20 minutes. No network can sustain that load. In fact, the fastest network is something that I call 'sneakernet'—you are taking your hard disk from the computer and putting another one in." A third impediment arises when a researcher switches crystals, which must be done often. "Yes, you might collect all the necessary data in 2 minutes, but changing and aligning the crystal takes a half-hour. Why? Because it's done in a very, very conservative way."

All that can be greatly improved with the proper software, Minor said. Indeed, after he applied "a little unconventional thinking" at a beam line at Argonne National Laboratory, he said, that beam line produced in 9 months as many protein structures as nine beam lines at Brookhaven National Laboratory turned out in a year. But good software for protein crystallography is not widely available, and Minor identified several reasons for that.

"The basic problem," he said, "is that you do not have tools to develop the software. You have to build tools basically from scratch, and sometimes, even if the tools exist, there are such restrictions on them that you prefer to build them from scratch." And the reason there are no tools for writing software, he said, "is that there is basically zero recognition for people who develop tools." Without such recognition, it has proved difficult to interest people with computer-science backgrounds to work on software for this specialized field. "None of the crystallographic software has been developed by anybody who had any training in computer science. It has been developed by scientists in the field."

Software distribution is another problem. If just just a few laboratories use a program, the developer can answer questions without too much difficulty. But the better programs get used in hundreds of laboratories, and researchers are not equipped to answer questions from and work with hundreds of users of their software."

Underlying all those issues is the question of funding. "If it's a 1-year project, it can easily be the component of another project. If it's a 3-year project, it's a Ph.D. project, or you may put a postdoctoral scientist into the job. But if it's a 10-year project, it has to be funded and recognized separately." The most important programs demand tens or hundreds of person-years to develop, so funding is critical, and it is usually not easy to find. Government agencies do not always recognize the importance of software, and the software generally can be commercialized only when it is already successful.

THE PROTEIN DATA BANK

Once researchers determine the structure of a protein, they are required to deposit the structure with the PDB, which has recently been moved from the Brookhaven National Laboratory to Rutgers University. Input into and access to data in the PDB now take place over the Internet, which is convenient for researchers, but Helen Berman identified several unresolved issues affecting access to the protein structures and other information in the database.

First, and most sensitive, is the question of how long data should be held before they are released to the scientific community. Historically, a 1-year hold has been placed on the information to allow the researchers who generated it to analyze it and reap the benefits of their own work. Without such a hold, some scientists worried, unscrupulous colleagues might swoop in and publish their own analyses first. Now, however, many in the research community are calling for quicker release of the data, and the organization that runs the PDB is trying to decide whether a new policy is needed.

A second question is whether and how thoroughly data should be validated before being put into the PDB. "Some people say they should be untouched, and others say they should be heavily checked," Berman said. Her own opinion is that "there should be minimal validation, in consultation with the author, to remove from the data what I would call obvious and embarrassing errors."

A third issue is the uncertainty of the implications of intellectual property rights legislation enacted in Europe and under consideration in the United States that might affect ownership of the data and the database, which are currently "unprotected". It is not clear if the database needs protection, said Dr. Berman; however, differences in national laws potentially complicate the PDB's relationship with secondary distribution centers in Europe and Asia.

Finally, Berman agreed with Vladek Minor that software development is a major bottleneck. Researchers depend on computer programs both for coming up with structures that will be deposited in the PDB and for analyzing structures that they retrieve from it, and the system does not do enough to encourage the production of such software. “The people who are developing software in an academic environment are not getting the salaries they could get in business, and they're not getting the normal academic recognition. You certainly don't get a paper out of making various kinds of tools available. So we have a difficult time convincing people in the academic realm to produce the kinds of software that are required for structural biology.”

“How software is developed and how software developers are recognized have to change,” Berman said, “and there has to be a way for people that have new algorithms, new software, or new tools, to get funded, even if it's not sexy. For the greater good of the community, we have to find a better way of handling software development for structural biology.”

CULTURE COLLECTIONS

For researchers who study bacteria and other microorganisms, culture collections are the only way to preserve a record of the creatures they have studied. Many culture collections are run by individual laboratories and departments, but these seldom have the resources or expertise to keep hundreds or thousands of different strains alive decade after decade. Thus, several major culture collections gather microorganisms from researchers around the world, keep them alive in culture, catalog them, and make them available to other researchers. They are a vital resource for microbiologists, and their success will strongly influence the health of the field.

Unfortunately, a large percentage of the microorganisms used in research are not retained, said Cletus Kurtzman, of the Agricultural Research Service Culture Collection. Scientific journals generally demand that researchers make available any microorganisms described in their published articles, but scientists often simply keep the cultures and respond to requests from other researchers themselves rather than depositing the material in a major collection. If researchers keep them, however, other researchers can find it difficult to get access to them, Kurtzman said. The original researcher might be planning to commercialize a strain and not be eager to share it, or a culture can be lost or allowed to die out. The point, Kurtzman said, is that samples “will not likely be distributed to anyone easily unless we do something about deposits at the beginning.”

Why do researchers not deposit microorganisms in a major, publicly accessible collection? “I'm sorry to report,” Kurtzman said, “that many investigators are saying, ‘This is my strain, and I want to control how it's used, and if you'd like that strain, I'd be happy to be a collaborator with you on your

next publication.'" The major concern, echoed Raymond Cypess, President and CEO of the American Type Culture Collection, is competition. Individual researchers worry "that when a material goes out into the public, the large factory-type research organizations will be able to capitalize on it and outmaneuver them for publication and for grants."

The other challenge facing nonprofit culture collections like his own, Cypess said, is the changes occurring in their funding. Support is shifting from federal programs to users of the collections. That means that the cultures most likely to be collected are the ones that have some commercial value, and this is causing a decrease in the diversity of the holdings of culture collections. As a result, the scientific community often must rely on places other than the major culture collections for its research materials; therefore, Cypess said, "80% of the materials that are currently used in the science establishment are undocumented and unstandardized."

MUSEUMS AND BOTANIC GARDENS

One often-overlooked source of research materials is the world's museums, said Leonard Krishtalka, director of the Natural History Museum at the University of Kansas. "I like to say that the massive amount of data housed in museums is really a stealth dataset. Nobody knows about it, nobody uses it. It is unmined."

Over the last 3 centuries, Krishtalka noted, researchers have catalogued 1.8 million species of animals, plants, and microorganisms and an enormous fossil record of animals and plants, and descriptions of the species and the samples have been placed in museums around the world. "At the University of Kansas, we have 7 million specimens of everything from algae to moose. At the National Museum of Natural History, about 120 million specimens. Worldwide, there are 3 billion specimens of animals and plants." And those specimens are accompanied by data on such things as taxonomic classification, geographic location, climate, ecology, anatomy, genetic makeup, and evolution. It all represents an incredible resource for scientists studying biodiversity or almost any aspect of life on Earth.

As an example, Krishtalka described a project with the Mexican government. By querying many natural-history museums around Mexico, a group of researchers at the University of Kansas accumulated a list of where in Mexico deer mice had been collected over the last century and the climatic conditions at the times of collection. Deer mice are carriers of the hantavirus; by analyzing their occurrence and the climate data, the group was able to predict where in Mexico future outbreaks of the hantavirus disease were likely to occur.

At the moment, however, only about 5% of the specimens in museums worldwide have been collected in digital databases. Entering the rest into databases and keeping up with the constant flow of new specimens from

scientists who cede them after their research is done will be a challenge to museums, Krishtalka said.

"We are going to need enormous physical and information technology resources to handle the voucher specimens and the data." There will also have to be an "informatics infrastructure" that allows researchers to access the data. In addition, the museums will have to address a series of questions concerned with the data they are collecting: Who owns the data? How can sensitive data be protected? How can profits generated from the data best be channeled back to their owners?

Those who gather specimens for the collections face a different set of hurdles, said James Miller, of the Missouri Botanical Garden. A botanic garden, he noted, is very much like a museum but with just one department: plants. In marshaling specimens from around the world, Miller said, museums and botanic gardens must abide by the recently signed Convention on Biological Diversity. "The convention calls for the tropical countries of the world, which are roughly equal to the developing countries of the world and are home to the vast majority of the world's species, to promote access to and study of the biologic resources that are held within their international borders. But at the same time, it calls for those countries to regulate that access. And therein lies one of the problems that we face with access."

The developing countries, Miller said, want a series of issues to be addressed before they will allow their plant or animal life to be shipped elsewhere for study. First, what is the intended use of the biologic materials? The countries will treat materials intended for academic study much differently from those intended for commercial applications. They are also sensitive to the ethics of the acquisitions. "For materials collected in their countries, they want to see information about the materials, duplicate specimens, and so on remain in their countries so that their countries will benefit from the increase in scientific knowledge that results from the collection of those materials."

Finally, they wish to share in the profits of any commercial exploitation of their resources, but it can be difficult to negotiate a sharing agreement that is acceptable to everyone involved. It is hard, for instance, to know how to value the biomaterials provided by a country, and the parties must decide on when and in what form there should be payback—royalties, up-front cash payments, shared research opportunities and the education of some of the developing country's scientists, or perhaps something else. So far, no norms for those sorts of agreements have been established.

"One impediment to the establishment of norms," Miller explained, "is that most of the bioprospecting agreements are proprietary. The specifics aren't shared. So despite the fact that the last 10 years has seen perhaps 50 or 100 international bioprospecting agreements, there is no consensus about how equivalent they are to one another, because the specifics relating to royalty rates and the benefits to be shared are not public information."

ECOLOGY

The field of ecology faces a situation very similar to that of museums. A great mass of ecological data has been gathered, but the data are scattered and in disparate forms. If they could be collected and put into large databases for analysis, they would constitute an invaluable resource for ecologists. But the hurdles to that assembling are formidable.

"In the past, we ecologists have made up data sheets, copied them, put them on a clipboard, and gone out and written down information," said Jim Reichman, director of the National Center for Ecological Analysis and Synthesis at the University of California in Santa Barbara. "Now we need to move into a new era where we gather the information electronically." Particularly important are the metadata—information about the data. "One thing an ecologist will always say is, 'nobody knows my data like I do', and that's certainly true. But the idea of metadata is to ensure that somebody else can know your data almost as well as you do, and that might include something as simple as a photograph of a field site. It also includes documenting the data so that when you put something about the biomass of an organism or of an area, you know whether it's in pounds per acre or grams per square meter."

Gathering all the data and metadata into usable databases will demand not just the storage and computing power to handle all the information, but also some sophisticated informatics technology, Reichman said. "It would be much easier if we all put our data in the same way and had access to them in the same way, but that has not happened, and it's unlikely to happen in the future. So we probably need to develop solutions after the fact—data-crawlers, in effect—that will go in and find the kinds of data we want whatever their format, and extract them for appropriate use."

But, he continued, "as difficult as some of these technologic issues are, I think that in the long run sociologic issues are of the greatest concern." For one thing, ecologists do not want to have to bother with putting their data into a database-ready format. "I would say that it's equivalent to washing glassware in the laboratory; nobody likes to do that." More important, ecologists worry that if they put their data into a database, someone else might scoop them, or someone might misinterpret their data or even use their data to prove them wrong.

The issue of intellectual-property rights to ecologic data is particularly tricky, Reichman noted. "Often, ecologic data have no value right away. The value comes from the packaging of the data, from understanding broad patterns in time and space; so the informatics element is much more important than simply knowing the name of a species or knowing that a particular specimen occurred in a particular place."

Finally, the culture of ecology needs to change. "We tend to have a mystique about the ecologist who goes to a new place, sleeps on the ground for a half-year, collects a lot of data, and stumbles back into the laboratory with some new results." But if ecology is to benefit from the new databases, the field will

have to accept and reward a new type of ecologist: one who rides a computer instead of a jeep or a burro. And that might be the most difficult adjustment of all.

DEVELOPMENTAL PSYCHOLOGY

Like ecologists, developmental psychologists have had little interest in databases. "Researchers were expected to share summary data if they were approached by someone who wanted to do a meta-analysis; or if someone challenged their data, they were obliged to share them," explained Sarah Friedman of the National Institute of Child Health and Human Development (NICHD). "And they were expected to keep their data for about 15 years. But there was no requirement to archive the data or to make them user-friendly if someone were interested in accessing them."

And, like ecologists, development psychologists have traditionally given little respect to those who did not collect and analyze their own data. "People who submit research proposals to NIH to do secondary data analysis to answer questions in child development don't do very well in terms of funding," Friedman said, "because the reviewers on the review panels think that the data that were collected in order to answer other questions are not the most appropriate for answering the new questions."

Both those attitudes are changing, Friedman said, as psychologists have come to see the potential of the new technologies. But to take advantage of that potential, psychologists must first address a number of issues that they have generally not faced in the past. As an example, Friedman described the NICHD Study of Early Child Care, which followed 1,300 children and their families for several years beginning in 1991.

Data were collected at 10 participating sites, and different investigators put their data into a central data center that all would have access to. "Several years into the study," Friedman said, "the funding agency, NICHD, told the investigators that the data set would need to be placed in the public domain. The idea was that the data set was paid for with public funds and that giving other investigators access to the data would lead to an increased scientific return on the investment in the study, which was a large investment." The data had never been intended to be put into the public domain, Friedman said, and that has caused the investigators several problems. The consent forms signed by the study participants, for instance, did not mention putting the results in the public domain. Future consent forms will include this, of course, but, Friedman said, "placing the data in the public domain breaches the agreement between the investigators and the research participants."

The investigators must also come to grips with exactly what the "data" of the study are. Part of the study entailed making videotapes of the participants—videotapes that would clearly identify the participants. Yet the investigators promised the study participants that they would remain anonymous. Should the videotapes be considered data?

As in other fields, the investigators are concerned that putting the data in the public domain will allow other researchers to profit unfairly from their work, perhaps scooping them by reporting analyses of the data first.

Finally, the question arises of who will pay to put the data in a useful format for use by others. "Preparing the data sets for use by people who did not develop the data and who do not know them inside out is time-consuming and expensive," Friedman said, and if someone must be available to answer questions about the data, that only adds to the cost.

HUMAN-POPULATION DATABASES

There is great value in collecting data on disparate groups of people around a country or around the world. It allows researchers to look for patterns and to spot trends or tendencies that might not otherwise be obvious. It also allows them to test hypotheses on different populations. But collecting data on people, particularly genetic data or detailed information about health and habits, is fraught with difficulties that researchers dealing with, say, protein structures or plants, do not face.

Consider, for instance, the National Longitudinal Study of Adolescent Health. It follows ten of thousands of children beginning in middle school and high school for some 7 years, or until the subjects are 18-25 years old. Its purpose is to trace the "health-related behaviors of adolescents and the consequences of those behaviors in their young adulthood," explained Richard Udry, its director. The data are deposited into a data set and, as soon as they are ready to use, are released to researchers. But many of the data, which include DNA samples and detailed personal histories, are sensitive, so the study has had to find ways to guarantee the confidentiality of the subjects. "We probably spent well over a million dollars in the extra security precautions," Udry said.

Identifiers are stripped from the data, and the identities of the subjects and links to the data are held by a security partner, separate from the database that contains the medical data. The study has taken steps to keep the data from ever being subpoenaed. And, Udry said, the study has instituted a complex series of defenses against "deductive disclosure"—the possibility that someone, knowing that a particular person had taken part in the study, could pick out that person by using the information in the database, such as sex, age, urban or rural setting, or participation in sports. It is not simple to protect identities absolutely,

Udry said, but it is necessary. "Researchers should never collect data whose confidentiality they cannot protect." Udry added that protecting sensitive data would be potentially complicated by a 1998 amendment to the Freedom of Information Act that would make scientific data produced with public funds subject to public disclosure under the Act.

The collection and analysis of human DNA samples can shed light on important questions in human evolution and genetic variability, but progress in this area of research has been slowed by misunderstandings and concerns about the way this information will be used. And if data are collected on people with different cultural beliefs and practices, a whole new set of considerations arises, said Lynn Jorde, of the Department of Human Genetics at the University of Utah School of Medicine. "An important issue is whether study subjects understand the issues addressed in the informed-consent document," he said. "That is a challenge in any population but perhaps a special challenge in populations whose technology is different from our own." Furthermore, the whole idea of "consent" might be different in other societies. In the United States, "consent" is understood to mean "individual consent"; but in some cultures, the more important consent could be that of an entire group. Finally, he said, using the DNA of subjects in immortalized cell lines—a standard way of preserving samples—can be a problem because of "cultural reservations about the long-term preservation of a part of you that is still living and that might go on living even after you're dead."

Once the data have been collected, their dissemination and interpretation can also be tricky. "Because genetic data can be sensitive, particularly when we're looking at ethnic variation in single-nucleotide polymorphisms, it might be reasonable to somehow restrict access to investigators with scientifically legitimate questions," Jorde said. And, he added, when the data have been analyzed, researchers should be careful to ensure that results are accurately understood by the population that has been studied.

Jorde described a study that he did in India that found less genetic variation in maternally inherited DNA between women in adjacent castes than in castes that were far apart in the caste hierarchy. That was simply the result, he said, of 3,000 years of a caste system that sometimes allowed women to marry up in rank to an adjacent caste. But some Indian newspapers, in reporting the results, interpreted them as implying that "your genes determine your caste" or that "scientists could now look at an individual's genes and determine which caste he or she came from." That was not what the study said at all, and Jorde concluded, "that it is our responsibility to try to disseminate these results in as accurate a way as possible to avoid misinterpretation."

Epilogue:

Chairman's Remarks

The NRC conference, “Finding the Path: Access to Research Resources”, presents us with a mixture of new and not-so-new problems, all of which strike to the heart of the health and future of the scientific enterprise and the effective application of science to further the public good. While there are few completely new issues revealed here, the conference brought together concerns in a wide range of fields, and there is a new urgency as science and its commercial applications accelerate. The stakes are increasing every year.

The changing face of patent issues in the life sciences, most participants felt, has had a major effect in recent years on the conduct and the atmosphere of academic research. The 1980 Supreme Court decision in *Diamond v Chakrabarty* opened the door to patents on microorganisms and genes, and was key to the rise of the US biotechnology industry in the subsequent 20 years. But disagreement over what types of genetic material should be eligible for patenting is strong and continues today, as illustrated in the comments of Steve Holtzman and Craig Venter at the conference. When panelists were asked why so few institutions or companies had responded to the Patent and Trademark Office's request for comments on its proposed guidelines for patenting expressed sequence tags (ESTs), several suggested that their organizations had no confidence that a satisfying and workable policy would be crafted. Lacking any other venue for resolution in advance of issuance of a patent, these matters will continue to be settled case by case in the Court of Appeals for the Federal Circuit.

Another set of closely related issues concern material transfer agreements and licensing negotiations between academia and industry, which many feel have for a decade been an impediment to the efficient and equitable sharing of research materials. Despite numerous meetings that have described the issues and put forward model agreements to simplify sharing, uniformity

remains only a goal, in part because many parties, in both academia and industry, seek terms that are not mutually compatible and resist streamlining the process in seeking their own perceived advantage. While there is great heterogeneity in this area, fault can easily be found on both sides. Many universities stand accused of overvaluing their nascent technology, and industry is often blamed for overreaching; seeking far-reaching rights to future discoveries in exchange for research support and proprietary information, and for efforts to perpetuate publication delays. While some institutions, individuals and companies are culpable, full understanding of the processes of commercialization and the mechanisms for protection of inventors and investors is not common. Between their institutions and the commercial sector many academic scientists appear to be caught in the middle.

The issues raised with a special new urgency at the conference are connected to the rise of the database as a critical research resource, not only for geneticists, molecular, computational and structural biologists, but also for ecologists, anthropologists, zoologists, botanists, crystallographers and social scientists as well. It is a simple fact that computer accessible databases will soon become a dominant research tool in research of all stripes. In the case of genomic databases, the prospect of the private sector holding close a large segment of valuable information presents a daunting challenge to publicly funded science: either commit to duplicating the databases for public use at significant cost or resolve to pay the costs of access to private resources. The public responsibility of assuring that the data will be there for the science of the future is a intimidating one. There are arguments that support one or the other solution for a variety of data types, but the answer will likely be a mixed strategy.

Databases in other fields have different, but no less real, costs associated with organization and maintenance. Making databases into usable public resources requires investment in their infrastructure. Responsibilities accruing to the providers of such databases include: protecting the underlying sources of information (for example, human identities or museum and culture collection specimens), developing standard formats for data contributions and data presentation on-line, supporting the creation of interoperable, flexible and powerful computational tools to search, retrieve and analyze the data, and establishing a common set of values to balance the interests of data contributors and database users.

Perhaps it is now time to re-examine both the old and new issues from a different perspective. For example, we often speak of the traditional missions and needs of the different parties involved in resource sharing as though they were well-defined for all time, but much of the tension over materials transfer agreements and licensing and even over data sharing reflects a more complex, changing system of interests and values. Although our attention at the conference was focused on the points at which conflicts arise, these issues cannot be solved until we grasp more fully the factors that shape the values and

interests of the several parties. To do that, we need a better understanding of the changing nature of the scientific enterprise, including the proper extent and structure of industry-university research alliances and the implications of the parallel research programs in the commercial and academic sectors.

Consequently, we need to evaluate the effects of basic federal policies, such as patent policy and the policies emanating from the Bayh-Dole Act, on university research broadly, and on the behavior of universities, industry, and individual scientists in both sectors. We should assess whether they are having their intended effects, what their unintended effects are, and to what extent they are promoting the public benefit from science.

We must be concerned not only with encouraging important commercial applications in the short term, but also with the long-term health of the scientific endeavor. If we are asking scientists to collaborate to develop and use new research tools, we must ask whether federal policies, funding levels and spending priorities are commensurate with our expectations.

Life science is evolving rapidly, and judging from the history of other scientific disciplines, it is now clear that there is no going back to the "way it used to be". Given the opportunities for furthering the science and the public good that so many kinds of cooperation present, I would argue that we should not wish to return to the ways of the past. As we approach the new century it is evident that the many public declarations that the 21st century will be the century of the life sciences are likely to be true. We must remember, however, that it is deeply in the national interest for us to understand the cultural, social, legal, and financial forces that shape the infrastructure and context of the enterprise of all science. As we develop the policies and structures for the future we need to understand their ramifications in all sectors and work hard to enable the enormous opportunities of science for the public good to be fully realized.

David Galas
Conference Chairman

APPENDIX A

Program and Discussion Questions

National Academy of Sciences Auditorium

Wednesday, January 27, 1999

8:30 am Welcome

Dr. Bruce Alberts, President, National Academy of Sciences

Dr. David Galas, Conference Chairman

8:50 am Remarks

Dr. Mary Clutter, Chairperson, Subcommittee on Biotechnology of the National Science and Technology Council Science Committee

9:00 am Panel 1 Resources in Biotechnology and Genomics

Moderator: Dr. David Galas, Chief Academic Officer, Keck Graduate Institute for Applied Life Sciences and Chief Scientific Advisor, Chiroscience, Inc.

Panelists:

Tom Caskey Sr. Vice President for Research, Merck and Co., Inc

Craig Venter, President and Chief Scientific Officer, Celera Genomics Corporation

Rod Wing, Coker Chair of Plant Molecular Genetics and Director, Clemson University Genomics Institute

Maria Freire, Director, Office of Technology Transfer, NIH

10:15 – 10:30 am Break

Barbara Mazur, Director, Biotechnology, E.I. Dupont de Nemours & Co.

Steven Holtzman, Chief Business Officer, Millenium Pharmaceuticals, Inc.

Fred Anderson, Attorney, Cadwalader, Wickersham & Taft

Michael Snyder, Professor of Molecular, Cellular and Developmental Biology, Yale University

11: 45 am General Discussion

DISCUSSION QUESTIONS: PANEL 1 ISSUES IN BIOTECHNOLOGY AND GENOMICS

What lessons have emerged from the 7- year debate on the patentability of ESTs that might inform our evaluation of how patents on other genomic materials could impact academic research and commercial development?

What insight do the experiences of different genomic research communities in the development of a common resource offer to structuring similar efforts in the future? Are duplicative public and private efforts in constructing sequence databases the inevitable consequence of competition or opposing interests, or simply a result of the uncertain effects of patent rights on the use of material described in the databases?

Given advances in sequencing technology and the interest of the for-profit sector to produce sequence data, and databases, how might public policy most effectively support access to and productive use of this information by the academic and commercial communities? What are the risks to turning sequencing over to the private sector?

What approaches can be used to distribute the cost of developing and providing access to a research tool such as micro-array technology so that the maximum rate of scientific progress can be achieved? What criteria can be used to evaluate different approaches to reducing cost and enhancing access to research resources?

While the “research exemption” exists in concept but not law, could specific “research” uses of a patented or unpatented resource be identified and authorized in the interests of scientific progress? Would an expanded effort of the federal government to negotiate blanket licensing agreements like the NIH-Dupont “Cre-Lox” agreement serve the same purpose?

In addition to intellectual property rights, what are other important legal and policy frameworks that constrain access to research resources?

Wednesday, January 28, 1999

1:30 pm Panel 2 Issues at the Interface of University, Industry, and Government Policy

Moderator: Bruce Alberts, President, National Academy of Sciences

Panelists:

Joan S. Leonard, Vice President and General Counsel, Howard Hughes Medical Institute

Dennis Stone, Vice President for Technology Development, University of Texas Southwestern Medical Center

Harry J. Klee, Eminent Scholar, University of Florida Department of Horticultural Sciences

Tracy D. Wilkins, Professor and Director, Fralin Center for Biotechnology, Virginia Tech

2:45 – 3:00 pm Break

Moderator: David Galas

Panelists:

Candace Voelker, Associate Director, Research Administration and Technology Transfer, Office of the President, University of California

Tony E. Hugli, Research Scientist, The Scripps Research Institute

Tom Caskey, Sr. Vice President for Research, Merck and Co., Inc.

Chris Scott, Director, Research Development and Executive Director, ACCESS, Stanford Medical School

4:15 pm General Discussion.

5:30 pm Adjourn.

DISCUSSION QUESTIONS: PANEL 2 ISSUES AT THE INTERFACE OF UNIVERSITY, INDUSTRY, AND GOVERNMENT POLICY

What criteria/experience exists to evaluate appropriate returns (including IPR rights, rights to data) to industry funded projects of academic research at institutions receiving public research funds?

How do different federal programs affect access to research tools (e.g. CRADAs, SBIR, SBTT, Advanced Technology Program and the Bayh-Dole Act)?

How does corporate support of (and rights to) university research affect the university's ability to negotiate straightforward access to the research tools of other commercial bodies (which, for example might also seek rights to research results in order to protect its IPR)?

What minimum level of administrative management of these issues should be expected of all institutions receiving federal funds? What obligations do universities have to educate their researchers on IPR issues, material transfer agreements, and their responsibilities under the law?

What is the proper role of the Materials Transfer Agreement (MTA) and how can the process be made less burdensome? How realistic is it to monitor recipients for their use of transferred materials ? Universities are in the awkward position of being both user and producers of research resources, of providing a setting where the pure knowledge is pursued and shared freely while at the same time launching small companies. Which has priority?

If universities regard the products of their academic labs as potential income streams, should products that can be identified as research tools be considered in the same way as any other licensable product? What criteria might be developed to determine the “value” of research tools for the purpose of licensing agreements?

What responsibility does the for-profit sector have to the progress of science? Is it within the enlightened self-interest of firms (and all resource-holders, like universities for that matter) to recognize and respond to academia's needs for straightforward access to research resources, or is that a luxury only large firms and wealthy universities can afford?

Thursday, January 28, 1999

8:30 am Panel 3 Issues of Access to Research Resources Across the Disciplines: Exploring common and unique issues in crystallography, microbiology, biodiversity, anthropology, sociology, psychology, and computer science.

Moderator: David Galas

Panelists:

Ray Cypess, CEO and President, American Type Culture Collection

Helen Berman, Professor of Chemistry, Rutgers University

Cletus Kurtzman, Director, Microbial Properties Research, National Center for Agriculture Utilization Research, USDA

Leonard Krishtalka, Professor of Ecology and Evolutionary Biology, University of Kansas

Jim Reichman, Director, National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara

10:00 – 10:30 am Break

Vladek Minor, Department of Molecular Physiology and Biological Physics, University of Virginia

Lynn Jorde, Professor, Department of Human Genetics, University of Utah

Sarah Friedman, National Institute for Child Health and Development

J. Richard Udry, Professor, Department of Maternal and Child Health, University of North Carolina, Chapel Hill

James Miller, Associate Curator and Head, Applied Sciences Division, Missouri Botanical Garden

12:30 pm Conclusion and adjournment

DISCUSSION QUESTIONS: PANEL 3 ISSUES OF ACCESS TO RESEARCH RESOURCES ACROSS THE DISCIPLINES

As research resources with exciting new potential are developed, how can we ensure the advancement of knowledge while protecting the rights of the individual?

What types of experimental data or materials should scientists be asked to deposit in publicly accessible databases, repositories or collections that is not currently deposited?

Under what circumstances is it acceptable to request that access to parts of that data or material be withheld for a certain period of time? What criteria should determine who gets access to data, a resource, or samples and how it can be used?

How should the scientific community evaluate the individual research effort of a) the collection of data or samples that are used by the community, b) the analysis of community generated data or materials c) the development of research resources, their maintenance and improvement?

Are criteria available to determine the best way to provide access to a research resource? How should the costs (time and money) of providing access (supplying, delivering, maintaining, developing) research resources be assessed?

What intellectual property rights considerations might affect future use of or access to these resources?

Should there be a research exemption for non-commercial use of patented materials? Do patents make any difference to a bench scientist?

What are the implications of recent legislation extending the Freedom of Information Act to scientific data?

SUPPLEMENTARY QUESTIONS FOR PARTICIPANTS IN PANEL THREE

Crystallography

Are there types of experimental data that structural biologists should be asked to deposit in the Protein Data Base that is not currently deposited? Under what circumstances is it acceptable to request that access to structural coordinates be withheld? How would academic research or industrial strategy be affected if coordinates are granted patents?

Software and Hardware

What aspects of hardware (beam lines) and software limit progress in structural biology research?

What influences access to electron beam line facilities? How does software development as a research activity receive recognition? What obligations do academic software developers have to the community that uses it as a tool (and vice versa)? How should software be distributed to the research community and on what terms?

Microbiology

Should journals require that organisms cited in papers be deposited in a national collection in order to ensure their access to the research community? How should the costs of this requirement be assessed?

Sociology

How can individuals' privacy be safeguarded when sociological data is made public?

What determines who gets access to data and what data is given out? How should credit be given for the maintenance and improvement of access to large data sets?

Psychology

When and on what terms should raw data from large studies be released? Should those who collected the data be able to reserve its use for further research? What reward system or cultural changes will encourage collaboration, data sharing and the secondary analysis of data?

Ecology

What are the positives and negatives associated with the deposit of unanalyzed ecological data in publicly accessible databanks? What control should scientists have over their raw data? How will new databases be maintained and what cost, and how is credit for resource maintenance allocated?

Anthropology

What are the positives and negatives associated with the systematic collection of human genetic diversity data? What considerations, safeguards or restrictions on access to this data are needed for the successful development of this kind of resource? How should analytical software be shared within this community?

Biodiversity

What criteria are appropriate in determining permission of access to or sampling of material in museum collections? On what terms should access to and sampling of national (U.S. and foreign) biodiversity resources (both data and samples) be permitted? What intellectual property rights considerations might affect future use of or access to these resources?

APPENDIX B

Participant Biographies

CONFERENCE CHAIRMAN

David Galas is Chief Academic Officer of the Keck Graduate Institute of Applied Life Sciences. Dr. Galas recently served as president and chief scientific officer of Seattle-based Chiroscience R&D, Inc., a company that adopted an integrated, multi-disciplinary approach to drug discovery. Prior to his involvement in the business world, Dr. Galas served as director for Health and Environmental Research at the U.S. Department of Energy's Office of Energy Research, where he headed up the Human Genome Project from 1990 to 1993.

DISTINGUISHED SPEAKERS

Bruce Alberts is President of the National Academy of Sciences. Born in 1938 in Chicago, Illinois, Alberts graduated from Harvard College in Cambridge, Massachusetts, with a degree in biochemical sciences. He earned a doctorate from Harvard University in 1965. He joined the faculty of Princeton University in 1966 and after ten years was appointed professor and vice chair of the Department of Biochemistry and Biophysics at the University of California, San Francisco (UCSF). In 1980, he was awarded the honor of an American Cancer Society Lifetime Research Professorship. In 1985, he was named chair of the UCSF Department of Biochemistry and Biophysics. Dr. Alberts is a principal

author of *The Molecular Biology of the Cell,* considered the leading textbook of its kind and used widely in U.S. colleges and universities. His most recent text, *Essential Cell Biology* (1997), is intended to approach this subject matter for a wider audience.

Frederick Anderson is a Partner with Cadwalader, Wickersham & Taft and former Dean of the law school at American University. His practice involves science, the environment, and natural resources including health risk assessment and management, and issues regarding the use of genetic information and research. He is a member of the D.C. and U.S. Supreme Court bars. He is a member of the NRC's Commission on Life Sciences.

Helen Berman, a structural biologist, is a Professor II in the Department of Chemistry and a member of the Waksman Institute at Rutgers, the State University of New Jersey. Dr. Berman is the Director Designate of the Protein Data Bank and is head of the Research Collaboratory for Structural Bioinformatics. Dr. Berman founded and is currently the director of the Nucleic Acid Database Project, and has been a leader in the national and international mmCIF effort. She is the chair of the International Union of Crystallography Database Committee. Dr. Berman was President of the American Crystallographic Association (ACA) and has served on numerous advisory boards.

C. Thomas Caskey is Senior Vice President, Human Genetics & Vaccines Discovery at Merck Research Laboratories, West Point, Pennsylvania. He is Trustee and President, The Merck Genome Research Institute, Inc.; Adjunct Professor in the Department of Molecular and Human Genetics, Medicine, Biochemistry and Cell Biology at Baylor College of Medicine, Houston, Texas; and Adjunct Professor in the Department of Molecular Genetics and Microbiology, University of Medicine and Dentistry of New Jersey. His many honors include: Presidency of the Human Genome Organization; Member of the National Academy of Sciences; Doctor of Science Honorary Degree, The University of South Carolina; The Giovanni Lorenzini Foundation Prize for Basic Biomedical Research; Past President of the American Society of Human Genetics; Member of the Institute of Medicine of the National Academy of Sciences; Member of the Department of Energy Advisory Committee on Mapping the Human Genome. His research interests include: Disease gene discovery, DNA-based diagnosis and gene therapy.

Mary E. Clutter is Assistant Director of the National Science Foundation (NSF). Dr. Clutter came to NSF from the Department of Biology at Yale University to be Program Director of Development Biology. Dr. Clutter is the U.S. Chair of the U.S.–European Commission Task Force on Biotechnology, Chair of the Biotechnology Subcommittee of the Committee on Science of the

National Science and Technology Council (NSTC) and a member of the NSTC Committee on Science's Interagency Working Group on Plant Genomes.

Raymond H. Cypess is President and CEO of ATCC (American Type Culture Collection), in Manassas, Virginia. He holds a D.V.M. from the University of Illinois and a Ph.D. in parasitology from the University of North Carolina. Dr. Cypess came to ATCC from the University of Tennessee, Memphis, in *1993*. He was Dean of the College of Graduate Health Sciences, as well as Professor of Microbiology and Immunology and Comparative Medicine, and Vice Provost for Research and Research Training. He has been an Associate Professor of Epidemiology and Microbiology at the University of Pittsburgh School of Public Health, and Professor and Chairman at the New York State College of Veterinary Medicine. In the course of his extensive professional career, Dr. Cypess participated on NIAID scientific review boards and various NIH Study Sections and served on several editorial boards. He is the author of more than 75 chapters, reviews, and journal articles, a fellow in the Infectious Disease Society, a member of the American Epidemiology Society, and principal investigator on numerous grants, contracts, and academic-industrial initiatives.

Maria C. Freire is the Director of the Office of Technology Transfer (OTT) for the National Institutes of Health (NIH). Prior to her appointment at the NIH, Dr. Freire established and headed the Office of Technology Development at the University of Maryland at Baltimore and at the University of Maryland Baltimore County to provide for the effective transfer of technology from academia to industry.

Sarah L. Friedman is Special Assistant to the Director of the National Institute for Child Health and Development. She earned her M.A. in Educational Psychology from Cornell University in 1971 and her Ph.D. in Developmental and Experimental Psychology in 1975 from The George Washington University. While employed by the National Institute of Mental Health (NIMH), the National Institute of Education (NIE) and the National Institute of Child Health and Human Development (NICHD), she has published scientific papers and edited books addressing (a) the effects of preterm birth on cognitive, educational and social development of children; (b) the interface of brain, cognition and education; (c) the development of planning skills, and (d) environmental influences on psychological development. Since 1989 she has also served as the NICHD scientific coordinator and one of the investigators of a collaborative longitudinal research project on the development of social, emotional, cognitive, linguistic and health development of children from birth through first grade.

Steve H. Holtzman is Chief Business Officer of Millennium Pharmaceuticals, Inc. and has responsibility in the areas of business development, mergers and acquisitions, licensing, intellectual property, corporate law and corporate communications. Mr. Holtzman was formerly President of DNX Biotherapeutics, Inc. and was instrumental in developing DNX from a start-up to a publicly traded company. He currently serves as Co-chair of the Biotechnology Industry Organization's Bioethics Committee and was appointed by President Clinton in 1996 to the National Bioethics Advisory Committee. He is also a member of the NIH Working Group on Research Tools.

Tony E. Hugli is a Professor in the Department of Immunology, at the Scripps Research Institute and an Adjunct Member of the Institute for Bio-Medical Engineering at UCSD. Included among his many awards and honors is the Distinguished Alumnus Award from Otterbein College. Dr. Hugli was co-founding associate editor of *Protein Science* from 1991-1995 and is currently editor of *Immunopharmacology.*

Lynn Jorde is Professor and Associate Chairman of the Department of Human Genetics at the University of Utah School of Medicine. His laboratory is actively involved in studies of human genetic variation and has collected DNA samples from populations in India, Finland, and Africa. His laboratory also conducts studies of the genetic basis of human limb malformations. Dr. Jorde has served on several advisory panels for the National Science Foundation and the National Institutes of Health. He is the lead author of Medical Genetics, a textbook that is widely used in medical schools in the United States and abroad.

Harry Klee is Eminent Scholar and Professor at the University of Florida in the Horticultural Sciences Department. His research involves the manipulation of plant hormone synthesis and perception with the goal of improving crop plants. He has served on several panels and advisory committees concerning plant genetics and molecular biology, including the Joint FAO/IAEA Advisory Committee on application of agricultural biotechnology to the third world. He holds two patents related to plant biotechnology.

Leonard Krishtalka is Professor of Molecular Physiology and Biological Physics, at the University of Virginia. Krishtalka grew up in Montreal, Canada, where he attended McGill University and later received his Bachelor (1969) and Master (1971) of Science degrees from The University of Alberta, Edmonton. Krishtalka completed his doctoral studies in paleontology and evolutionary biology at the University of Kansas, Lawrence, and Texas Tech University, Lubbock (Ph.D., 1975). In 1989, he became Assistant Director for Science at the Carnegie Museum. In 1992, Krishtalka took a two-year leave from the Carnegie Museum to serve as a Program Director at the National Science Foundation, Washington, D.C., where he directed two research programs:

worldwide surveys of biodiversity; and biodiversity collections housed in national museums. Krishtalka left Carnegie Museum for The University of Kansas, Lawrence to become the Director of the Natural History Museum and a Professor of Systematics and Ecology. Krishtalka's research involves the origin and early evolution of the modern groups of mammals and he has been involved in the search for human origins with Richard Leakey's research team in Kenya and an international expedition to Ethiopia. Most recently, Krishtalka has worked with the museum community to use technology and information systems to bring the wealth of museum-based biodiversity data to science and society. In 1996 he helped form US-OBI, the U.S. Organization for Biodiversity Information, and is working with NAFTA's Commission on Environmental Cooperation, which is attempting to implement a North American Biodiversity Network.

Cletus P. Kurtzman is Research Leader of the Microbial Properties Research Unit at the National Center for Agricultural Utilization Research, Agricultural Research Service, of the U.S. Department of Agriculture in Peoria, IL. Dr. Kurtzman's responsibilities include directing a research program on molecular systematics of agriculturally and industrially important microorganisms, and enhancing the ARS Culture Collection (NRRL), an International Depositary Authority under the Budapest Treaty, which maintains 80,000 microbial strains (http://nrrl.ncaur.usda.gov). His personal research is on molecular systematics of yeasts.

Joan S. Leonard is Vice President and General Counsel for the Howard Hughes Medical Institute. She, along with her staff of six attorneys, is responsible for the whole range of legal matters affecting the Institute, including the implementation of its intellectual property policies as they affect its more than 320 investigators. Ms. Leonard served as a member of the Working Group on Research Tools, which was convened by the Advisory Committee to the Director of the NIH. The Working Group, which was charged with inquiring into the problems faced by NIH grantees in gaining access to research tools and identifying and assessing possible NIH responses, issued its report in June, 1998.

Barbara J. Mazur is currently the Director for Biotechnology Research in the DuPont Agricultural Products Enterprise. In this position she has responsibility for research directed towards developing novel grain quality products and for developing crop protection chemicals and products. The research includes programs with DuPont's alliance partner, Pioneer Hybrid International, as well as with Protein Technologies International and the Cereals Innovation Centre, members of the Agricultural Enterprise. Dr. Mazur has led the agricultural biotechnology program in DuPont for the past ten years. Prior to that she led a

research program that focussed on identifying genes conferring resistance to the sulfonlyurea class of herbicides, and creating herbicide resistance transgenic crop species with those genes.

James S. Miller is Head of the Applied Research Department at the Missouri Botanical Garden and coordinates a series of programs that explore possibilities for economic development of biological resources. In the past decade, more than 30,000 plant samples have been collected and screened for pharmaceutical or agricultural activity at the National Cancer Institute, several pharmaceutical or agricultural companies, and a number of university laboratories. In addition to being one of the most active programs promoting bioprospecting, the Garden has also been involved in developing the appropriate ethical and legal framework for operation of these activities. The Garden was one of the first institutions to sign an international agreement insuring that profits would be shared equitably with the source country if a products was developed and also one of the first to develop an institutional policy guiding research interactions with commercial programs.

Vladek Minor is an Associate Professor in the Department of Molecular Physiology and Biological Physics at the University of Virginia. Previously he was a Research Scientist in the Department of Biological Sciences at Purdue University. He received his Ph.D. in Solid State Physics at the University of Warsaw, Poland in 1978 and was a Visiting Professor at the Royal Institute of Technology in Stockholm, where he received the ASEA award for research work on the microcrystallization of metallic glasses. Since 1991 his research has focused on determining the crystallographic structure of macromolecules. He is a co-developer of the software program DENZO, used to process x-ray diffraction data.

O.J. Reichman is Director of the National Center for Ecological Analysis and Synthesis at the University of California, Santa Barbara. He currently serves as Editor of *Ecological Applications*; as President of the American Society of Mammalogists; on the Advisory Board for the journal *Ecosystems*; as Editor of Special Features for the *Journal of Mammalogy*; and on the Board of Trustees for BIOSIS Corporation. Dr. Reichman's current research involves the interactions between plants and animals in restored communities.

Christopher Scott is currently the Director of Research Development at Stanford University Medical Center and Executive Director of ACCESS, a clinical trials program at the School of Medicine. He is also the Associate Director of the Beckman Center for Molecular and Genetic Medicine and has held appointments as Director of Corporate Initiatives and in the Dean's office as Special Counsel in industrial relations and program development. He is the Administrative Director of the Program in Molecular and Genetic Medicine; a

multidisciplinary program designed to translate basic science discoveries to clinical application. He is the co-creator of Spectrum, a technology and information transfer program created to bring the biomedical industry in closer alignment with Stanford's academic research. His work interests and expertise center on new revenue strategies for clinical and basic biomedicine and the development of novel multidisciplinary research centers. His work career has spanned private industry and biotechnology.

Michael Snyder is Professor and Chair of the Department of Molecular, Cellular and Developmental Biology and Professor of Molecular Biophysics and Biochemistry at Yale, where he joined the faculty in 1986. His laboratory studies cell structure and division in yeast and was the first to carry out largescale functional genomics, characterizing the many genes of the yeast genome.

Dennis K. Stone is Vice President for Technology Development, Professor in the Departments of Internal Medicine, Biochemistry, and Physiology, and NCH Chair in Molecular Transport at UT Southwestern Medical Center at Dallas. Dr. Stone maintains an active laboratory that is focused on the structure and function of vacuolar type proton translocating ATPases. Before assuming his current position, Dr. Stone served as Associate Dean for Medical Student Research, Director for Internal Medicine Clerkship, and Director of the Graduate Program in Physiology. Dr. Stone is a former Searle Scholar and Established Investigator of the American Heart Association.

J. Richard Udry, Kenan Professor of Maternal and Child Health and Sociology, University of North Carolina at Chapel Hill, received his Ph.D. from the University of Southern California in 1960, and joined the faculty of the University of North Carolina at Chapel Hill in 1965. His special interest is integrating biological and social science models of behavior. His main current activity is directing the National Longitudinal Study of Adolescent Health.

J. Craig Venter is the President of Celera Genomics and Chairman of the Board of The Institute for Genomic Research (TIGR). At Celera, Dr. Venter is leading the effort to sequence the human genome by 2001. During his career focused in genomics and biomedical research, Dr. Venter has revolutionized the methods by which genomes are sequenced and analyzed. He has pioneered the use of automated gene sequencers and developed the expressed sequence tag (EST) technique for identifying expressed genes. With his lab at TIGR, Dr. Venter was the first to sequence the genome of an entire living organism.

Candace Voelker is Associate Director of The Office of Technology Transfer (OTT), University of California, Office of the President, which encompasses all the University of California campuses and includes Lawrence Livermore National Lab, Los Alamos National Laboratory, and Lawrence Berkeley

National Laboratory. She has spent 14 years in licensing intellectual property and five years as a patent advisor at Lawrence Berkeley National Laboratory. Her current responsibilities include managing the patenting and licensing staff at the University of California, OTT, that is responsible for the invention portfolios in the fields of chemistry, agriculture, and some biotechnology.

Tracy D. Wilkins is Director of the Fralin Biotechnology Center at Virginia Tech and holds the J.B. Stoobant's Professorship in Agricultural Biotechnology. His research interests are varied but primarily concern diseases of the intestine and, more recently, development of nasal and oral vaccines. In 1990, he co-founded a private company, TechLab, which develops and manufactures immunological tests to diagnose intestinal disease; and in 1992, he founded TransPharm, Inc. to produce transgenic farm animals that would contain human genes as part of their chromosomes and produce human proteins in their milk. He has been awarded eight United States patents on his work.

Rod A. Wing is Associate Professor and Coker Endowed Chair of Plant Molecular Genetics, in the Departments of Agronomy and Biological Sciences at Clemson University. He received his Ph.D. in 1987 at the University of California, Davis. Dr. Wing is also the Director of the Clemson University Genomics Institute (CUGI), which focuses on research, service and teaching in genomics. The research component of the CUGI is presently focused on 1) the development of frameworks to sequence the genomes of rice and rice blast; 2) the development of integrated physical maps between rice, sorghum and corn; and 3) the development and analysis of 50,000 cotton fiber and 50,000 barley ESTs. The service component of the CUGI is to provide affordable access to plant and fungal BAC libraries and related technologies through a five-year NSF funded BAC Resource Center. The Center presently produces, maintains and distributes the majority of BAC libraries used in agriculture today. The training component of the Institute includes both undergraduate, graduate, postdoctoral and visiting scientists training and education in genomics, primarily in the areas of BAC technologies, physical mapping, DNA sequencing and Bioinformatics.