

# COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 99-2					<b>FOR NSF USE ONLY</b>	
<b>BDI-99</b> <span style="float: right;"><b>07/12/99</b></span>					<b>NSF PROPOSAL NUMBER</b>	
FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.)						
<b>Database Activities</b>						
<b>DATE RECEIVED</b>	<b>NUMBER OF COPIES</b>	<b>DIVISION ASSIGNED</b>	<b>FUND CODE</b>	<b>DUNS#</b> (Data Universal Numbering System)	<b>FILE LOCATION</b>	
EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)		SHOW PREVIOUS AWARD NO. IF THIS IS <input type="checkbox"/> A RENEWAL <input type="checkbox"/> AN ACCOMPLISHMENT-BASED RENEWAL		IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? YES <input type="checkbox"/> NO <input checked="" type="checkbox"/> IF YES, LIST ACRONYMS(S)		
NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE <b>Indiana University Bloomington</b>			ADDRESS OF Awardee ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE <b>Indiana University Bloomington P.O. Box 1847 Bloomington, IN. 474021847</b>			
AWARDEE ORGANIZATION CODE (IF KNOWN) <b>0018093000</b>						
NAME OF PERFORMING ORGANIZATION, IF DIFFERENT FROM ABOVE			ADDRESS OF PERFORMING ORGANIZATION, IF DIFFERENT, INCLUDING 9 DIGIT ZIP CODE			
PERFORMING ORGANIZATION CODE (IF KNOWN)						
IS Awardee ORGANIZATION (Check All That Apply) (See GPG II.D.1 For Definitions) <input type="checkbox"/> FOR-PROFIT ORGANIZATION <input type="checkbox"/> SMALL BUSINESS <input type="checkbox"/> MINORITY BUSINESS <input type="checkbox"/> WOMAN-OWNED BUSINESS						
TITLE OF PROPOSED PROJECT <b>IUBio Archive and World-Wide High Performance Network Distribution of Bioinformatics Data</b>						
REQUESTED AMOUNT \$ <b>67,000</b>		PROPOSED DURATION (1-60 MONTHS) <b>12</b> months		REQUESTED STARTING DATE <b>01/01/00</b>		SHOW RELATED PREPROPOSAL NO., IF APPLICABLE
CHECK APPROPRIATE BOX(ES) IF THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW						
<input checked="" type="checkbox"/> BEGINNING INVESTIGATOR (GPG 1.A.3) <span style="float: right;"><input type="checkbox"/> VERTEBRATE ANIMALS (GPG II.D.12) IACUC App. Date _____</span> <input type="checkbox"/> DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.D.1) <span style="float: right;"><input type="checkbox"/> HUMAN SUBJECTS (GPG II.D.12)</span> <input type="checkbox"/> PROPRIETARY & PRIVILEGED INFORMATION (GPG II.D.10) <span style="float: right;">Exemption Subsection _____ or IRB App. Date _____</span> <input type="checkbox"/> NATIONAL ENVIRONMENTAL POLICY ACT (GPG II.D.10) <span style="float: right;"><input checked="" type="checkbox"/> INTERNATIONAL COOPERATIVE ACTIVITIES: COUNTRY/COUNTRIES</span> <input type="checkbox"/> HISTORIC PLACES (GPG II.D.10) <span style="float: right;"><u><b>Au</b></u> <u><b>Cn</b></u> <u><b>Jp</b></u></span> <input type="checkbox"/> SMALL GRANT FOR EXPLOR. RESEARCH (SGER) (GPG II.D.12) <span style="float: right;"><input type="checkbox"/> FACILITATION FOR SCIENTISTS/ENGINEERS WITH DISABILITIES (GPG V.G.)</span> <input type="checkbox"/> GROUP PROPOSAL (GPG II.D.12) <span style="float: right;"><input type="checkbox"/> RESEARCH OPPORTUNITY AWARD (GPG V.H)</span>						
PI/PD DEPARTMENT <b>Department of Biology</b>		PI/PD POSTAL ADDRESS <b>1001 E. 3rd Street</b>				
PI/PD FAX NUMBER <b>812-855-6705</b>		<b>Bloomington, IN 47405 United States</b>				
NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Electronic Mail Address		
<b>Donald G Gilbert</b>	<b>phd</b>	<b>1981</b>	<b>812-855-0587</b>	<b>gilbertd@bio.indiana.edu</b>		
CO-PI/PD						
CO-PI/PD						
CO-PI/PD						
CO-PI/PD						

## **Project Summary**

---

A central focus of bioinformatics is the large and rapidly growing knowledge of the genetic code of life. The Human Genome Project and related international efforts are on track with a full unraveling of this code, resulting in a flood of important data to bioscientists. The Internet is the primary means for distributing this data. As it is growing to tens of gigabytes, and rapid increase in use of the Internet has introduced delays, dissemination of this data in timely manner has suffered. A need exists for use of high performance network methods to speed its distribution.

To help address this need, this project supports IUBio Archive and Bio-Mirrors, centered at Indiana University with international partners, for high speed, world-wide distribution of genome data.

The principle investigator has a ten year record of providing public access to molecular biology data and software through IUBio Archive, which serves a worldwide community of bioscientists (<http://iubio.bio.indiana.edu/>). Indiana University is well positioned with its High Performance Network initiatives to address needs for rapid distribution of these data. Based on these strengths, a Bio-Mirror project has been initiated with international collaborators to provide rapid public access to and exchange of genome data. Bioinformatics centers in Japan, Australia, Singapore, China and Korea are currently involved with Bio-Mirrors.

# PROJECT DESCRIPTION

---

## I. Introduction

The end of the 20th century has seen an explosion of information discovered from living organisms, especially in areas of molecular biology and genetics. The focus of bioinformatics is dealing with this flood of information, coming from academic, industry and government labs, and turning it into useful knowledge. At the center of this information flood is the rapidly unraveling genetic code of life, and new research directions in understanding this code, how life develops from it, and how it can be engineered to advance health and well being of humans in many aspects.

The flow of information from the Human Genome Project and related worldwide efforts has revolutionized many fields of biosciences. This knowledge of the human genome code will revolutionize medical practice and biological research in the coming century, including understanding of most inherited diseases. The genome code coupled with new understanding of its organization, regulation and function in cells and in development of organisms will form the basis for designing new treatments for many diseases, and for understanding and modulating health problems with aging. Genome information is quickly becoming the basis for designing most new pharmacology and drugs. It is also central to the improvement of genomes of economically important crops and animals.

Detailed analysis and investigation of genome data has not been done, as the data are accumulating too rapidly to fully implement such analyses without the most powerful of computing methods and equipment. In fact this goal is the focus of several of the largest companies in the area of pharmaceutical, medical, agricultural bioengineering and biotechnology. Celera, a new company focusing on the extraction and sale of knowledge from genome data is building a bioinformatics center for this that is rivaled only by classified US government centers <ref Marshall 1999>.

Current Internet speeds and delays are a drawback to timely sharing of genome data. This genome data is published electronically from world bioinformatics centers, including the U.S. National Center for Biotechnology Information (NCBI), home to the GenBank <ref Benson *et al.*, 1999> DNA sequence databank. From primary collection centers, this data is distributed worldwide to researchers via the Internet. It is accessible to researchers in many ways, including Web searches for small portions of the data, Internet file transfers for bulk exchanges, and even slower CDROM distribution. Timely access to this daily changing data is of importance to

the continued rapid advance of biosciences research and discoveries, as it forms the basis for integrating the common genome codes used by all life, so new discoveries pertaining to one section of the gene code carry over to discoveries in many other sections of the code, understanding of biological pathways and gene functions, and of different organisms.

This project will build on Indiana University (IU) expertise and standing in the growing, important field of bioinformatics. Over the last few years, US biomedical research and funding has started to focus much more on biocomputing and bioinformatics as a nationally important field <ref Malakoff 1999>. IU Biology department has a ten year presence as a center for distribution of bioinformatics data and software, at IUBio Archive (<http://iubio.bio.indiana.edu/>). This will be a research area associated with the developing IU School of Informatics (<http://informatics.indiana.edu/>). It may well be a focus point to attract students and further national and industry support for this School.

## **II. Research objectives**

### **Objectives**

Specific objectives of this proposal are to provide computing infrastructure for continuing services of IUBio Archive of molecular biology software and data, including a new Bio-Mirrors project. A large (144 Gigabyte) redundant array of inexpensive disks (RAID) will make it possible to house and distribute important biosequence and related genomic data to the world community of bioscientists, using high speed Internet2 connections available at Indiana University. A multiprocessor Unix computer will permit indexing, searching and other analysis Internet-available public services based on these biosequence and genomic data.

### **Current support**

The IUBio Archive and Bio-Mirrors service rely now on equipment purchased for other purposes. These include:

- Sun Microsystems E3000 server and 24 Gigabyte RAID array, purchased for and used by the FlyBase genome informatics project <ref FlyBase Consortium 1999>.
- a 18 Gigabyte disk purchased by IU Biology department for this project. This has been filled to capacity, necessitating removal of data and delay in plans to add new bioinformatic data sets.
- a 100baseT connection through an ethernet switch to the IU campus backbone network.

As well, Indiana University's developing high performance network infrastructure and collaborative help have been essential to this project. The P.I. has and will devote needed time to maintaining this hardware, network connectivity, and biology data and services for the IUBio and Bio-Mirrors projects.

## **Broad impacts**

This project will enhance the infrastructure for bioinformatics in the US and world-wide, including a bioinformatic information distribution center utilizing emerging high-speed Internet2 network infrastructure. It will enhance partnerships with several Asian-Pacific nations in areas of bioinformatics and high-speed networking.

The results will be disseminated broadly to enhance scientific and technological understanding. At the core this is a public-oriented service, to improve access to a range of important biosciences information for average biologists, the general public, and for bioinformaticians working closely with large data sets.

Benefits to society include improved world scientific partnerships and technology transfer to areas that are now building bioinformatic and Internet centers for sciences.

## **III. Project background**

### **IUBio Archive history**

IUBio Archive for biology data and software has been an important Internet information resource for biologists since 1989 <ref Baxevanis and Ouellete 1998>. Its use has grown dramatically over its life, when it started as an archive of public molecular biology software, accessible by anonymous FTP (file transfer protocol). It provides access by HTTP (World Wide Web) and FTP file transfers, at the address iubio.bio.indiana.edu. It currently handles over 100,000 transactions (data queries, file transfers) per month. The PI pioneered in making DNA and protein data available for keyword searching through Internet Gopher and HTTP <ref Gilbert 1990, 1992>, and that is a continuing important resource for biologists seeking current biosequence information.

The archive is currently managed through time volunteered by the PI and those who contribute software and data. The popularity of IUBio archive has at times strained this volunteer effort. Started on a computer used as the IU Biology department data analysis computer, it has since migrated through three other computers, as an unfunded guest on equipment purchased for other uses. There have been rough times over the past years, with disk failures and running out of disk space. In 1993, several of its public users kindly contributed funds for a disk

drive. In 1996, the FlyBase genomic database server, with a large RAID for disk storage, became its current home, where the archive has occupied the lion's share of this disk. This RAID array has provided a very stable and reliable home.

## **Current contents of IUBio Archive**

Public software categories served from IUBio include molecular biology, general biology, chemistry, science, and utilities. Molecular biology software sections include alignment, codon, browsing, consensus, phylogenetics, pattern matching, primer selection, restriction enzymes, rna structure, searching, mswindows/ibmpc, macintosh, and unix software <ref Gilbert 1999>.

IUBio Archive has for years been the sole US public site <ref Gilbert 1995> for the widely used Sequence Retrieval System (SRS, <ref Etzold and Argos 1993>). SRS is an important tool for biosequence and related genomic data searches, as it provides a tested and working way of linking, or federating in a simple sense, many related databanks, as well as offering automated up to date access to these data. Unfortunately recent growth of databanks combined with other needs for the equipment IUBio resides on have forced cutbacks in available datasets available through IUBio's SRS service.

Services include search of GenBank nucleic sequence databank, SWISS-PROT, TrEMBL, PIR protein databanks, BLOCKS, ENZYME, PROSITE, REBASE associated biosequence data. Full release and daily updates of these data are searchable.

The SRS-FastA service <ref Gilbert 1995b> for sequence similarity searches against any user-defined subset of DNA and protein data is a unique service at IUBio. It includes subsets for the popular organisms *Arabidopsis*, *Caenorhabditis*, *Drosophila*, *Oryza*, *Poaecaeae*, *Rattus*, *Murinae*, and *Saccharomyces*.

An archive of the Bionet public news articles comprises a large and widely used section of this archive. Besides serving the biosciences community, this news archive is a source of science information well-used by the general public. Bionet news groups have been archived here since December 1990, and contain much useful information on molecular biology materials and methods, software, organism and techniques oriented science news and discussion.

The GenBank sequence search is among the most popular service, along with Swiss-Protein and other genomic databanks. The archive provides a widely used resource for biosequence information. IUBio services also include the largest U.S. collection of publicly available software for molecular and general biology, and a Bionet network news archive where researchers can find quick answers to many current questions in biosciences.

## Bio-Mirrors project

The Bio-Mirrors project is a worldwide bioinformatics public service for high-speed access to up to date DNA and protein biological sequence databanks. In genome research, these databanks have been growing tremendously, so much that distribution of them is hampered by existing Internet speeds. The Bio-Mirror project is devoted to facilitate timely access to important large data sets for this research.

High speed access between the sites is provided by the network infrastructure developed by Very High Speed Backbone Service (vBNS), TransPAC (Trans-Pacific network), and Asia-Pacific Advanced Network (APAN), and these sites are well connected to national research and education networks within each country.

The original Bio-Mirrors proposal was developed by Japanese partner Y. Ugawa <ref Ugawa 1997>, following discussions with the PI about ways to address the needs of timely international distribution of biosequence data. Subsequently A. Mizushima (Japan) and Tan Tin Wee (Singapore) joined this group, bringing the support and approval of the Asian Pacific Bioinformatics working group (APBionet).

The IUBio biology data and software archive has been an ongoing project of this IU department for this decade. Its international reputation and wide use as a public bioinformatic resource formed the starting point for the Bio Mirrors project.

Indiana University's developing high performance network infrastructure and collaborative help has enabled this project to start initial service in December 1998. Project documents, data sets and services are available at the Indiana University Biology department server, <http://www.bio-mirror.net/>, and at international partner sites (see below).

DNA and protein biosequence databanks are essential for advanced studies in genome research. These databanks are growing rapidly <ref Benson *et al.*, 1999>, doubling every 14 months since 1982, with the rate of increase accelerating as new genome sequencing projects are undertaken and new technology for this improves the output of high throughput genome sequencing centers for Human genome and other organism sequencing efforts. Bidirectional and multidirectional mirror services of this genome data, taking advantage of new Internet2 infrastructure and new methods, will help to distribute this data collected including in many countries and deliver them to researchers. This mirror service could not be established without high speed and reliable Internet2 connections.

## **Current Bio-Mirrors data sets**

Data in the Bio-Mirrors project currently totals about 10 Gigabytes in compressed format, and are updated from the primary sources nightly. DNA biosequence data include GenBank, EMBL, DDBJ. Protein biosequence data include SWISS-PROT, TrEMBL, PIR. Other data include BLOCKS, ENZYME, PROSITE, REBASE. These are mirrored from originating sites for these data banks, in the US, UK, Switzerland, and Japan.

Additional large data sets of importance to this project include the protein structure databank (PDB), and several whole genome and genome sequence integration data sets. These data sets will be added when disk space becomes available.

The Indiana University Bio-Mirrors offshoot of IUBio Archive, as well as several of the other Bio-Mirrors project sites, provides search and analysis services on uncompressed sets of these data, requiring some 20 Gigabytes of disk space for data, indices and updates, as well as CPU processing power for the analyses.

## **Institutions participating in the Bio-Mirrors project**

Agriculture, Forestry and Fisheries Research Council (AFFRC), Japan

File transfer: <ftp://bio-mirror.jp.apan.net/pub/biomirror/> Web:

<http://bio-mirror.jp.apan.net/>

Contact: Akira Mizushima [goddila@maffin.ad.jp](mailto:goddila@maffin.ad.jp)

Indiana University, Department of Biology, USA

File transfer: <ftp://iubio:iubio@bio-mirror.net/> Web: <http://www.bio-mirror.net/>

Contact: Don Gilbert [gilbertd@bio.indiana.edu](mailto:gilbertd@bio.indiana.edu)

Advanced Computational Systems CRC, Australia

File transfer: <ftp://bio-mirror.au.apan.net/biomirrors/> Web:

<http://bio-mirror.au.apan.net/>

Contact: Markus Buchhorn [markus@acsys.anu.edu.au](mailto:markus@acsys.anu.edu.au)

BioInformatics Centre (BIC), National University of Singapore, Singapore

File transfer: <ftp://bio-mirror.sg.apan.net/biomirrors/> Web:

<http://bio-mirror.sg.apan.net/>

Contact: Mark De Silva [mark@bic.nus.edu.sg](mailto:mark@bic.nus.edu.sg)

The Institute of Microbiology, Chinese Academy of Sciences (IMCAS), China

File transfer: <ftp://bio-mirror.cn.apan.net/> Web: <http://bio-mirror.cn.apan.net/>

Contact: Juncai MA [ma@sun.im.ac.cn](mailto:ma@sun.im.ac.cn)



The addition of a Bio-Mirror site in Korea is expected soon. Bioinformatics centers in other countries, including Canada and Europe, have expressed interest in the project. The support of Asia Pacific Bioinformatics Network APBioNet (<http://www.apbionet.org/>) through the APBioNet-APAN advanced networking project, with the Agriculture Working Group and Bioinformatics Working Group of APAN, has been instrumental in aiding this project.

## **Internet2 infrastructure organizations**

The following organizations for Internet2 infrastructure are employed by the Bio-Mirror project. Expertise and information from these organizations will be important to the most efficient and rapid exchange of Bio-Mirror data.

Abilene - <http://www.internet2.edu/abilene/>

Asia-Pacific Advanced Network (APAN) - <http://www.apan.net/>

Trans-Pacific network, TransPAC - <http://www.transpac.org/>

Very High Speed Backbone Service (vBNS) - <http://www.vbns.net/>

Singapore Internet Next Generation Advanced Research and Education Network (SINGAREN) - <http://www.singaren.net.sg/>

Agriculture, Forestry and Fisheries Research Council (AFFRC), Japan and Advanced Computational Systems CRC, Australia, are well connected to APAN, sharing data through APAN between the sites, and through TransPAC with Indiana University, and through SINGAREN with National University of Singapore. SINGAREN is connected to STARTAP vBNS, to CA\*Net2 and to Tokyo APAN.

Indiana University is well connected to high speed initiatives of Internet2, as a member in vBNS, and center of Abilene network operations and the US TransPAC connections to APAN. Abilene is an advanced research and education network in the United States. See <http://www.iuinfo.indiana.edu/ocm/releases/Abilene2.htm>

## **Future directions**

Potential additions for the Bio-Mirrors project may include new biosequence data such as the annotated rice genome sequence, estimated at 450 megabases, for which several international partners have planned to accomplish under the Rice Genome Project (RGP) <ref Burr 1997>. This project biosequence data would find a useful partner in the high speed data mirroring of the Bio-Mirrors project, as many of the countries participating in rice sequencing also have partners in the Bio-Mirrors project through APAN/APBionet centers for bioinformatics. Japan is the leading country for RGP, with Korea, China, Singapore, India, USA, Canada, and European Union involvement expected.

Publishing data from rice or other genome sequencing projects electronically in a

timely manner, in a way which all participating countries can have up to date access to the data is important. Bio-Mirrors can be an important partner in such work by offering high speed daily transfer data among Bio-Mirrors sites. The US effort for rice sequencing, in partnership with international groups, and is expected to start the fall of 1999.

A new bioinformatic project in development at IU Biology department is the Model Eukaryotic Organism Genome Information Database, which has a preliminary service at <http://iubio.bio.indiana.edu/meow/>. This project may also provide a highly important resource to the US and world biosciences community <ref Model Eukaryotic Organism Workshop, 1998>. It is designed to be portable and mirrored to other sites around the world, and would make a good addition to the Bio-Mirrors high speed distribution framework.

The Internet2 Distributed Storage Infrastructure (I2-DSI) project may offer important methods for the Bio-Mirrors project. Bio Mirrors members have agreed to continue for now with tested methods of FTP mirroring for this production service, but will investigate potential of DSI for improving network throughput and distribution of these data. As Indiana University is a partner in the I2-DSI project, it may be possible to tap knowledge of this for Bio-Mirrors.

## **Bio-Mirrors Project milestones**

- June 1997 -- initial discussions between Y Ugawa and D Gilbert about Internet 2 uses for biosequence data between USA and Japan.
- August 1997 -- "Development of Mirror Server by using High Speed Data Transfer in Genome Science" proposed by Y Ugawa to APAN organization. See <http://www.jp.apan.net/HPIIS-Applications/JP-AFFRC>
- February 1998 -- APAN Singapore link with APAN Japan established.
- April 1998 -- an expanded Bio-Mirrors project, by A Mizushima and Y Ugawa, approved by APBioNet.
- May 1998 -- APAN resource allocation secretariat approved the resource allocation which A Mizushima and Tan Tin Wee applied for through APBioNet.
- June 1998 -- Bio-Mirrors project approved by APAN.
- September 1998 -- Transpac link to APAN Japan established.
- December 1998 -- Initial mirroring between [bio-mirror.jp.apan.net](http://bio-mirror.jp.apan.net) and [bio-mirror.us.apan.net](http://bio-mirror.us.apan.net), including GenBank, EMBL, and protein data.
- January 1999 -- Australian site [bio-mirror.au.apan.net](http://bio-mirror.au.apan.net) joins; DDBJ databank added.
- March 1999 -- TransPAC, the international connection between the vBNS and APAN is operational at 70 Mbps.
- April 1999 -- Singapore site [bio-mirror.sg.apan.net](http://bio-mirror.sg.apan.net) joins.
- May 1999 -- China site [bio-mirror.cn.apan.net](http://bio-mirror.cn.apan.net) joins.

## REFERENCES CITED

---

- Baxeavanis, A. and B.F.F. Ouellete, eds, 1998. *Bioinformatics: A practical guide to the analysis of genes and proteins* J. Wiley and Sons, Inc., NY.
- Benson DA, et al, 1999. GenBank. *Nucleic Acids Res* 1999 Jan 1;27(1):12-7 See also <http://www.ncbi.nlm.nih.gov/>
- Burr B, 1997. An International Collaboration To Sequence The Rice Genome, <ftp://genome1.bio.bnl.gov/pub/maize/rice.html>
- Etzold, T and P. Argos, 1993. SRS – an indexing and retrieval tool for flat file data libraries. *Comp. Appl. Biosci.*, **9**: 49-58.
- FlyBase Consortium, 1999. The FlyBase database of the Drosophila genome projects and community literature. *Nuc. Acid Res.*, **27**:85-88. Also <http://flybase.bio.indiana.edu/>
- Gilbert, D.G., 1990. GenBank Search, a Hypercard stack for network searching and fetching sequences from the GenBank e-mail server. October 1990. <ftp://iubio.bio.indiana.edu/molbio/mac/gbsearch-ncbi.hqx>
- Gilbert, D.G., 1992. IUBio Archive through Internet Gopher. Feb. 1992. <<news://1992Feb13.023656.13616@bronze.ucs.indiana.edu>>. Also <gopher://iubio.bio.indiana.edu>
- Gilbert, D., 1995. New SRS node at IUBio, Indiana Bionet.Software.SRS, March 1995. <[news://3jdd3d\\$5t@usenet.ucs.indiana.edu](news://3jdd3d$5t@usenet.ucs.indiana.edu)> Also <http://iubio.bio.indiana.edu/srs/>
- Gilbert, D.G., 1995b. SRS-FastA: Similarity searches of Genbank subsets. Bionet.Software, Sept. 1995. <[news://434ddb\\$llc@usenet.ucs.indiana.edu](news://434ddb$llc@usenet.ucs.indiana.edu)> Also <http://iubio.bio.indiana.edu/srsfasta/>
- Gilbert, D., 1999. Free Software in Molecular Biology for Macintosh and MS Windows Computers. in *Bioinformatics Methods and Protocols* (S. Misener and S.A. Krawetz, eds.) Humana Press, NJ. Also <http://iubio.bio.indiana.edu/soft/molbio/Listings.html>
- Malakoff D., 1999. Biocomputing: NIH Urged to Fund Centers to Merge Computing and Biology. *Science*, **284**, p. 1742, 11 June 1999.
- Marshall E, 1999. A High-Stakes Gamble on Genome Sequencing. *Science*, **284**, p. 1906-1908, 18 June 1999.
- Model Eukaryotic Organism Workshop, 1998. Workshop report, December 1998. <http://www.nhlbi.nih.gov/nhlbi/sciinf/modeldb/model.htm>
- Ugawa Y, 1997. Development of Mirror Server by using High Speed Data Transfer in Genome Science. Proposal to Asian-Pacific Advanced Network (APAN) organization, <http://www.jp.apan.net/HPIIS-Applications/JP-AFFRC>