New developments at BIONET

David Roode, Rob Liebschutz, Sunil Maulik, Terry Friedemann, David Benton and David Kristofferson

BIONET, c/o IntelliGenetics, 700 E. El Camino Real, Mountain View, CA 94040, USA
Submitted August 17, 1987

Abstract

BIONET has made considerable progress in developing communication links among molecular biologists and biochemists worldwide. We describe these efforts and also note the many new enhancements to the BIONET system itself.

Introduction

BIONET is a non-profit resource for molecular biology computing funded by the National Institutes of Health Division of Research Resources and run by IntelliGenetics, Inc. through a five year cooperative agreement. BIONET provides molecular biology software, recent versions of databases relevant to molecular biology, and electronic communications facilities. An annual fee of \$400 is charged to academic and non-profit users to cover telecommunications costs. This fee allows unlimited use of the facility. Access limited to the communications facilities is available for \$100 per year. Non-profit users outside of the U.S. do not pay the BIONET access fee, but must pay their own telecommunications costs. For further information about BIONET contact Dr. David Kristofferson at the address above or send electronic mail to kristofferson@bionet-20.arpa.

Growth of the BIONET Computer Network

Work is in progress to expand BIONET as a logical communications network reaching molecular biologists worldwide. Many existing physical networks are used by molecular biologists. BIONET's aim is to utilize them all. Some users access BIONET by establishing a direct connection to the DARPA/NSF Research Internet (see below). BIONET assists sites needing information and advice on how to join the Internet. Remote terminal access via Telenet and Compuserve continues to be the means many use to reach BIONET. This is an important service for researchers lacking developed network connections. To include sites outside the Internet, BIONET is establishing a mechanism which allows a connection via the dialup telephone network. Computers which utilize this mechanism are known as BIONET satellites. Exchange of electronic mail and bulletin boards with BIONET satellites is provided by this method. Satellite Computers that run the Unix operating system make use of the existing Unix¹ Network News software and UUCP, a protocol used for telephone networking of Unix systems. BIONET is assembling software which provides similar functionality for VAX/VMS² computer systems. In August 1987, 16

¹UNIX is the registered trademark of AT&T Technologies, Inc.

²VAX and VMS are registered trademarks of Digital Equipment Corporation.

Internet sites and one phone link site were participating in the BIONET logical network with two to four new sites being added each month.

BIONET expects most of its network growth to occur through the TCP/IP-based Internet in the U.S., which is funded largely by DARPA and NSF. BIONET maintains a network connection to the Arpanet, the DARPA-sponsored backbone of the NSF/DARPA Internet. This network ties together various research and academic institutes throughout the United States, permitting them to exchange electronic mail, transfer files, and establish remote terminal sessions. Through the Arpanet, BIONET currently reaches the NSFnet backbone and all affiliated nets.

The Internet is linked to networks based on other protocols both in the U.S. and internationally. Mail relay is available on BIONET from Arpanet to BITNET, another widespread scientific network, which extends into Europe via EARN, the European Academic and Research Network, to Canada via NETNORTH, and elsewhere in the world. BITNET is likely to convert partially to interoperability with the NSF/DARPA Internet. Numerous research networks exist outside the U.S., including the Joint Academic Network (JANET) in England and ACSNET in Australia. Different access procedures are needed for each of these, e.g. as used in routing electronic mail. BIONET is working to clarify these procedures for its participants. Although we intend to promote common procedures for all international research networks, the services available today are impressive.

BIONET is working on plans to provide molecular biologists with access to one or more supercomputers or parallel processing resources. This will increase the computational power available to the community, allowing, for example, extremely rapid searches of nucleic acid databases. Special programs will be developed to provide an easy-to-use interface to submit the supercomputer jobs. We are investigating the benefits of directly connecting BIONET to one of the regional branches of NSFnet to facilitate resource sharing.

BIONET Bulletin Boards and Electronic Mail

In early 1987 BIONET began distributing copies of its bulletin board messages to SEQNET, a bulletin board service operated from Cambridge, England, and to other computer sites that could be accessed via the networks described above. Messages posted on any of BIONET's scientific or informational bulletin boards are distributed automatically to scientists around the world. Scientists interested in receiving copies of BIONET bulletins on their local computers should contact kristofferson@bionet-20arpa.

Especially active are the METHODS-AND-REAGENTS bulletin board (for requesting information on lab protocols and/or experimental reagents) and the RESEARCH-NEWS bulletin board. The latter is used for posting interesting scientific developments and as a place where scientists can introduce their labs and work to the electronic community. Bulletin boards exist for the GenBank and EMBL nucleic acid sequence databases. Copies of messages on these bulletin boards are forwarded to the database staff members for their attention. These bulletin boards allow discussion of issues relating to the databases and as a place where database users can obtain assistance. BIONET has also developed the GENPUB

program which facilitates submission of sequence data and author-entered annotations in computer-readable form directly to GenBank and EMBL via the electronic mail network.

Recent Enhancements to BIONET

BIONET has established a series of "HELP ME" menus to acquaint users with the capabilities of the resource. By typing HELP ME, the user is introduced to menus describing the software, E-mail and bulletin-boards, databases, and other information. Each menu leads in tree-fashion to help on a topic. Annotated examples of program usage are available which cover the major uses of the BIONET software for data entry, gel management, sequence, structure and restriction site analysis, cloning simulations, database searches, and sequence similarities and alignments. A manual of standard molecular biology lab protocols, donated by BIONET user Dr. Wajeeh Bajwa, has also been added to HELP ME.

One of BIONET's major goals is to serve as a focus for the development and sharing of new software tools. BIONET makes available to the community a wide variety of important computer programs donated by a number of software developers. This includes software developed by William Bains, Minoru Kanehisa, William Pearson, George Rose, Don Roufa, John Thompson, Michael Waterman, and Michael Zuker. The BIONET staff has collaborated with the software authors in making programs compatible with a number of hardware and user-community constraints. The software is available to users on the BIONET Resource's DEC 2065 computer and on a variety of other systems including VAX/VMS minicomputers, SUN workstations, and PC's.

BIONET provides an increasing number of databases online. These include the GenBank™ Genetic Sequence Data Bank, the European Molecular Biology Laboratory (EMBL) Nucleic Acid Sequence Database, the Protein Information Resource (PIR) Protein Sequence Databank, the SWISS-PROT Protein Data Bank, and the Brookhaven Protein Data Bank. A restriction enzyme database is provided to BIONET by Dr. Richard Roberts. The IntelliGenetics VectorBank™ of common cloning vector restriction maps and vector sequences is available as well as the KeyBank database which contains a growing set of regular expressions derived from published consensus sequences. A recent revision of the full text of *Genetic Variations of Drosophila melanogaster* by Dan L. Lindsley and E. H. Grell (the Drosophila "Red Book") is available online and can be searched by allele, author, journal, or other keyword. A revised version of Sprinzl's database of 871 tRNA (and tRNA gene) sequences and the Maize Genetic Map (including RFLPs) may soon be available. BIONET invites curators of genetic and physical genome maps to use this resource for the collection, maintenance, and distribution of their databases.

Acknowledgements

We thank the Biomedical Research Technology Program of the Division of Research Resources at the National Institutes of Health for their generous support (grant number RR01685-04).