Volume 14 Number 1 1986 Nucleic Acids Research

BIONET[tm]: national computer resource for molecular biology

Dennis H.Smith[1], Douglas Brutlag[2], Peter Friedland[3] and Laurence H.Kedes[4]

[1]IntelliGenetics, 124 University Avenue, Palo Alto, CA 94301, [2]Department of Biochemistry, [3]Department of Computer Science and [4]Stanford University School of Medicine, Stanford University, Stanford, CA 94305, USA

Received 22 July 1985

Abstract. This paper describes briefly the BIONET National Computer Resource for Molecular Biology. This presentation is intended as information for scientists in molecular biology and related disciplines who require access to computational methods for sequence analysis. We describe the goals, and the service and research opportunities offered to the community by BIONET, the relationship of BIONET to other national and regional resources, our recent efforts toward distribution of the resource to BIONET Satellites, and procedures for investigators to gain access to the Resource.

The BIONET Resource

BIONET is funded by a five year, cooperative agreement with the Biomedical Research Technology Program, Division of Research Resources, National Institutes of Health, for the period March 1, 1984 - February 27, 1989. The IntelliGenetics Division of IntelliCorp, Palo Alto, California, provides the computer facilities, core software and support. Responsibility for overseeing the Resource rests with a National Advisory Committee.

Goals

There are three major goals of BIONET:

- To provide computational assistance in data analysis and problem solving for participating scientists. Thus, software is available that can aid all phases of protein and nucleic acid sequence analysis, from initial data entry through design of cloning experiments and construction of probes.

- To serve as a focus for development and sharing of new software tools. Thus, participating scientists can contribute software for use by the community, and use Resource facilities to develop new software.

- To promote collaboration and rapid sharing of information among a national community of scientists. Thus, electronic mail and bulletin boards are available through which messages, experimental techniques, sequence data, and so forth, can be exchanged easily among geographically remote colleagues.

Charter of the Resource

The charter of a national resource funded through the Division of Research Resources specifies five major areas through which the resource interacts with the scientific community. Thus, BIONET provides:

- **Service.** Access to the computer facilities and staff is provided for scientists whose major activity is use of the programs for sequence analysis and electronic mail and bulletin boards.

- **Collaborative Research.** Access to additional system resources and staff is provided to those scientists who wish to develop new programs on BIONET and make them available to the rest of the community.

- **Core Research.** Longer-term research projects are undertaken by staff, based on the community's need for new computational approaches (hardware and software) to difficult problems.

- **Training.** BIONET has provided local and regional trainings in use of the computer system and its software. In addition, introductory materials, documentation and training manuals are available.

- **Dissemination.** The results of research carried out on BIONET can be disseminated to the community using electronic media (mail and bulletins) in addition to standard publications in scientific journals.

Staff and Computer Facilities

The staff of BIONET includes persons responsible for maintaining the computer facilities and system software, plus user consultants and scientists responsible for carrying out activities related to the five area summarized above. This staff, about six people, is very small compared to the size of the scientific community with access to BIONET. Therefore, we rely on the community for assistance in maintaining bulletin boards, contributing software, and providing assistance to their local colleagues in use of the system. We are continually working on more automated methods for providing high-quality service, and on developing new training materials to make it simpler for people to be able to use the Resource effectively.

The BIONET computer is a central time-shared Digital Equipment Corp. 2060 machine. Access is provided via local dial-up telephone lines and the nationwide UNINET telecommunication network. A suite of systems programs including editors, language compilers, mail, and so forth, are available to the community. Because of the interest in the community in use of personal computers to access the 2060, and to transfer files between machines, we maintain a lending library of public domain software that performs these functions, in addition to providing the KERMIT file transfer program for a variety of mainframe and microcomputers.

The Core Library of software for sequence analysis is provided by IntelliGenetics. We encourage the community to contribute additional software and will provide some assistance in making the software available to the community.

Current Status

BIONET is now in the second year of its five year grant. Access to BIONET was first provided in late November, 1984. Since that time, about 500 Principal Investigators have been granted access, representing a total of more than 1600 individual scientists. We expect this number to double in the next year. Access to BIONET is currently limited by the number of communication lines available, and we are working to increase this number in order to serve more simultaneous users. It is a characteristic of computing in molecular biology that use of programs is interspersed with substantial laboratory work. Even so, BIONET is rapidly approaching its capacity to serve the community. We are, therefore, beginning an effort to distribute the Resource among other computers at institutions across the country. This effort is described in a following section on BIONET Satellites.

In our first year of operation, we concentrated on building the Service component of the resource, and in holding several training sessions on use of BIONET. In the second year we are expanding our efforts to develop the Core and Collaborative components of the Resource. In addition, we are working on better training materials based on our experience with the current community.

Relationship to Other Resources

There are several other national and regional resources, funded by the US and other governments, with a primary focus on molecular biology. We have established joint computer accounts with these other resources, and are working toward better mechanisms of communication and data exchange with them. We summarize here our existing interactions with these resources.

GenBank[tm] and the European Molecular Biology Laboratory have as their primary focus the collection, organization and distribution of databases of nucleotide sequences. BIONET makes these databases available for use with its software for sequence analysis. We are developing mechanisms to allow BIONET scientists to communicate with and contribute sequence data to, these other organizations.

The Protein Identification Resource has responsibilities for a database of protein sequences, and development of and access to programs focused more on analysis of such sequences. BIONET makes this database available as well.

A regional resource has recently been established at the Dana-Farber Cancer Institute. Its primary focus is on developing new methods for sequence analysis, together with distribution of these methods to a wide community of scientists. As these methods mature, BIONET will participate in dissemination of them to its scientific community.

BIONET Satellites

The enormous demand for access to programs for sequence analysis, as evidenced by the growth rate of the BIONET community, clearly cannot be met by the relatively limited computer facilities of BIONET and related resources. Therefore, we have begun an effort to distribute BIONET among a set of computer facilities that can run BIONET software locally, but still communicate with the central time shared computer to maintain contacts with Resource staff and the rest of the BIONET community. Each remote facility will be given the status of a BIONET Satellite, and we will provide the necessary software to simplify exchange of electronic mail and data files with BIONET's 2060 as well as other Satellite machines. This will result in a tremendous amplification of the Resource itself while maintaining the contacts, through exchange of mail, bulletins, data and methods, that are critical to a *community* of participating scientists.

Procedures for Access to BIONET

Persons who wish to apply for access to BIONET, including the BIONET Satellite program, should write to The BIONET Administrator, IntelliGenetics, 1975 El Camino Real West, Mountain View, CA 94040. Access to BIONET is restricted to academic and other not-for-profit organizations. A subscription fee of $400, payable after an investigator has become familiar with the system, is required. The imposition of a subscription fee was required recently for two reasons: 1) the enormous response from the community has led to higher than anticipated telecommunications costs; and 2) NIH budget cuts for our current grant year.

Acknowledgement

We wish to thank the Biomedical Reseach Technology Program, Division of Research Resources, National Institutes of Health, for their generous financial support, Grant number RR01685-02.