23. User Services: Helping You Find Your Way

David Wheeler and Barbara Rapp

Created: October 9, 2002
Updated: August 13, 2003

Summary

The User Services team is the primary liaison between the public and the resources and data at NCBI. User Services disseminates information through outreach training programs and exhibits at scientific conferences and responds to incoming questions by email and telephone assistance. The team instructs people in the use of NCBI resources, responds to a wide range of questions, receives comments and suggestions, and coordinates with the NCBI resource developers to implement suggestions from users. In addition, User Services develops documentation, tutorials, and other support materials; produces the NCBI News; and publishes articles on NCBI resources.

The User Services Team

User Services consists of a staff of scientists and information specialists with diverse backgrounds and experiences. Scientist members of the staff hold Masters or Ph.D. degrees in an area of molecular biology, biochemistry, or biotechnology. Information specialists have Masters degrees in Library and Information Science and extensive experience using online databases of scientific information.

The Help Desk

Help Desk assistance is available from 8:30 a.m. to 5:30 p.m. Eastern Time, Monday through Friday. Two email addresses are available, info@ncbi.nlm.nih.gov and blast-help@ncbi.nlm.nih.gov. The info@ncbi.nlm.nih.gov address is for any type of inquiry, including questions about services; how to get started using the NCBI tools for a particular research problem; reports of technical problems; press inquiries; and comments or suggestions about NCBI resources. The blast-help@ncbi.nlm.nih.gov address is for questions and comments regarding sequence similarity searches using BLAST tools and databases. The Help Desk phone number is 301-496-2475.

email questions are answered as expeditiously as possible, usually within a day of receipt of the question. However, those that require extended investigation may take longer. Questions are usually handled directly by members of the User Services staff, although some are referred to a specific database development team for attention.

Examples of question topics include: data submission protocols, including the use of BankIt and Sequin (Chapter 12); finding summary information about a gene or disease; using Entrez (Chapter 15) to find sequence records for genes and proteins; choosing the best BLAST service to use for a particular application (Chapter 16); how to interpret BLAST results; how to display and manipulate three-dimensional structures with Cn3D (Chapter 3); how to display genome data

using the Map Viewer (Chapter 20); how to print or save search output; how to set up sequence databases and install NCBI software locally; and which databases to use for a specific research question. We also accept reports of possible data errors, suggestions of new features or content to include in databases, and reports of system bugs or Web problems. The NCBI also receives a number of press inquiries about the research applications of our services, bioinformatics in general, and various genome projects. Occasionally, a high-school student will submit questions for a classroom assignment, providing a special outreach opportunity to young scientists.

Because of the genetic focus of many of NCBI resources, we receive a number of questions from the general public regarding medical issues. The NCBI Help Desk staff can neither provide direct answers to medical questions nor give medical advice or guidance. However, we do provide suggestions on how to search our resources for information on the gene or condition of interest and refer users to the National Library of Medicine (NLM) customer service group for further assistance with PubMed (Chapter 2), MEDLINEplus [http://www.nlm.nih.gov/medlineplus/], and ClinicalTrials.gov [http://www.clinicaltrials.gov/]. We also refer them to outside organizations that can provide information on such topics as support groups and sources of medical advice.

Questions about PubMed are handled by a separate customer service group within the NLM. Their direct address is custserv@nlm.nih.gov, and their phone number is 1-888-FIND-NLM. PubMed questions that are received at the NCBI Help Desk are forwarded to NLM.

Development of User Support Materials

Because of its ongoing personal contact with our users, the User Services group plays an important role in communicating with database development and production teams, making suggestions, testing new releases and new features, and keeping them informed of problems that people are having with the services. The team also collaborates with developers in creating help documents, frequently asked questions (FAQs), tutorials, and workshop materials.

Tutorials on the Web

Web-based tutorials for BLAST [http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/information3.html], Entrez [http://www.ncbi.nlm.nih.gov/Database/tut1.html], Cn3D [http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3dtut.shtml], and PubMed [http://www.nlm.nih.gov/bsd/pubmed_tutorial/m1001.html] are currently available, with additional topics under development. Tutorials are produced on a collaborative basis by database development and User Services staff.

About NCBI

In keeping with the "plain language initiative" at NIH, the *About NCBI* [http://www.ncbi.nlm.nih.gov/About/index.html] section of the NCBI Web site presents many fundamentals of NCBI's bioinformatics tools and databases, including a science primer covering such topics as molecular genetics, genome mapping, Single Nucleotide Polymorphisms (SNPs) (Chapter 5), and microar-

ray technology (Chapter 6). A model organism guide presents various model organisms and their uses in laboratory settings. As an introduction and orientation to NCBI's multifaceted Web site, the *About NCBI* section appeals to the general public, educators, and researchers alike.

NCBI Site Map

The NCBI Site Map [http://www.ncbi.nlm.nih.gov/Sitemap/index.html] serves as a guide to NCBI resources. It provides a comprehensive, linked list of resources, along with a brief description of each resource. An effective way to locate a resource of interest within the Site Map is to perform a Find in Page search, a function that is built into all commonly used Web browsers.

Publications

The *NCBI News* is a quarterly newsletter that includes articles on new services, new features, and basic research at NCBI, as well as how to use selected resources for common applications. The newsletter is available free of charge and is offered online and by print subscription.

The User Services group also prepares fact sheets, brochures, and other public information materials to describe and illustrate NCBI services. A list [http://www.ncbi.nlm.nih.gov/About/news/print.html] of available materials is provided in the *About NCBI* section of the Web site, under News.

Overview articles entitled *GenBank* and *Database Resources of the NCBI* have also been published recently in the annual database issues of *Nucleic Acids Research* (1-3).

Outreach

NCBI's continuing emphasis on outreach to the scientific community is evident in its multifaceted program that includes exhibiting its services at scientific meetings, offering a variety of training courses, and developing Web-based tutorials and workshops.

Exhibits at Scientific Meetings

NCBI exhibits at approximately 15 scientific meetings per year, providing an opportunity for a wide range of researchers, students, and teachers to see demonstrations of NCBI resources and interact directly with NCBI staff. The current exhibit schedule [http://www.ncbi.nlm.nih.gov/About/glance/index.html] is posted on the NCBI Web site in the *About NCBI* section, under *NCBI at a Glance*.

Workshops are offered at select scientific meetings and include the standing workshops described below in the Training section, but workshops also can be customized for particular audiences. Meeting organizers who would like to invite NCBI to offer a workshop are encouraged to do so.

Training Courses

NCBI has a growing training program consisting of full-day, half-day, and two-hour courses that are usually a combination of lecture and computer-based formats. There are also advanced courses that are given over a more extended time period. Each is described briefly below, and further information on the training programs can be found in the Education section of the Web site, under *NCBI Courses* [http://www.ncbi.nlm.nih.gov/About/outreach/courses.html].

A Field Guide to GenBank and Other NCBI Resources

A Field Guide to GenBank and NCBI Molecular Biology Resources is a training course offered in a lecture format, followed by hands-on computer sessions. It is designed as a basic but broad introduction to NCBI tools and resources.

Field Guide topics include the following: description and scope of the primary database, GenBank (Chapter 1); derivative databases, such as UniGene (Chapter 21), LocusLink (Chapter 19), and Reference Sequence (RefSeq) (Chapter 18); effective database searching using Entrez; NCBI structure databases and the structure viewer, Cn3D; sequence similarity searching using the BLAST programs; the Conserved Domain Database (CDD) and associated search engine; and genome resources, including the NCBI assembly of the draft human genome, access to both finished and unfinished microbial genomes, and the genome Map Viewer.

The course is offered by invitation at academic institutions as well as at selected scientific conferences. If you are interested in hosting a course at your institution or conference, write to info@ncbi.nlm.nih.gov, and your request will be routed to the course coordinator.

The course is also offered four times a year at the NLM on the NIH campus in Bethesda, Maryland, and is free and open to anyone who would like to attend.

More information on this course [http://www.ncbi.nlm.nih.gov/Class/FieldGuide/] is available in the Education section of the NCBI Web site, under *NCBI Courses*. The Web site includes the course handout, slide presentations, and problem sets with answers. A schedule of planned courses at NLM and elsewhere is posted under *Upcoming Courses*.

Molecular Biology Information Resources

This course is designed primarily for medical and science librarians or other professionals who are providing support services for molecular biology information resources. It provides an introduction to four categories of molecular biology information available from NCBI: nucleotide sequences, protein sequences, three-dimensional structures, and complete genomes and maps. An overview of search systems available at the NCBI, particularly Entrez and BLAST, emphasizes how search skills related to other types of information resources also apply to molecular biology databases. The course concludes with a discussion of various levels of molecular biology information services provided by librarians.

The Medical Library Association approves this course for eight continuing education credit hours. The course has been given at 24 locations since May 1997. Because of the increase in NCBI services, courses are being revised and are not being scheduled at this time.

More information on this course [http://www.ncbi.nlm.nih.gov/About/outreach/courses.html] is available in the Education section of the NCBI Web site, under *NCBI Courses.* The Web site includes the course materials used for the lecture and a set of exercises.

NCBI Advanced Workshop for Bioinformatics Information Specialists

A new 5-day advanced course on NCBI resources has been developed as part of a collaborative project with a group of scientists and librarians who currently provide bioinformatics support services at their universities. The course provides detailed descriptive information as well as hands-on experience with handling a wide range of user questions. The course is designed for bioinformatics support staff based in university medical libraries so that they can, in turn, assist students, faculty, staff, and clinicians at their institutions in the use of molecular biology information resources. Additional information on this course [http://www.ncbi.nlm.nih.gov/About/outreach/index.html] is available in the Education section of the NCBI Web site.

Specialized Mini-Courses

The Service Desk staff also offers four mini-courses: BLAST QuickStart, Unmasking Genes in the Human Genome, Making Sense of DNA and Protein Sequences, and GenBank and PubMed Searching. Each is described briefly below. The purpose of the mini-courses is to focus on specific research application areas and address how to use multiple NCBI resources together to answer a research question. Additional problem-oriented mini-courses are under development.

The courses are 2 hours each in length. An overview is given during the first hour in lecture format, followed by a 1-hour hands-on session. Although primarily given on the NIH campus, NCBI is beginning to offer these workshops at outside institutions as well. Although the mini-courses were originally designed to be presented by an instructor, they are constructed in an online notebook format; therefore, it is possible to take the course on your own. Revisions to augment the online notebooks with lecture material and make the courses completely self-guided are currently under way.

BLAST QuickStart!

This mini-course is a practical introduction to the BLAST family of sequence-similarity search programs. Exercises range from simple searches to creative uses of the BLAST programs.

Unmasking Genes in the Human Genome

This mini-course covers how to find genes, promoters, and transcription factor-binding sites in human DNA sequences. It is designed around a program developed within User Services called Greengene, which integrates the output of several gene-finding tools and allows a coding sequence and accompanying protein translation to be assembled from the exons detected by these programs. Because the output of several programs is integrated, there is increased reliability in exon selection.

Making Sense of DNA and Protein Sequences

In this course, participants find a gene within a eukaryotic DNA sequence. They then predict the function of the derived protein by seeking sequence similarities to proteins with documented function using BLAST and other tools. Finally, a 3D modeling template is located for the protein sequence using the Conserved Domain Search (CDD-Search).

During the first hour, an instructor walks the class through an analysis of an uncharacterized *Drosophila melanogaster* genomic sequence from a GenBank record. During the second hour, participants perform the same analysis independently, using a different genomic sequence.

GenBank and PubMed Searching

This mini-course provides an overview of literature searching and sequence retrieval using the PubMed and Entrez database search interfaces. Exercises illustrate advanced search tips for using Entrez, many of which explore the use of the **Preview/Index** options for specifying parameters to limit the search results. The course also features 21 self-scoring exercises for GenBank.

The NCBI Learning Center

In addition to communication by email and phone provided through the Help Desk, a regular research consultation service provides one-on-one support for researchers in the NIH community. The consults are available by appointment and are provided in 1-hour time slots at the NIH Library as well as the NCBI training facility. Because of the success of the program, this type of service may be offered by appointment at selected scientific meetings in the future.

CoreBio

An innovative training program that began in 2001 aims to train molecular biologists for a new type of career as bioinformatics specialists who provide institutional support for users of computational biology tools. The NCBI Core Bioinformatics Facility (referred to as the CoreBio program) currently functions to train and support a network of bioinformatics specialists serving individual Institutes at NIH. NCBI's CoreBio facility trains Core members identified by their respective institutes in the use of its bioinformatics tools. The Core members, in turn, support the use of NCBI tools and databases by researchers at their institutes.

The training is provided over a 9-week period, with students attending lectures and completing practical exercises in the morning and returning to their regular workplace in the afternoon. The coursework centers on one major topic each week and follows the rough schedule given below:

WEEK 1: Introduction to the Sequence Databases
WEEK 2: BLAST
WEEK 3: The Human Genome
WEEK 4: Genomic Biology
WEEK 5: Molecular Modeling
WEEK 6: Web Page Development
WEEK 7: Setting Up a BLAST Web Server
WEEK 8: Interaction with Users

WEEK 9: Practicum

During week 9, the students pursue an institute-related project with the assistance of NCBI instructors. These projects run the gamut from the compilation of specialized datasets and data mining to the creation of novel BLAST interfaces and the construction of new data display tools. Students also develop a Web page to support the services they are developing for the respective Institutes at the NIH.

Although currently a NIH-based program, other organizations are welcome to consider using the program as a model for development of similar initiatives to meet their bioinformatics support needs.

Conclusion

At NCBI, we encourage our users to contact us with questions, suggestions, and requests for training or presentations on NCBI services. We invite feedback on tutorials, FAQs, and other support materials and welcome suggestions regarding additional materials that would be useful in guiding users through the wide range of services offered by NCBI.

References

1. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL. GenBank. Nucleic Acids Res 30:17–20; 2002. (PubMed)

2. Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, Pontius JU, Schuler GD, Schriml LM, Tatusova TA, Wagner L, et al. Database resources of the National Center for Biotechnology Information: 2002 update. Nucleic Acids Res 30:13–16; 2002. (PubMed)

3. Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, Pontius JU, Schuler GD, Schriml LM, Tatusova TA, Wagner L, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 29:11–16; 2001. (PubMed)