TECHSIGHTING
SOFTWARE

e-Grants?

Submitting a grant proposal usually causes anxiety for both principal investigators and their support staff. GrantSlam version 3.1 from Cayuse Software takes some of the pain out of NIH grant submissions [Public Health Service (PHS) 398 and 1250 forms], especially for those who process the grant applications for academic departments as opposed to those who write the main body of the proposal.

GrantSlam
Cayuse, Inc.
Portland, OR.
PC, $279 for PHS 398 or $379 for the Duo (PHS 398 and 2590); Mac, $399 or $499.
503-297-2108
www.cayuse.com

GrantSlam acts as a front-end negotiator to a database of user-entered grant information. Applicants accustomed to word processors will only need to learn how to enter data in the program. The database format simplifies the tracking of different grants as well as the revising of grants for resubmission by reassembling data from earlier applications. Most fields have context-sensitive pop-up windows with program help and information about NIH guidelines, features that make the database entry a bit more manageable.

Most of the information required on NIH forms is formatted automatically by the program as the user enters it. Biographical sketches, for example, are a standard component of many grant applications, so GrantSlam stores this information as a separate file that is available for each new submission. For the main body of the grant proposal, GrantSlam provides document templates for Microsoft Word and Corel WordPerfect with the correct margins, headers, and footers. Another handy template creates shipping labels with the correct addresses.

GrantSlam manages the database workflow fairly effectively, but it lacks some of the standard features of a word processor. For example, the program's lack of a spell checking function and its inability to copy or paste formatted text makes some operations cumbersome. Text formatting and entry of special characters require keystroke combinations that are unique to the program. GrantSlam does not handle graphical material or permit dynamic updating to other programs. NIH does not yet support submitting and receiving forms by e-mail, so printouts are essential; fortunately, the program does print to laser printers, yielding clear copies that comply with NIH formatting guidelines.

GrantSlam manages and balances budgets better than the widely distributed spreadsheet templates, which tend to violate the NIH form-layout requirements after data is entered. The program calculates dollar totals for the detailed and the modular budget pages. Applicants who have incorrectly entered different yearly budget totals on the "Check List" and "Face Page" will appreciate the automation and checking abilities of these forms.

GrantSlam provides excellent contextual guidance, embedded formatting, and simplified budget calculations. Small- to medium-sized departments may be able to justify the cost of the program, but single users will probably prefer the grant templates that are available for free from NIH. GrantSlam will not eliminate the need for expertise in the subtleties of word processing, graphical design, table formatting, and mathematical expressions. Those skills remain in demand until the NIH more fully participates in the use of enterprise software, where paperless forms are more common.

The program is compatible with Windows 3.1 or later and requires a 386 PC, but a Pentium is recommended. Macintosh users must emulate Windows with the Connectix Virtual PC (bundled with the program) to use GrantSlam.

—J. BRUCE MCCALLUM

Department of Anesthesiology, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, USA. E-mail: mccallum@mcw.edu

TECHSIGHTING
SOFTWARE

Sequence Analysis Star

As the sequencing of the human genome approaches completion, the task of analyzing and annotating the finished sequence will become increasingly important. Comprehensive sequence analysis products provide alignment of DNA and protein sequences, primer design, sequence editing, secondary structure prediction, mapping, and database homology analysis. One of the most popular and powerful of these has been the Lasergene sequence analysis software suite published by DNASTAR. In recent years, this package has lost ground in the highly competitive market for sequence analysis packages. Last year, an update to this classic program named Lasergene99 was released. The most noticeable change in this update is the integration of Internet capabilities into the seven different modules that compose the Lasergene99 suite, although the user interface and selected modules have also been improved and a number of bugs have been fixed.

Lasergene99
DNASTAR Inc.
Madison, WI.
Pricing available from DNASTAR.
608-258-7420
www.dnastar.com

Despite these changes, the Lasergene99 package is virtually identical in look and feel to the previous version. The entire package retains a modular format, which allows users to select only those modules useful for their work. Rather than the previous bimonthly release that included CD-ROMs with DNA and protein databases, the new version of Lasergene99 is released annually without the accompanying sequence databases. The rapid growth of the public sequence databases, and their availability online, makes the CD-ROMs unnecessary. The individual Lasergene99 modules are accessed through the Navigator, which also allows the user to access help files for the Lasergene99 suite.

Integration of Internet analyses within the various program modules is a major addition to Lasergene. Sequence files can be directly imported into each module, from the National Center for Biotechnology Information (NCBI) Entrez database, either by accession number or by locus name. NCBI, established in 1988 as a national resource for molecular biology information, also provides the Basic Local Alignment Search Tool (BLAST), which is available from every module of Lasergene through a server at NCBI. BLAST can identify similarities between a nucleotide or protein sequence and other such sequences in the public databases. Entrez sequence retrieval and BLAST searches are both configured in Lasergene to use servers at NCBI, although local versions of the programs can be employed if desired. Lasergene is compatible with BLAST versions 1.4 and 2.0 (which allows the introduction of gaps into the sequence alignments and is thus more sensitive, but slower).

EditSeq is a sequence editor module for importing and exporting DNA and protein sequence files. Although the other modules can import from Entrez, as well as from automated sequencer data in the PE Applied Biosystems (ABI) and Standard Chromatogram Format version 3.0 (SCF) file types, other DNA or protein sequence formats must be converted to a DNASTAR configuration before they can be used. EditSeq allows the export

of sequence data in many file types, such as GenBank Flat File, FASTA File, or Genetics Computer Group (GCG) File. EditSeq is now able to create individual sequence files from a multisequence FASTA File, as well as combine multiple individual sequences into a single multisequence FASTA File. The latter format is useful in conjunction with a number of bioinformatic tools. The program can also find open reading frames in nucleotide sequences and translate them into polypeptides.

SeqMan II is a module that assembles overlapping DNA sequence fragments into a stretch of continuous sequence called a contig. Before assembling the fragments, SeqMan II can remove poor quality data and trim vector or other contaminating sequences. Poor quality data is "masked" in such a way that it can be recovered at a later stage, should it be useful in helping to resolve conflicts or to join contigs. SeqMan II has doubled its capacity and can now assemble up to 64,000 sequences in any given project. Consensus sequence is generated by the use of DNASTAR's Trace Quality Evaluation scheme. The algorithms used to mask the poor quality data and to generate the consensus sequences for the contigs have been updated in Lasergene99. Graphical interfaces are available to display contig coverage and data quality indices and to provide tools for editing the individual sequences within the contigs. Automated and manual tools help determine whether multiple contigs should be joined. Consensus sequence of a contig can be exported in DNASTAR, GenBank Flat File, or FASTA File formats for use in the other modules or in other analysis packages.

MapDraw is a module that creates a variety of linear and circular restriction maps from DNA sequence. The maps can be used for experimental design, sequence analysis, or the presentation of experimental results. MapDraw also has the capability to display annotated features for sequences imported from GenBank. A variety of filters are available to help select restriction enzymes for the analysis. The filters can be combined to select specific enzyme sets that meet multiple criteria. Enzymes can also be selected manually. Enzymes can be added to or deleted from the default library, and information about each enzyme can be modified.

The PrimerSelect module aids in the design and analysis of primers or probes for polymerase chain reaction (PCR), sequencing, and hybridization experiments. The program can use DNA or RNA sequences or it can back-translate a protein sequence. PrimerSelect performs strand-strand melting temperature and hybridization free energy calculations based on a set of user-specified conditions. When possible, multiple primer sequences are displayed in order of suitability for a given experiment based on user-selected criteria. Modifications to primers can be analyzed for effects on translated reading frames, secondary structure, false priming sites, and restriction sites. Once optimal primers have been selected, they can be printed out for oligonucleotide synthesis.

The MegAlign module is used for the construction of pair-wise and multiple alignments of DNA and protein sequences. In addition, the program can construct phylogenetic trees on the basis of the alignments and calculate sequence distances and residue substitutions between the sequences. A number of tools are available to customize the display of the alignments. Similarities or differences between sequences within an alignment can be clearly illustrated, and colored histograms that illustrate sequence similarity or disparity can be created. Alignments can be exported in PAUP or GCG pileup formats.

GeneQuest is the module that is used for the discovery and annotation of genes and other biologically significant features in DNA sequences. This program contains a rich array of tools to characterize unknown DNA sequences. GeneQuest can open DNASTAR, ABI, and GenBank files directly. Sequences in other formats must be converted to one of these formats by the EditSeq program. Sequences are analyzed by specific analytical methods, such as algorithms for finding repeats, finding genes, restriction mapping, pattern matching, and codon prediction. The matrices used to perform the gene prediction methods have been updated to improve the accuracy of the gene-finding process. A default group of methods are presented at the beginning of each analysis. Methods can be added to or removed from an analysis of a given sequence, allowing the user to customize the analysis for each sequence. A summary of the results is presented graphically on a common horizontal scale to facilitate comparison between the different types of analyses performed. When a properly formatted GenBank features table is available, the features from the table are available as annotations. The user can also label regions of interest within the DNA sequence.

Protean is the module used for the analysis and prediction of protein structures. The methods in Protean are grouped by the type of analysis to be performed. Some protein analytical groups may have more than one method, while others are represented by a single method. Current groupings consist of algorithms for the prediction of secondary structure, hydropathy, antigenicity, amphiphilicity, charge density, surface probability, and flexibility. Like the results for DNA analysis, the results for the analysis of a protein sequence are plotted with a common horizontal scale. Protean can also simulate protease digests resolved on SDS–polyacrylamide gels, calculate and display titration curves, and create models of secondary structures. Protean also has a summary screen that provides numerous statistics about the protein sequence as well as a breakdown of the protein composition to its amino acids.

Installation of the package was trouble free, but it requires the use of both a floppy disk drive and a CD-ROM drive, which might be a problem for some users. Although the modular nature of the program can be disruptive at times, the design across modules ultimately works well and the workflow remains efficient.

Running the programs is virtually intuitive. The Macintosh version of Lasergene99 was used for this review, and the programs conform to the general look and feel of the Macintosh user interface. Most of the analyses can be quickly mastered, although the powerful analytical methods in GeneQuest and Protean may take more time. The online help provided with the modules is usually useful. For most of the methods, a summary of purpose is provided. Ample documentation is provided with the package. One manual describes the installation process and provides quick tutorials for the different modules. A second manual describes the features that have been added to Lasergene99. Finally, another manual documents the various features in the modules. Despite extensive documentation describing how to use Lasergene99, the best sources of information are in the originally published scientific papers.

Lasergene99 is available for Macintosh and Windows 95, 98, and NT (4.0 or later) platforms. The minimum system requirements for Macintosh are System 7.0 or later (Power Macintosh recommended), CD-ROM drive, 8 MB RAM (32 MB recommended), and 40 MB of free hard disk space. The minimum system requirements for Windows are a Pentium 100 MHz processor, a CD-ROM drive, 32 MB RAM (64 MB recommended), and 40 MB of free hard disk space. Internet access is recommended for both platforms.

—**DARRYL NISHIMURA**

Department of Pediatrics, 440G EMRB, University of Iowa, Iowa City, IA 52242, USA. E-mail: darryl-nishimura@uiowa.edu

Tech.Sight is published in the third issue of each month. Contributing editor: Kevin Ahern, Department of Biochemistry and Biophysics, Oregon State University. Send your comments by e-mail to techsight@aaas.org.