NEW MEDIA: SOFTWARE

Comprehensive Sequence Analysis

James P. Calvet

We are now adept at determining genomic and complementary DNA sequences and consequently are awash in DNA and protein sequence data. The need for analyses of these sequences can be simple (finding a match to a specific sequence of nucleotides in DNA) or complex (aligning multiple protein amino acid sequences that are related, but not identical). The software, which has evolved to meet these needs, ranges from specialty applications that provide a handful of specific functions (*1*) to so-called comprehensive products that attempt to cover the full range of user needs. As one might expect, niche products offer considerably greater depth of function than comprehensive ones, but comprehensive products often better meet the diverse sequence analysis needs of workers in a laboratory.

OMIGA 1.1
Oxford Molecular Group Inc.
Oxford, UK. Industry $1995; academic/government $1495.
Phone: (800) 876-9994
products@oxmol.com

Oxford Molecular's OMIGA 1.1 sequence analysis package for the Windows 95 or NT operating systems is one such comprehensive software package. All of the basic functions of the program—importing sequence data, analysis, editing, organizing, formatting, and exporting—operate in an intuitive and user-friendly manner. The software contains an impressive battery of functions for nucleic acid and protein sequence analysis that will provide compositional data, locate motifs, predict structures, or align related sequences. Relevant data from several such analyses can be integrated into a single on-screen figure to display a comprehensive sequence analysis profile. OMIGA will also help to generate figures for slides or for publication. The three strongest points of this software are its ability to organize projects, its power and versatility, and its flexibility and ease of use once the basics are mastered.

In order to use the program, DNA or protein sequence data can be entered in several ways. OMIGA imports sequence data from all the major file formats, including GenBank, EMBL, GCG (Genetics Computing Group), MacVector, SwissProt, and IG-Suite. It also accepts nucleic acid and protein sequence information from an online BLAST search that has been copied and pasted from the BLAST screen to OMIGA. Unrecognized symbols are discarded by the program and the sequence is automatically converted to the GCG format. Sequences may also be imported as files created in word processing programs, if they have been saved as ASCII text documents. Export options from OMIGA include these same major sequence formats. Figures created in the program can be exported either as Bitmap files or as Windows Metafiles.

One of the best features of OMIGA 1.1 is its use of the Windows Explorer file format for organizing files and data. The Explorer-like window has three panels: a Tree panel that contains a list of data files within each project folder, a Summary panel showing a list of sequence files, and an Attributes panel containing a detailed description of any highlighted sequence. This so-called Project View is the starting point for any analysis. After bringing the Project View onto the screen, one selects a sequence from the Summary panel to run an analysis. The outcome is automatically saved as a new file that fits logically into the Tree panel, from which it can easily be recalled as needed. Tree panel files are given sufficiently descriptive names to make them easy to find at a later time.

A related benefit of the automatic project-based system is that various kinds of data can be integrated into a single view, or Feature Map. For example, a DNA sequence can be displayed along with a variety of analyses of that sequence, such as a restriction map, open reading frame, and amino acid sequence, in a single view. The Windows Explorer format facilitates this process by having a panel in view on the left of the screen that lists all of the available features. It is then a simple process to point and click on each feature (putting a check in a box) that is desired for the map that is being built. Maps containing different features can then be named and saved. Editing a sequence file automatically modifies all derivative files, and the program will alert the user if the searches and calculations have to be redone.

The OMIGA 1.1 package is loaded with sequence analysis options, giving it a lot of versatility. Restriction enzyme cut sites are identified in DNA sequences with REBASE. The most recent update of this database (and the others used in the package) can be downloaded from the Oxford Molecular Web site as desired. Numerous user-selected search options are available, making it possible, for example, to look specifically for six-base cutters, blunt-end cutters, or specific overhangs. Restriction maps may be viewed on the screen in their entirety or can be zoomed to the nucleotide level, showing the sequence surrounding the cut sites. Details about the sizes of restriction fragments can be obtained, and a list of sites is generated that includes the enzymes that will not cut a sequence—a useful piece of information for many cloning procedures.

Open reading frames in a DNA sequence are located with GENMOTIFS. Structural elements, like Z-DNA–favoring sequences, and functional sites, such as promoter elements, are found with NASITE. OMIGA can also identify optimal sequencing or polymerase chain reaction primer sites in DNA sequences by using default or user-defined parameters that include primer length, product length, G/C content, and melting temperature. Data are displayed in either tabular form or as a map, the latter indicating which primers are optimal (in red) and which are suboptimal (in blue). A helpful feature of the sequence maps is that informational pop-ups can be made to appear when specific sites are highlighted. This enables one to rapidly scan a map for detailed information about structural elements or functional sites before zooming down to the nucleotide sequence.

OMIGA will provide basic information about protein sequences, such as amino acid composition, isoelectric point, and molecular weight. It will also predict several structural and chemical features (referred to in the program as property profiles). These include antigenicity (four algorithms), hydrophobicity (seven), transmembrane domains (two), and protein flexibility (one). The Chou-Fasman and the GOR II secondary structure algorithms are also available. Output from the analyses shows the amino acid sequence being analyzed together with colorful symbols depicting regions predicted to have any of the various structural elements.

Proteolytic cut sites (both chemical and enzymatic) in amino acid sequences are located with PABASE. Protein motifs are identified by using amino acid sequence information in the PROSITE sequence database. For example, PROSITE contains more than 40 amino acid sequence motifs that function as protein kinase phosphorylation sites. OMIGA will use that sequence information to identify putative protein kinase phosphorylation sites in a user's sequence. The program will identify all such sites, or users may restrict the search to only one or a selection of kinases. OMIGA allows users to alter a preex-

The author is in the Department of Biochemistry and Molecular Biology, University of Kansas Medical Center, Kansas City, KS 66160–7421, USA. E-mail: jcalvet@kumc.edu

isting motif or to create their own, enabling new motifs to be identified as desired. The new motifs are treated in the same way as any of the preexisting database options and can be given a name and placed on any Feature Map. Almost any entry or any search function in the program can be edited, and once the searches are done, they can be fine-tuned by filtering the results so that only selected data sets are displayed.

OMIGA's ability to perform alignment of multiple related protein sequences (up to 500) with the CLUSTAL W algorithm demonstrates the power of the program. Amino acid residues are color-coded to depict their properties (for example, acidic or basic) and can be viewed with a corresponding secondary structure analysis. Many aspects of the display can be edited, including the color-coding. In addition, the alignments can be adjusted by the user, residue by residue, to specifically remove or introduce gaps, giving one freedom to experiment with alignments outside the constraints of the program.

OMIGA 1.1 is supported by full documentation and online help. The on-screen help files are comprehensive and easy to follow, although the set of instructions for how to carry out a secondary structure analysis of an amino acid sequence did not work in our trial. The full User Guide and Reference Manual can be viewed on screen or printed at high resolution. In general, the OMIGA 1.1 package is free of bugs, with one notable exception. The software was tested with numerous DNA and protein sequences of various lengths, which it handled flawlessly. Nevertheless, the program consistently caused crashes whenever a compositional analysis was performed in a fully expanded window. Of serious concern was that it purged all of the project files one of the times it crashed. The compositional analysis can also be done in an unexpanded window, so it is possible to avoid crashes, but this bug should be fixed.

OMIGA 1.1 is a powerful and comprehensive sequence analysis package that is very easy to learn and use. Any user with some experience with Windows 95, or with other sequence analysis programs, can quickly adapt to the software and immediately find it useful. Although the novice may have some difficulty using the product at first, the intuitive feel of OMIGA, together with the excellent help files, should make the experience quite painless.

Reference

1. Available from Textco at www.textco.com.

PERSPECTIVES: ANCIENT METALLURGY

Metallic Reflections

Jeffrey Quilter

Recent discoveries of gold tombs, such as the Lord of Sipán of the Moche culture, hailed as the "New World's King Tut," and the remarkably preserved ice mummies of sacrificial Inka maidens, have refocused attention on ancient Peru's rich cultural legacy (*1*, *2*). Both, however, were late blooms attached to roots deep in the ancient past, exemplified by early maritime settlements in Peru (*3*). Now, as reported on page 1108 of this issue, Burger and Gordon (*4*) have uncovered the earliest New World metallurgy—cold-worked copper foils—providing the opportunity for a comparative examination of the development of this important technology. The new findings also reveal the unique ways in which ancient Peruvians used technology for their own means and ends.

Enhanced online at www.sciencemag.org/cgi/content/full/282/5391/1058

The study of the prehistoric civilizations of ancient America has often swung between two poles: a search for generalized stages of development in comparison with the Old World or the perception of New World societies as unique. Nineteenth century views of cultural evolution did not fit comfortably with societies such as the Maya, who had a full written language and excelled in art and architecture yet had little in the way of metallurgy, or the Inca, who ran a vast empire and had advanced metal skills but no writing. Did New World cultures follow unique cultural trajectories? Were they influenced by outside civilizations? And were they comparable in any way with cultural developments known elsewhere? In the case of Andean metallurgy, the answers to these questions seem to be that the technology was a local development that followed its own path, both similar to and different from that of the Old World.

Metallic sunshine. Central American gold disk, demonstrating the use of the metal's reflective properties.

PHOTOS: COURTESY OF DUMBARTON OAKS LIBRARY

The author is in the Pre-Columbian Studies Department, Dumbarton Oaks, Washington, DC 20007, USA. E-mail: quilterj@doaks.org

The early date determined by Burger and Gordon for the cold-hammered copper foil at Mina Perdida (*4*) strongly suggests that metallurgy had a long period of development in America and was neither a transoceanic import nor developed full-blown, ex nihilo. Although New World metallurgy had its own origins, it shared many patterns with the Old World in its subsequent development. In both hemispheres, cold working was eventually augmented with casting (both mold and lost wax) and elaborate assembly techniques, including folding, bending, hammering, and the use of staples and solder. Furthermore, the first purposes to which metal technology were applied were not industrial but rather tied to social and ideological concerns. In Central Europe, precious metal jewelry enhanced the status of chiefs and similar social leaders (*5*). At Mina Perdida, foil was likely used to elaborate ritual display, probably also in aid of a growing, although still relatively weak, social elite.

We do not know exactly what kinds of rituals took place on top of the pyramids at Mina Perdida, but they almost certainly involved activities that reenacted fundamental social myths with much spectacle and display. Part of the evidence for this is the remarkable puppetlike figurine that Burger and Salazar-Burger (*6*) found on the same terrace where four of the copper fragments (*4*) were retrieved. The fact that the recov-