TECHVIEW
SOFTWARE

Planning Plasmids

David R. Wessner

Vector NTI Suite from InforMax fulfills three important needs of molecular biologists: cataloging a growing number of complex plasmid molecules, analyzing and comparing the molecules, and simplifying the design of new recombinant molecules. In addition, Vector NTI Suite provides seamless Internet connectivity and can be used effectively as a learning tool for beginning students of molecular biology.

Vector NTI Suite
InforMax, Inc.
Bethesda, MD.
Price available on request.
800-357-3114
www.informaxinc.com

Vector NTI Suite consists of four independent, yet interconnected components: Vector NTI, a molecular biology program; Align X, a multiple sequence alignment program; BioPlot, a sequence analyzer program; and ContigExpress, an application for sequence assembly and management. The strength of the suite lies in Vector NTI, which allows one to analyze existing biomolecules and to design the construction of new ones. Within Vector NTI, molecules can be placed into one of five different sets: nucleic acid molecules, protein molecules, restriction endonucleases, oligonucleotides, or gel markers. Sets of common molecules for each category are preloaded, and the user may add new molecules to each database. For DNA or RNA, new molecules may be imported from outside databanks, created from scratch with the use of Vector NTI's design capabilities, or constructed from other molecules. Proteins can be imported, user-specified, or translated from a DNA or RNA sequence. Oligonucleotides can be either designed by the user or generated by Vector NTI. Restriction endonuclease recognition sites and gel markers can be created from scratch and loaded manually or imported from outside sources.

The most useful feature of Vector NTI is its plasmid-related functions. The ultimate goal of many recombinant DNA projects is to remove a specific piece of DNA from one source and insert it into another. With Vector NTI's ability to automate cloning strategies, the task of designing plasmids on laboratory white boards is a thing of the past. This function of the program is most clearly manifested in the generation of cloning strategies, the identification of acceptable PCR primers, and the identification of acceptable hybridization probes. For most cloning operations, the first step involves the identification of useful restriction sites that are present in both the donor and recipient plasmids. Restriction endonuclease sites must be identified that will digest the donor molecule on both ends of the fragment to be cloned, but not elsewhere within the molecule. Ideally, these restriction endonuclease sites should also be present in the recipient plasmid at the desired location. In the graphics pane of the plasmid display window (Fig. 1) of Vector NTI, one can highlight the fragment of interest and define the recipient plasmid. The program will then determine the most feasible cloning strategy. One can specify according to personal preferences that only cohesive-cohesive strategies, for instance, be considered or that only a limited set of restriction endonucleases, such as the ones currently stocked in the freezer, be considered. In this manner, a cloning strategy that matches one's resources will be generated. The text pane then displays a detailed account of the steps required to conduct the desired cloning, including a list of restriction endonuclease digestions, type of ligation (cohesive-cohesive or cohesive-blunt for example), and recommended oligonucleotides for PCR and hybridization identification of recombinant clones. One shortcoming, though, is that Vector NTI does not provide a graphical flowchart of the required cloning steps.

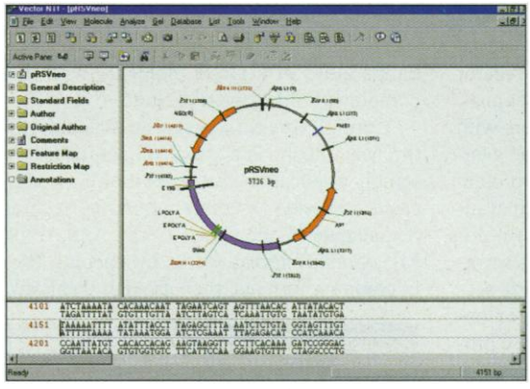

Fig. 1. Plasmid display window. When a nucleic acid or protein plasmid in the database is opened, a three-paned plasmid display window is displayed. A graphical representation of the plasmid, its complete nucleotide or amino acid sequence, and textual information about the plasmid are shown.

As molecules are manipulated, Vector NTI maintains parent-descendant connections. This feature allows one to keep track of the origin, or lineage, of recombinant molecules. In addition, one can assign key words to certain molecules or place them in defined sub-clone databases, thereby categorizing the molecules and making future access of them more efficient.

When selected from the database, information about the plasmid and its features are presented in a window that has three panes: a text pane, which contains a description of the molecule; a sequence pane, which displays the entire sequence of the molecule; and a graphics pane, which displays an image of the molecule. The relative sizes of these panes can be adjusted easily. Clicking within a pane activates that area. One also can switch between active panes by clicking on a toggle icon located on the display window toolbar, but the small size of this icon (even on a 17-inch monitor) makes it difficult to determine the toggle's position. It also seems the relative sizes of the panes should revert to a default setting, instead of resizing the panes upon returning to a task.

Within the graphics pane, important features of the molecule, including restriction sites, promoter regions, and open reading frames (ORFs) of interest are displayed. It is possible to alter the labeling and color schemes of these highlighted regions, thereby customizing the graphic for one's intended purposes. If one clicks on a specific feature, a promoter sequence for instance, information about that feature will be presented, such as its exact nucleotide location. At the same time, the sequence corresponding to that feature will be highlighted in the sequence window. The text pane provides important information and user input notes about the molecule. For plasmids, general information is provided, such as the source of the vector and its size. In addition, descriptions of features of the molecule are supplied. Descriptions of restriction sites, for example, contain the number of sites for a given endonuclease, their locations, and the recognition site of the enzyme.

Vector NTI also provides a functional, easy-to-use polymerase chain reaction (PCR) primer identification package (Fig. 2). By simply highlighting the region of DNA to be amplified by PCR, the user can instruct Vector NTI to generate a series of acceptable primer pairs. Salient information, such as locations of primers and length of predicted product, is reported for each pair. Again, various features of the oligonucleotide primers, such as acceptable minimum and maximum lengths and desired annealing temperatures, can be specified. One can also specify the importance of duplex melting temperature (T_m) difference between members of a pair, the likelihood of primer-dimer formation, and the presence of potential in-

The author is in the Department of Biology, Davidson College, P.O. Box 1719, Davidson, NC 28036, USA. E-mail: dawessner@davidson.edu

tramolecular base pair formation before instructing the software to produce a set of primers. In this respect, Vector NTI is not unique. Several programs are available that perform this function equally well. Few, however, integrate this feature into an all-encompassing molecular biology package.

In most cloning and PCR procedures, it is necessary, at some point in the process, to confirm the identity of a recombinant molecule or PCR product through hybridization with a specific oligonucleotide probe. As with the identification of PCR primers, this process can be more challenging than it may first appear. Various criteria, including homology, hybridization temperature, and location of the probe, may vary between experiments. With Vector NTI, the task of identifying potential hybridization probes that meet these criteria becomes trivial. Furthermore, a given oligonucleotide's usefulness as a hybridization probe can be assessed by the program.

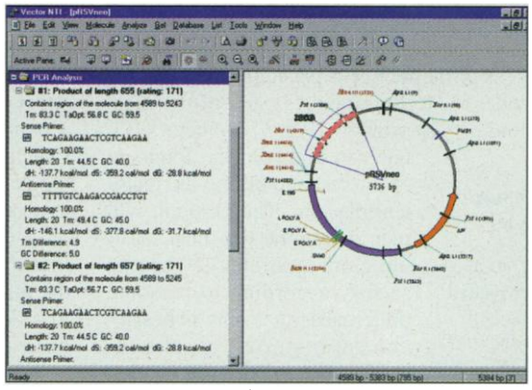

Fig. 2. PCR primer analysis. If PCR amplification of a DNA fragment is desired, Vector NTI can generate a series of acceptable oligonucleotide primers based on user-defined criteria (such as preferred oligonucleotide lengths, annealing temperatures, and so forth). As shown in this figure, various attributes of the generated primer pairs and the resulting PCR product are displayed.

Vector NTI also competes effectively with other molecular biology software products, such as MacVector and LaserGene, in simplifying Internet analyses. The program provides researchers with access to a series of search engines and analysis programs. For nucleic acid and protein molecules, for instance, one can search available databases like GenBank for homologous entries. Likewise, one can access several Internet programs designed to analyze secondary structures and hydrophilicity or hydrophobicity of proteins. Although these programs are available through a browser, the tight integration between Vector NTI and these Internet sites considerably increases their ease of use. Transitions between Vector NTI and the required World Wide Web sites appeared seamless. As noted in the manual, Web servers continually change. Thus, the examples provided in the manual may not remain correct. Also, if URLs change, the transparent links provided by Vector NTI may become outdated. Additions and corrections to these links are maintained on the Vector NTI home page.

The Vector NTI Suite also contains three additional molecular biology programs, Align X, BioPlot, and ContigExpress. Align X allows the user to align multiple nucleic acid or protein sequences. Once sequences are selected for alignment, Align X produces a display window that has four panes: a text pane, an alignment pane, a phylogenetic tree pane, and a graphical pane. Within the graphical pane, three graphs are displayed that show the quality of alignment, the statistical significance of alignment, and the statistical significance of alignment between a given molecule and the consensus sequence. With BioPlot, linear plots of various DNA, RNA, or protein properties can be generated. For instance, one can display the GC base pair content or the melting temperature of a DNA molecule with this component of Vector NTI Suite.

ContigExpress (soon to be available for the Macintosh) is a powerful sequence assembly application that is as useful as dedicated sequence assembly programs, such as Sequencher (Gene Codes Corp., Ann Arbor, MI). ContigExpress works by opening files in common formats, such as ABI, GenBank, FASTA, and EMBL, and assembling the DNA sequences into a "contig" (contiguous sequence file) by finding overlapping regions. The settings used by the program to identify overlaps can be modified as desired.

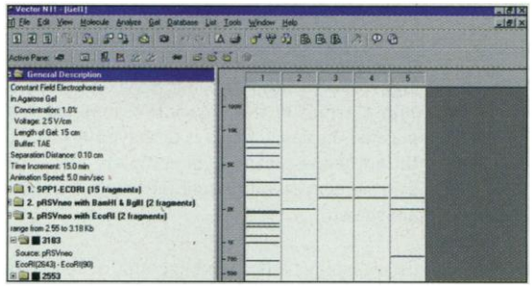

Fig. 3. Virtual gel electrophoresis. Animations of gels containing user-defined molecular weight markers and samples can be displayed. Plasmids can be "digested" with a desired restriction endonuclease and then "run" on a virtual gel. Various attributes of the gel, including buffer, concentration of agarose or polyacrylamide, and voltage, can be altered. Time required for adequate separation of selected fragments also can be displayed.

Several features of Vector NTI could make it quite useful in the classroom as well as the laboratory. An interactive medium for restriction digestion and electrophoresis, as provided in Vector NTI, could dramatically improve a novice's understanding of these basic procedures (Fig. 3). With the Create Gel Sample function, one can specify a plasmid of interest and subject that molecule to digestion with any of the available restriction endonucleases. After creating the gel sample, a researcher or student can load it, along with a desired gel maker, onto a virtual gel. Numerous attributes of this gel, including buffer, concentration, and voltage can be altered. A run of the samples on the gel then can be animated. By altering various parameters, from the restriction endonucleases used for a digestion to the concentration of agarose in the gel, students can quickly grasp the basic theories of cutting and separating DNA fragments.

The manual for Vector NTI Suite is well written and quite clear. The first several chapters provide thorough, easy-to-follow tutorials, which offer an excellent introduction to the various capabilities of the software. Within the Vector NTI portion of the manual, a much larger emphasis is placed on manipulations of DNA and RNA molecules than on proteins. To a large extent, this disparity in coverage simply reflects the relative importance of and number of manipulations available for plasmids and other nucleic acid vectors. The sections of the manual devoted to the other components of this suite consist mainly of brief descriptions of the functions present in these packages. No tutorials are provided for them. It is clear that these components of the suite, while useful, are not the major attractions of this software.

Vector NTI Suite is a very solid program for both researchers and educators. In a market flush with software that claims to take the headaches out of cloning, Vector NTI delivers. The programs are user-friendly and adaptable. The molecule design and oligonucleotide-generating features shine the brightest and will greatly assist researchers in experimental design of recombinant molecules. The cataloging properties are impressive as well, allowing workers to track laboratory reagent inventory and link it with the design of new experiments. In concert, these features make Vector NTI a worthwhile addition to the molecular biology laboratory.

Vector NTI 5.0 is supplied as a single CD-ROM for Windows 95, 98, or NT or for Macintosh platforms, and its installation is simple. System requirements include 20 MB of hard drive space and 32 MB RAM. An Internet connection is required for some features.