

Visual Cloning 2000: A DNA Sequence Analysis Program

Bernhard Kusian* and Botho Bowien

Institut für Mikrobiologie und Genetik,
Georg-August-Universität Göttingen, Grisebachstrasse 8,
37077 Göttingen, Germany

Abstract

VC2000 is a recommendable, easy-to-use and affordable program suited particularly to plot maps of both plasmids and linear DNA fragments for the documentation of cloning procedures. It comes with basic tools for sequence analysis that assist this major utility. A special feature of the program is the integration of web-based tools, providing additional analytical power and flexibility. In this way the user is supplied with the latest versions of these applications.

Redasoft (Toronto, Canada) recently introduced the software Visual Cloning 2000 (VC2000) as a relatively small and easy-to-use tool for DNA sequence handling and analysis. The program can be downloaded as a 6.6-MB executable file from the Redasoft internet home page (<http://www.redasoft.com>). Its installation on a PC running Microsoft Windows 98 presented no problem although it required an update to the latest version of the Microsoft XML parser that was performed automatically by the installation routine via the internet. However, the installation on a Microsoft Windows NT4 (service pack 6) system failed at first attempt. This problem was solved only after a complete reinstallation of the Microsoft Internet Explorer (version 5.5), indicating that VC2000 interacts extensively with the web browser.

VC2000 consists of seven subprograms directly accessible from the program window, allowing fast switching between the different modules. Map View is the central program section displaying the graphical map of a sequence, which can be modified by deletion or insertion of segments. Fragments to be inserted may originate from either existing maps or DNA sequences files. The generation of maps is assisted by analysis data obtained by the other program sections described below. The import of sequences from public databases is accomplished easily by means of the web browser integrated into the Internet section of VC2000. The features of imported sequences are represented in the map as colored bars or arrows together with the feature labels. Style, size and/or color of features and labels are editable. However, it is not possible to mix styles within labels as required for e.g. the proper designation of restriction endonucleases (*HindIII*) in publications. An automatic rearrangement function would be helpful in fitting a complete map to page size in cases

where the attached feature labels are placed beyond the upper margin. Such minor weaknesses of the program can be tolerated if the maps are primarily intended for lab documentation. Maps in publication quality will definitely require subsequent editing, but the bitmap files exported by VC2000 are not suitable for this task. Whereas circular plasmid maps are visualized nicely, those of larger, fully annotated linear sequences from databases might cause confusion because all features are shown simultaneously in a very compressed layout. An extended zooming function together with the possibilities to split up the sequence and to select the types and hierarchy of features to be shown would provide much more comfort. Surprisingly, the scaling does not offer the use of the metric system.

The Sequence Viewer serves to display a DNA sequence along with the corresponding map. It allows to highlight a sequence region and to copy it into other maps generated by VC2000 or into other programs, but sequence manipulations or modifications are not possible. The latter require editing of the sequence by an external editor and subsequent reimporting with the consequence of a partial or complete loss of the originally attached features. Therefore, documentation of modified clones is not very convenient.

The Restriction Analysis subprogram makes use of restriction enzymes compiled in the Rebase database (<http://rebase.neb.com/rebase/rebase.html>). Two sets of enzymes are accessed with one including all known restriction enzymes and the other only those available commercially. In addition information related to the enzymes, like the occurrence of isoschizomers and the commercial sources, may be obtained. The enzymes are selected individually or as customized sets. Such selections can be defined further by setting the cutting frequency within the target DNA sequence, the size of the recognition sequence, and the type of fragment ends generated after digestion as well as by limiting the analysis to a specified region of the sequence. Restriction sites identified are shown in the map. It is possible to sort the list of detected restriction sites by the name of the enzyme, the position of the site, the type of recognition site, or the type of ends. Functions missed, however, were a listing of restriction fragments produced and the possibility to simulate multiple digestions for evaluating restriction maps of new constructs.

Potential protein coding regions within DNA sequences are identified employing the Open Reading Frames module. Searches may be restricted to defined regions and/or specified frames. Further selectable parameters include the minimal size of an open reading frame (ORF) and alternative start codons or genetic codes. Detected ORFs are listed by number (ORF1 etc.), position (start, end), and length (bp) along with the length (number of amino acid residues) and calculated molecular weight of the putative peptide product. An ORF marked in the list can be added to the map, but renaming it within the map will not be acknowledged in the list. Consequently, the program does not accept names of ORFs/genes in annotated sequences.

*For correspondence. Email bkusian@gwdg.de; Tel. +49 551 393828; Fax. +49 551 393793.

Various functions for protein sequence analysis are provided by referring the sequence to Expasy (<http://www.expasy.ch/tools>) via the Redasoft web site. The transfer of individual amino acid sequences to other programs is possible by copying and pasting, however, saving within VC2000 is not implemented. Also, access to amino acid sequences of ORFs in existing maps requires a new search.

The Primer and Oligonucleotide calculator is a web-based application integrated into VC2000. The settings for this subprogram allow the selection of all important parameters necessary for the design of oligonucleotides to be used in PCR or as probes for hybridizations. Target and search ranges within a sequence can also be specified, although the search range is limited to a maximal size of 10 kb. Larger sequences are not accepted even if the range is restricted to this size. Primer sites may be added to the map. Another useful subprogram is the Subsequence Search. It allows searches on the sequence displayed as map. The search range can be confined to specified segments such as the sequence features or by entering a region. The sequence sections are listed by name, relative orientation as well as their start and end positions. Again, they may be integrated into the map. All analytical functions of VC2000 lack the possibility to save, print, or copy the lists generated as outputs, thereby limiting their use to the current session of the program. The Redasoft web site (<http://www.redasoft.com/rsn>) referred to by the browser offers several links to sequence databases and to other molecular biology resources such as the ones to lab protocols (<http://www.protocol-online.net/Protocol.htm>) and to the Redasoft Sequence Converter or the Formatter. The links for online ordering restriction enzymes or oligonucleotides would be even more valuable if they also included manufacturers based outside of North America. Redasoft should make any effort in keeping the page up-to-date to maintain the utility of the program in the future.

In summary, VC2000 is a recommendable, easy-to-use and, last but not least, affordable program suited particularly to plot maps of both plasmids and linear DNA fragments for the documentation of cloning procedures. It comes with basic tools for sequence analysis that assist this major utility. A special feature of the program is the integration of web-based tools, providing additional analytical power and flexibility. In this way the user is supplied with the latest versions of these applications. VC2000 is offered at a reasonable price of US\$ 399 for academic or US\$ 599 for corporate customers. Update subscriptions are available for US\$ 20 or 30 per month, respectively.