# Free Software in Molecular Biology

for Macintosh and MS Windows computers

Don Gilbert

June 1998

archive@bio.indiana.edu
Biology Department
Indiana University, Bloomington, IN 47405

# 1. Introduction

You will find a large collection of free applications for molecular biology and chemistry through Internet servers and from other sources. Most of these are written by biologists, chemists, and software developers, sometimes as part of a university or government funded project, sometimes as an unfunded project. Some of these offer a single function useful to you, but one that may be found nowhere else.

What is free software for bioscientists, and especially where can you get it? How do you set it up and use it? What are some good programs, and what overall is available? These are some of the questions you can find answers to in this review, which concentrates on software for molecular biology for the commonly used Macintosh and Microsoft (Wintel) operating systems.

What you won't find answers to here include the many free programs that now run as an Internet service through your web browser. Neither listed are programs available mainly as source code (not ready to run) or Unix programs, nor the choices you have with commercial software, nor software for related areas of chemistry, medicine, population biology, ecology and others.

## 1.1 Free or commercial software?

Free data analysis software is common in the sciences, as the scientists in need of new analyses develop algorithms for it, then crystallize the algorithms as software. Most of the basic biosequence analyses are scientist-developed, including FastA, BLAST, Clustal, MFOLD, PHYLIP, Paup, CAP, to name a few. The source code of these is often shared freely. But often these algorithm programs lack ease of use and integration with other functions. Commercial software developers have incorporated such algorithms, along with their own, and added a much greater usability and integration, to allow you to analyze your data without spending oodles of time learning how to run the programs.

Besides adding integration and user-interface, companies add the great value of good documentation, telephone and other support for their wares. When your funds permit, and a commercial package does what you need, it is usually a better choice than free software. Given market place realities, companies charge what seem large fees to you, but these fees are needed to cover the many aspects of publishing: advertising, technical support, software developers, and so forth.

For scientists with limited budgets, students and teachers, free software is often the only choice. You can also find unique programs that do things no commercial package does. Another great advantage of free software is that it often includes source code you can use to modify and extend an analysis.

Today with the rapid growth in the fields of bioinformatics and biocomputing, more good programmers are developing software, many making it freely available. You will find more sophistication in attention to user interfaces and ease of your learning to use the software. Still there is no common means for funding development of free software: government agencies do not generally fund projects from individual research/programmers, the main source of many of the free software packages. The shareware concept of users paying for software they use has never worked well. For most scientific software, the potential market is so small that we see a strong distinction between expensive commercial packages and free software.

Usability of free programs is variable, and depends in part on your computer and needs. Be aware that older programs may fail to work on new computers. Tolerance for program flaws and limitations is generally needed, and self-reliance in learning how they work. Often the authors either cannot be contacted, or do not have time to spare for supporting a product that they freely distribute.

## 1.2 Internet sources

Access to and use of the Internet has become so ubiquitous in our society, especially among scientists, that it is assumed you have or can get such access. This wasn't true a few years back, but since the 1980s, free software for biosciences has been most widely available through the Internet, since distribution that way is free (from cost and time) to software authors. There are a few widely used archives of biosciences software, and today many authors prefer to provide their own network server to distribute software, as that is very

easy to do. Archives which collect such software still play an important role in providing these collections, and in archiving them for years past when an author may be able to.

Two commonly used archives of molecular biology software are IUBio Archive, at Indiana University, and the European Bioinformatics Institute (EBI) software archive. Internet resource locators (URLs) for these are

**EBI** at **http://www.ebi.ac.uk/ or ftp://ftp.ebi.ac.uk/**

The European Bioinformatics Institute, home to EMBL databank and others, and a home to a large molecular biology software archive, including the very useful Bio Catalog of software (http://www.ebi.-ac.uk/biocat/biocat.html).

**IUBio** at **http or ftp://iubio.bio.indiana.edu/**

IUBio Archive is home to a large collection of biology software, and also provides services for keyword searching of current GenBank, SwissProt and PIR databanks, and the useful Bionet network news archive. It has been operating since 1989, as a user- and author- supported self-service archive. The molecular biology software collection at IUBio is also mirrored (copied) to sites around the world, including the Finland, Sweden, Japan, United Kingdom, France, Spain, and Israel. See the following software listings section for URLs to these.

A few of many other Internet servers of note include

**http://www.ncbi.nlm.nih.gov/** - home site for up-to-date access to Entrez software (databank lookups), Sequin for publishing your sequence in GenBank, Macaw (multiple alignment), and others.

**http://kiwi.imgen.bcm.tmc.edu:8088/search-launcher/launcher.html** - Search Launcher at Baylor College of Medicine, home to a selection of sequence analysis methods. Recommended is their search launcher program (a Perl script that is usable on Macintosh and Wintel systems) for batch access to this valuable resource.

**http://expasy.hcuge.ch/** - home of SwissProt, ProSite and home to a collection of useful database analyses functions, especially for protein data.

## 1.3 Fetching software

File Transfer Protocol (FTP) is an Internet service designed specifically for file transfer. It is an antecedent of the popular Hypertext Transport Protocol (HTTP), or web, Internet method. Generally FTP is still a better method for transferring large files such as software packages, than is a web browser. One popular and easy to use FTP program for MacOS is the *Fetch* program by Jim Matthews of Dartmouth College. Find more information at http://www.dartmouth.edu/netsoftware/fetch.html. The ubiquitous *Netscape* or other web browser also lets you fetch software either via HTTP or FTP methods.

Generally software is stored and transferred to you in encoded formats from archives. Most current fetching software, like Fetch and Netscape, will decode this automatically for you. However, other software sometimes is used or needed for decoding. General Internet software packages and servers will get you started on what is needed. If in doubt, your local library or bookstore will have a few shelves full of books on using Internet services and general software archives.

The main trick in fetching software these days is knowing where to look. Besides this review, you can make use of web links made by others, perhaps with better perspective than this author. See also the common Internet search services, like Yahoo (http://www.yahoo.com/), AltaVista (http://www.altavista.-digital.com/), Lycos (http://www.lycos.com/) or others.

Keep in mind that software is updated; if the version you try fails in some way, a newer version may be available now or soon. The author's preferred or home server is the best place to check for updates, as archive servers don't always have the current release.

## 1.4 Setup and use

Once on your computer, the steps of installing and configuring free software range from easy to trying. Instructions are usually included, but are not always detailed enough, or cover all the range of problems one can run into. Special installing software is common with commercial software, but not with the free variety.

Report problems you have to the author, and suggestions for improving the software. Asking the frequently overworked authors for help with installation isn't always good solution though, so attention to their pre-written installation directions is essential.

Special kinds of software, especially those written for Java or Perl, will require that you also fetch and install other free, general software to use. In the case of Java, this is becoming less a problem as new MacOS and UNIX systems are shipping with a Java runtime as part of the system.

## 1.5 Multi-platform software

By 1998 counts of Internet browser contacts at IUBio and other biology web servers, from 30% to 50% of biologists use Macintosh computers, 40% to 70% use Wintel systems, with somewhat less than 10% using XWindows systems as their workstation (though many use UNIX or VMS for other things). Bioscientists remain diverse in their computing system choices and needs for software. Many use or have access to multiple operating systems, depending on what software runs on. Also of help, there are commercial emulator programs that allows one to run most MS Windows and MSDOS software on Macintoshes, and conversely to run some Macintosh programs on Wintel systems

Some software is written to run on many operating systems. This may be a holy grail of those who develop scientific applications - one hopes a program can be used by anyone needing it, on any computer system. It isn't as easy to write multi-platform software as might be. Outside of the recent arrival of Java, there hasn't been an easy, good way to do this inexpensively. Even large commercial developers expend effort that is not always successful (e.g., WordPerfect on UNIX pales compared to the MacOS or Wintel versions).

With the case of graphical interface software that most of us expect today, multi-platform programs may not look quite right on your particular system, even though they operate as intended. In the more common case of software with no graphical interface, you may need to devote extra time to learning its command-line or menu driven syntax.

The new Java development system born at Sun Microsystems (http://www.javasoft.com), now is providing a means for developing useful software that works well on common systems, albeit at present Java software is frequently slower than its counterpart written in C++, C or other languages. We may expect to see much more software in biocomputing written in Java in coming years. See for instance the new sequencing analysis package from Licor (www.licor.com).

## 1.6 Client-Server biosequence software

There are various developers working on the concept of separating the user interface from the analysis programs, and I think this is be a useful approach to making such programs easier to use. This is the basis of a client-server design for software. Simple examples of this are web interfaces that abound now for various data analyses.

This author's own work-in-progress *SeqPup* takes this approach: it allows you to use the analysis software you need, whether *Clustal*, *CAP*, *tacg*, *fastDNAml*, or others, running on your own computer or on a server computer. SeqPup provides a graphical interface and standard user interface methods for editing sequences, basic manipulations and alignments, and sophisticated display and output options. It also links to analysis "engines" in a way you can configure to your taste. These analysis programs encode the complex data analysis algorithms but generally without a user-interface beyond command-line options. With a client program such as SeqPup, use of these programs is simplified and tied into a way for you to organize your sequence data.

Martin Senger at EMBL/EBI is working on a general CORBA interface to sequence analysis software called *AppLab* (http://industry.ebi.ac.uk/applab/), which is a similar approach.

Peter Rice's *EMBOSS* (http://www.sanger.ac.uk/Software/EMBOSS/) will be a freely distributable set of analysis programs, still under development at this writing. It will run on UNIX server computers, with a command-line interface. EMBOSS will include various sequence analysis topics, and will include a major effort to provide easy integration of other public domain packages. Analyses include rapid database searching with sequence patterns, and for sequence overlaps, simple and species-specific repeat identification, nucleotide sequence pattern analysis, codon usage analysis for small genomes, gene identification tools for genomic sequencing, rapid identification of sequence patterns in large scale sequence sets, protein motif identification, and presentation tools for publication.

These could be the basic analysis engines for sequence analysis, and some client program with a good user interface, perhaps like SeqPup, Java applets, or web forms, can be the program you use to run the analyses.

# 2. Free software highlights

## 2.1 Clustal sequence alignment

**Clustal** provides automatic multiple sequence alignment. The current version is called Clustal W, and is available for MacOS, Wintel, UNIX, and VMS computers. The simultaneous alignment of many nucleotide or amino acid sequences is now an essential tool in molecular biology. Multiple alignments are used to find diagnostic patterns to characterize protein families; to detect or demonstrate homology between new sequences and existing families of sequences; to help predict the secondary and tertiary structures of new sequences; to suggest oligonucleotide primers for PCR; as an essential prelude to molecular evolutionary analysis. The Clustal program fits this need very well. It is available at ftp://ftp-igbmc.u-strasbg.fr/pub/-ClustalW and at EBI and IUBio archives. There is a companion program ClustalX which provides a graphic interface to Clustal, and Clustal can be used from other sequence editors such as SeqPup.

## 2.2 Entrez to search genome data

The **Entrez** program is used for keyword searches of gene sequence data and Medline literature. It has been written by the programming staff at the National Center for Biotechnology Information (NCBI). It can be obtained from http or ftp://ncbi.nlm.nih.gov/. Entrez runs on a variety of computer systems. One big advantage of Entrez is the inclusion of a subset of MEDLINE, which covers the abstracts of entries submitted to the sequence databases. The WWW service of NCBI also offers an Entrez type of capability through your web browser. The program source for Entrez developed at NCBI has been instrumental in providing a software framework for other biosciences applications, including a version of SeqPup and Clustal-X.

## 2.3 NIH Image for image analysis

A very useful Macintosh program for general image analysis is **NIH Image**, written by Wayne Rasband. Image can be used to measure the area, average density, center of gravity, and angle of orientation of a user defined region of interest. It also performs automated particle analysis and can be used to measure path lengths and angles. Measurement results can be printed, exported to text files, or copied to the clipboard. Results can be calibrated to provide real world values. Find Image at ftp://zippy.nimh.nih.gov/pub/nih-image/, or http://rsb.info.nih.gov/nih-image/. There is now a version for MS Windows, from http://www.scioncorp.com/ .

## 2.4 PHYLIP for phylogeny analyses

The widely used Phylogeny Inference Package, **PHYLIP**, from Joseph Felsenstein, is a package of programs for inferring phylogenies (evolutionary trees), and written to work on as many different of computer systems as possible. It includes analyses of DNA and protein sequences, restriction sites, distance matrices and gene frequencies, quantitative and discrete characters, and plotting of evolutionary trees. Algorithms used include parsimony, maximum likelihood, neighbor joining and several others. Many

options for precise control of the analyses are available. The home for PHYLIP is at http or ftp://evolution.genetics.washington.edu/.

### 2.5 RasMol for molecular modeling

**RasMol** is a widely used, free molecular graphics program for the visualization of proteins, nucleic acids and small molecules. The program is aimed at display, teaching and generation of publication quality images. RasMol runs on all common computer systems. The program reads in molecule co-ordinates and interactively displays the molecule in a variety of color schemes and molecule representations, including depth-cued wireframes, space-filling spheres, ball and stick, solid and strand biomolecular ribbons, atom labels and dot surfaces. The home for RasMol is ftp://ftp.dcs.ed.ac.uk/pub/rasmol/

### 2.6 SeqPup for sequence editing

**SeqPup**, and its predecessor SeqApp, are biological sequence editor and analysis programs. They includes links to network services and external analysis programs. SeqPup is usable on common computer systems, using the new Java language.

Features include multiple sequence alignment and single sequence editing, read and write several sequence file formats, pretty print of alignments and sequences with boxed and shaded regions, sequence feature editing, manipulation and marking in prints, consensus, reverse-complement, distance/similarity, translate DNA to/from protein. Print file formats include PICT, Postscript and GIF.

User-definable links to external analysis programs, including ClustalW multiple alignment, CAP contig assembly, tacg restriction maps, fastDNAml phylogenetic analysis are included, and others can be added by you. One can use these running on your own computer or on an Internet server computer, using a new CORBA protocol (www.corba.org). Internet sequence analysis services include fetching sequences using SRS keyword search, and performing NCBI-BLAST similarity searches.

The home for SeqPup is http://iubio.bio.indiana.edu/soft/molbio/seqpup/. Note that this application is a work in progress; it has bugs. SeqApp is the Macintosh-only predecessor to SeqPup. Many folks find this currently a more useful program than SeqPup. It is faster, but lacks newer features of SeqPup.

## 3. Software use issues

### 3.1 Copyrighted versus public domain

Most free software is copyrighted by the author or sponsor, who retain all rights. They specifically grant you a right to use this software freely, perhaps only for non-commercial uses. Use of copyrighted software in commercial packages or other uses that make money require consent from the author. Many of the free software programs come with source code, so you can modify and extend it. This is a great boon to let sophisticated users do a needed analysis, but keep in mind use of such in a commercial product isn't allowed. If the author explicitly places his work in the public domain, s/he retains no control, and it can be used in commercial applications.

### 3.2 Citing software publications

Publicly available software is a publication, and free software that you use should be treated with consideration that you give other publications used in your research. While some free software has a companion paper publication to cite, some does not. It is usual practice to cite a software publication with its Internet URL in place of the journal/volume portion, e.g.,

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author at ftp://evolution.genetics.washington.edu/. Department of Genetics, University of Washington, Seattle.

Gilbert, D.G., 1996. SeqPup, biosequence editor & analysis platform, version 0.6. Bionet.Software, July 1996. <news://4rb7hr$6rc@usenet.ucs.indiana.edu> See also ftp://iubio.bio.indiana.edu /molbio/seqpup/

Some of these programs have been available at the same location for 10 years or more, so there isn't a general problem of impermanence with Internet locators.

# 4. Acknowledgments

The many developers of free software for biosciences, some of them mentioned herein, are the real authors of this document. If you use their software, please let them know you find it useful. Often one program builds upon others. This author would like to thank Jonathan Kans, Joseph Felsenstein, Michael Zuker, Gary Olsen, Dan Davison, Rob Harper, Dave Kristofferson, Reinhard Doelz, Rainer Fuchs, Peter Markiewicz, Thure Etzold, Xiaoqiu Huang, Des Higgins, Harry Mangalam, Jim Brown, Bill Pearson, and many others who provided their sweat and ideas to help make his own works useful. The many users of software who offer suggestions, criticisms and insights on how software should work, also contribute enormously to making free software better for all to use. The former GenBank home at Intelligenetics in the 1980s held an archive for free molecular biology software, to which many of us owe thanks for their pioneer efforts.

# 5. Software listings

This list of over 150 free software programs in molecular biology and related areas is not exhaustive by any means, but includes much of what is available for Macintosh and/or MS Windows computers.

Operating system key: **M** - MacOS, **W** - MS Windows or MS DOS, **O** - Other (Unix usually)

Software archive abbreviations:

**ebi** - ftp://ftp.ebi.ac.uk/pub/software/ or http://www.ebi.ac.uk/software/software.html

**iubio** - ftp://iubio.bio.indiana.edu/molbio/ or http://iubio.bio.indiana.edu/soft/molbio/

Alternate sites for the IUBio molecular biology collection:

ftp://ftp.funet.fi/pub/sci/molbio/iubiomolbio

ftp://ftp.sunet.se/pub/molbio

ftp://ftp.nig.ac.jp/pub/mirror/IUBIO/molbio

ftp://ftp.uam.es/pub/mirror/molbio,

ftp://ftp.pasteur.fr/pub/GenSoft/mirrors/IUBio/molbio

http, ftp://mic3.hensa.ac.uk/hosts/iubio.bio.indiana.edu/molbio/

ftp://bioinformatics.weizmann.ac.il/pub/software/mac and software/ibmpc

Some Email addresses and home URLs may be out of date. Unless otherwise indicated, all software listed is copyrighted by the author, and is available free for non-commercial use. Specific copyright restrictions should be noted. Some of this software is shareware, the author requesting a fee for use.

**ABaCUS M, W, O**

ABaCUS is a no-frills program to investigate the significance of the putative correspondence between exons and units of protein structure.

*Author:* Arlin Stoltzfus, arlin@is.dal.ca

*Archive:* iubio/evolve/abacus/

**ADE-4 M**

ADE-4 is a multivariate analysis and graphical display software package for Macintosh micro-computers.

*Author:* Olivier J.M. and others, Jean-Michel.Olivier@biomserv.univ-lyon1.fr

*Home:* ftp://biom3.univ-lyon1.fr/pub/mac/-ADE/ADE4, http://biomserv.univ-lyon1.fr/ADE-4.html

**Amplify M**

This Macintosh software is for use in designing, analyzing, and simulating experiments involving the polymerase chain reaction (PCR). Amplify will search a target sequence for near matches and display the results of using various primers. It can check oligos for matching sequence and for internal repeats.

*Author:* Bill Engels, WREngels@macc.wisc.edu

*Archive:* iubio/mac/amplify*, ebi/mac/

**AnalyzeSignalase M**

A Macintosh program for applying the algorithm of von Heijne to the prediction and analysis of mammalian signal sequences. It uses a weight-matrix method to try to predict the site at which signal peptides in secretory peptides are cut off by the signal peptidase.

*Author:* Ned Mantei

*Archive:* iubio/mac/analyze-signalase*

**Ancestor W**

Ancestor is designed to infer ancestral amino acid sequences from a set of homologous amino acid sequences whose phylogenetic relationships are known.

*Author:* Jianzhi Zhang, zhang@imeg.bio.psu.edu

*Archive:* iubio/ibmpc/ancestor*

**Annhyb W**

Annhyb is a little program for Windows 95 that is able to calculate various DNA sequence (degenerated or not) parameters.

*Author:* O. Friard & G. Stefanuto, friard@ba.cnr.it

*Home:* http://area.ba.cnr.it/~e105of01/-annhyb221.zip+

**Antheprot W, O**

ANalyze THE PROTeins (ANTHEPROT) is a package includes study of physico-chemical properties: hydrophobicity, antigenicity, flexibility, solvent accessibility, amphiphilicity; secondary structure prediction: Chou & Fasman, Garnier, Gibrat, Deleage, Levin; prediction of transmembranous regions & of structural domains; multiple alignment; search for biological sites using PROSITE, and PATMAT; search for homologous protein using FASTA; identity level between several sequences. Look and handle protein structures from PDB.

*Author:* G. Deleage, deleage@ibcp.fr & C. Geourjon, geourjon@ibcp.fr

*Home:* http, ftp://www.ibcp.fr/

*Archive:* iubio/ibmpc/antheprot*

**Automatic-BLAST M**

This is an AppleScript that automatically sends sequences by email to the BLAST server at ncbi.nlm.nih.gov at prescribed times, daily or weekly. The script uses the scriptable email program Eudora.

*Author:* Brian Osborne, bosborne@nature.berkeley.edu

*Home:* http://pgebaker4.pw.usda.–gov/bio/bio.html

*Archive:* iubio/mac/automatic-blast.*

**BCM Search Launcher M, W, O (Perl)**

The BCM Search Launcher is an integrated set of World- Wide Web pages that organize molecular biology- related search and analysis services available on the WWW by function, and provide a single point-of-entry for related searches. There is a batch client interface for Unix and Macintosh computers that allows multiple input sequences to be automatically searched as a background task, with the results returned as individual HTML documents. Requires Perl.

*Author:* Randall F. Smith et al.

*Home:* http://gc.bcm.tmc.edu:8088/search-launcher/launcher.html

**BufferStack M**

Given the appropriate information, the BufferStack will construct a complete recipe for a buffer that is defined in terms of both pH and ionic strength, at the temperature of use.

*Author:* Rob Beynon

*Archive:* iubio/mac/buffer*, ebi/mac/bufstack*

**Cabuffer W**

This program allows you to calculate the concentrations of all ionic species present in a mixture of up to four divalent cations and four ligands for these ions. Examples of such buffers are EDTA, EGTA, NTA, HEDTA, citrate, Ca-binding proteins, etc. Corrections for temperature, ionic strength and pH are provided.

*Author:* Jochen Kleinschmidt, kleinschmidt@mcclb0.med.nyu.edu

*Archive:* iubio/ibmpc/cabuf*

**CAIC M**

CAIC, Comparative Analysis by Independent Contrasts, computes phylogenetically independent contrasts from comparative data, allowing valid statistical testing of adaptational hypotheses. The manual gives guidance on testing hypotheses of correlated evolution among traits, rate variation among traits or taxa, and grade shifts.

*Author:* Andy Purvis & Andrew Rambaut, Andrew.Rambaut@zoology.ox.ac.uk

*Home:* http://evolve.zps.ox.ac.uk/-CAIC/CAIC.html

**CAP M, W, O**

Contig Assembly Program based on sensitive detection of fragment overlaps. C source, command-line program, can be used from SeqPup, others.

*Author:* Xiaoqiu Huang, huang@cs.mtu.edu

*Home:* ftp://cs.mtu.edu/pub/huang/

*Archive:* iubio/align/cap*

**CGR M**

A HyperCard stack for presenting nucleotide sequence data using Chaos Game Representation.

*Author:* Heikki Lehva, lehvaslaiho@cc.helsinki.fi

*Archive:*: iubio/mac/cgr*, ebi/mac/cgr*

**Clustal W M, W, O**

A multiple sequence alignment program, which is widely used and described in the article Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680(1994).

*Author:* D. Higgins et al.

*Home:* ftp://ftp-igbmc.u-strasbg.fr/pub/-Clustal*

*Archive:* ebi/mac/clustalw*, ebi/dos/-clustalw*, iubio/align/clustal*

**ClustalX M, W, O**

ClustalX is a graphic interface for the ClustalW alignment program. The sequence alignment is displayed in a window on the screen, with pull-down menus.

*Author:* Thompson J.D et al., julie@igbmc.u-strasbg.fr

*Home:* ftp://ftp-igbmc.u-strasbg.-fr/pub/ClustalX

**Codon Frequency Analyzer W**

This program helps to identify coding regions of DNA by comparing the codon frequencies in known coding regions with a sequence of DNA, the coding regions of which are unknown. The program will work with any organism.

*Author:* Ballyclaire Analysis

*Archive:* iubio/ibmpc/codon*

**CodonBiasIndex M**

The codon bias index is a statistic created by Bennetzen and Hall to quantify the extent to which more frequently used codons are used in preference to less frequently used codons.

*Author:* Tom Ritch, ritch@seas.ucla.edu

*Archive:* iubio/mac/codonbioasindex*

**Comap W**

A program for helping with the construction of restriction maps of small DNA fragment from digestion data. The program works under a graphical user interface.

*Author:* Kay Hofmann, khofmann@cipvax.-biolan.uni-koeln.de

*Archive:* iubio/ibmpc/codon*, ebi/dos/

**ConsInspector M, W, O**

ConsInspector uses a precompiled library of extended weight matrix descriptions (consensus profiles) of transcription factor binding sites to scan nucleic acid sequences for matches to these sites.

*Author:* K. Frech, et al. frech@gsf.de

*Home:* ftp://ariane.gsf.de/pub/

**Covariation M**

A Hypercard stack for phylogenetic comparative analysis of aligned RNA sequences.

*Author:* James W. Brown, jwbrown@mbio.ncsu.edu

*Archive:* iubio/mac/covariation*, ebi/mac

**CPrimer M**

CPrimer evaluates oligonucleotides as possible PCR primers. It shows melting points, interfering structures, and can search for optimum amplification pairs.

*Author:* Greg Bristol, gbristol@ucla.edu

*Archive:* iubio/mac/cprimer*

**DCSE W, O**

Dedicated Comparative Sequence Editor is a multiple alignment editor. It can be used to edit protein, DNA or RNA alignments. The structure of the molecules can be incorporated in the alignment. It offers lots of features such as color display of characters and structure, automatic alignment relative to sequences already aligned with others, sequence grouping, sequence or pattern searching, marker system, checking of incorporated RNA structure, on-line hypertext help, macros, and a lot more.

*Author:* Peter De Rijk, derijkp@reks.uia.ac.be

*Home:* http://www-rrna.uia.ac.be/~peter/dcse

*Archive:* ebi/dos/dcse

**Digest W**

Digest scans DNA sequence files for restriction sites. It prompts the user to specify which enzymes to cut with, and if they are in the enzyme database, it writes out the positions of all the cuts and sorts the fragments by size.

*Author:* Ramin Nakisa, ramin@ic.ac.uk

*Archive:* iubio/ibmpc/digest*, ebi/dos/digest*

**Digispeak M**

A program for reading sequencing gels with the aid of a Graf-Bar or similar sonic digitizer.

*Author:* Ned Mantei, bcmantei@-aeolus.vmsmail.ethz.ch

*Archive:* iubio/mac/digispeak*, ebi/mac/

**Dispan W**

DISPAN (genetic DIStance and Phylogenetic ANalysis) is designed to compute the following: average heterozygosity and standard error for each population; gene diversity and associated parameters; standard genetic distances and errors; DA distances between populations. It also

constructs phylogenetic trees and does bootstrap tests.

*Author:* Tatsuya Ota, imeg@psuvm.psu.edu

*Archive:* iubio/ibmpc/dispan*

**DNA Runs M**

DNA Runs is a program for performing a significance test of the number of runs in DNA sequence polymorphism and divergence data.

*Author:* John H. McDonald, mcdonald@udel.edu

*Home:* http://udel.edu/~mcdonald/

*Archive:* iubio/mac/dna-runs*

**DNA Slider M**

DNA Slider is a program for performing a significance test of heterogeneity in the ratio of polymorphic sites to fixed differences in DNA sequence data.

*Author:* John H. McDonald, mcdonald@udel.edu

*Home:* http://udel.edu/~mcdonald/

*Archive:* iubio/mac/dna-slider*

**DNA Stacks M**

DNA Stacks is a software package of HyperCard stacks providing utilities for viewing and manipulating molecular data. DNA Translator includes a gene mapping facility, draws and displays 2 linearized gene maps for comparison. Aligner is a stack for editing and display of multiple alignments. Codon Usage displays codon and amino acid usage data for a variety of organisms and organelles.

*Author:* D. J. Eernisse, DEernisse@fullerton.edu

*Home:* http://biology.fullerton.edu/-people/faculty/doug-eernisse

*Archive:* iubio/mac/dnastacks*, ebi/mac/

**Dna Workbench M, W, O (Perl)**

A program for sequence searching and manipulation. It offers powerful and fast searches on Genbank and other databases. Client-server access to remote databases and programs. Its many sequence manipulation functions include calculating the reverse complement, displaying reading frames and nucleotide to protein translations, editing, searching for restriction enzyme sites, searching for human repeat or vector in a sequence, comparing a sequence against a library or a user file, searching for a regular expression in a sequence. It requires Perl.

*Author:* James Tisdall, tisdall@-cbil.humgen.upenn.edu

*Home:* ftp://cbil.humgen.upenn.edu/-pub/dnaworkbench

**DNAdraw M**

DNAdraw is a program designed for preparing DNA and protein sequences for publication. A large selection of highlighting options is available. It has special features for formatting raw data into a style commonly used for publication, and for doing automatic highlighting of aligned sequences.

*Author:* Marvin Shapiro, mbs@kias.com

*Archive:* iubio/mac/dnadraw*

**DNAfrag W**

This program is used in restriction mapping of DNA or sizing of proteins from gels. It calculates the size of restriction fragments or peptide bands if standards are run on the same gel, and a standard curve of the standard bands using their mobilities.

*Author:* John Nash, Nash@biologysx.lan.nrc.ca

*Archive:* iubio/ibmpc/dnafrag*, ebi/dos/dfrag*

**DnaSP W**

DnaSP is a package for Windows that performs extensive population genetics analysis on DNA data, for hundreds of sequences of thousands of bases. It estimates several measures of polymorphism within and between populations, linkage disequilibrium, recombination, gene flow, gene conversion, and neutrality. DnaSP can do analyses by a sliding window method, and will make graphic representations.

*Author:* Julio Rozas & Ricardo Rozas, julio@porthos.bio.ub.es

*Home:* http://www.bio.ub.es/-~julio/DnaSP.html

*Archive:* ebi/dos/dnasp

**Dotplot W**

A dotplot program for MS-DOS.

*Author:* Ramin Nakisa, ramin@ic.ac.uk

*Archive:* iubio/ibmpc/dotplot*, dprel3*, ebi/dos/dotplot*

**Dotty Plotter M**

Dotty Plotter is a tool for drawing dot matrix comparisons of sequences in molecular biology. Dot plots are used to view all areas of homology between two nucleic acid or protein sequences.

*Author:* D. Gilbert, software@bio.indiana.edu
*Home:* iubio/mac/dottyplot*, ebi/mac/

**Double Digester M, O**

This is a program designed to help researchers in molecular biology assemble restriction maps of DNA using data from double digest experiments.

*Author:* L. Wright, wright-lawrence@yale.edu
*Archive:* iubio/restrict-enz/, ebi/mac/

**dPrimer M**

This Macintosh software is for use in calculating Tm values for degenerate primer.

*Author:* Haoyuan Chen, hchen@bimcore.emory.edu
*Archive:* iubio/mac/dprimer*

**EditView M**

DNA Sequence Viewer for ABI Sequencer trace data. EditView is a software application that allows you to view and print analyzed sample files containing sequence data from an ABI PRISM Genetic Analyzer.

*Author:* EditView@perkin-elmer.com
*Home:* ftp://ftp.abd.perkin-elmer.com/-pub/public/Sequencing/EditView/-EditView1.0.1.sea.hqx
*Archive:* iubio/mac/editview*

**Entrez M, W, O**

Entrez is a molecular sequence retrieval system developed at the National Center for Biotechnology Information (NCBI). Entrez provides an integrated approach for gaining access to nucleotide and protein sequence information, to the MEDLINE citations in which the sequences were published, and to a sequence-associated subset of MEDLINE.

*Author:* various at NCBI
*Home:* ftp://ncbi.nlm.nih.gov/entrez/

**Enzyme Kinetics M**

Enzyme Kinetics is a Hypercard stack for Macintosh computers. It calculates and plots the biochemical values for the kinetics of enzyme-catalyzed reactions.

*Author:* D. Gilbert, software@bio.indiana.edu
*Archive:* iubio/mac/enzymekinetic*, ebi/mac/enzymekin*

**Esee W**

Eyeball SEquence Editor, for MS DOS.

*Author:* Eric L. Cabot, cabot@gcg.com
*Archive:* iubio/ibmpc/esee*

**FastA M, W, O**

The FASTA sequence comparison programs, improved versions of the FASTP program, originally described in Science (Lipman and Pearson, (1985) Science 227:1435-1441)

*Author:* Bill Pearson, wrp@virginia.edu
*Home:* ftp://ftp.virginia.edu/pub/fasta/
*Archive:* iubio/search/fasta*

**fastDNAml M, W, O**

fastDNAml is a faster version of Joseph Felsenstein's DNAML (part of PHYLIP). Users should consult the documentation for DNAML before using this program.

*Author:* Gary J. Olsen et al., gary@phylo.life.uiuc.edu
*Archive:* iubio/evolve/fastdna*

**FoldIt M**

FoldIt (light) is a molecular modelling program to visualize and manipulate proteins. It has an integrated environment in which statistical analysis as well 3D observations can be realized on PDB files. It can analyze proteins up to 1600 residues. It can extract a number of structural features: Ramachandran plots, SS-bond plots, H-bond plots, and statistics on atomic parameters.

*Author:* Jean-Claude JESIOR, jean-claude.jesior@imag.fr
*Home:* ftp://ftp.imag.fr/pub/-TIMC/FoldIt.html
*Archive:* iubio/mac/foldit*, ebi/mac/

**GCUA M, O**

General Codon Usage Analysis. This program is designed to calculate various parameters that might be relevant in accessing the codon usage patterns of a group of genes. The user can look at codon usage (or any other statistic) in the dataset as a whole or for each gene individually. Features of this program include: Multivariate analyses of codon usage (RSCU) and amino acid patterns. Calculation of codon usage frequency, RSCU values, amino acid frequency data, base composition; distances between genes; ability to analyse complete prokaryotic genomes .

*Author:* James O. McInerney, J.mcinerney@nhm.ac.uk

*Home:* ftp://ftp.nhm.ac.uk/pub/gcua/

*Archive:* iubio/mac/gcua*

**Gel M, W**

An application to calculate the size of DNA fragment in an agarose gel.

*Author:* Jean-Michel Lacroix, lacroix@-medac.med.utoronto.ca

*Archive:* iubio/mac/gel-jml, iubio/ibmpc/gel-jml, ebi/mac/gel-jml, ebi/dos/gel-jml

**Gel W**

GEL takes a set of standard DNA fragment sizes and mobilities and predicts the sizes of unknown fragments, using a least squares fit to the relationship of mobility and fragment length.

*Author:* John R. Thompson

*Archive:* iubio/ibmpc/gel*, ebi/dos/gel/

**Gel Frag Sizer M**

Gel Frag Sizer is a HyperCard stack which calculates restriction fragment sizes from their mobilities. Two methods for estimating sizes are provided: the local reciprocal method of Elder & Southern or the cubic spline method.

*Author:* D. Gilbert, software@bio.indiana.edu

*Home:* iubio/mac/gelfragsizer.*

*Archive:* ebi/mac/

**Gel Manager W**

Gel Manager is a user friendly program which runs in MS Windows. It includes techniques of image processing along with options for data analysis. It can deal with different kinds of data such as: RFLP, RAMM, RAPD, microsatellites and other fingerprinting techniques. It is useful for studies including genetic relationships, taxonomy and classification , epidemiology, etc.

*Author:* Carlos Vaquerizo, Joaquin Dopazo, dopazo@samba.cnb.uam.es

*Home:* ftp://ftp.cnb.uam.es/software/-molbiol/gel_man

**GeneDoc W**

GeneDoc is a full featured multiple sequence alignment editor and shading utility. It is intended to help you bring your genetics research work to publication with shading, page and font layout features.

*Author:* Karl Nicholas, ketchup@cris.com

*Home:* http://www.cris.com/~ketchup/-genedoc.shtml

*Archive:* iubio/ibmpc/genedoc*

**GeneMaster W**

A small gene analysis package that performs searches for sequences, looks for regions of GC richness, translates using a variety of start codons and genetic codes, and restriction analysis.

*Author:* Shawn Abigail, ad873@freenet.carleton.ca

*Archive:* iubio/ibmpc/genemast*

**GeneTree M, W**

GeneTree is a program for comparing gene and species trees using reconciled trees. The program can compute the cost of embedding a gene tree within a species tree, visually display the location and number of gene duplications and losses, and search for optimal species trees.

*Author:* Roderic D. M. Page, r.page@bio.gla.ac.uk

*Home:* http://taxonomy.zoology.-gla.ac.uk/rod/genetree/

**Gepasi W**

Gepasi is intended for the simulation of the kinetics of systems of chemical and biochemical reactions. Gepasi is able to simulate the steady-state and time course behaviour of reactions in several compartments of different volumes. Results can be plotted in 2D and 3D graphs directly from the program. Steady states are analysed with metabolic control analysis and linear stability analysis.

*Author:* Pedro Mendes, prm@aber.ac.uk

*Home:* http://gepasi.dbs.aber.ac.uk/-softw/gepasi.html

*Archive:* ebi/dos/

**HDProbe M**

HDProbe matches a probe sequence against a set of alleles and tabulates stable and unstable mismatches. HDProbe accepts probe and allele sequences as input, then displays the sequences with respect to their orientation in a heteroduplex molecule.

*Author:* Marvin Shapiro, mbs@pa.net

*Home:* iubio/mac/hdprobe.*

**HelixVu W**

HelixVu illustrates an 80 bp region of DNA as a helix with the sequence listing printed above the

helix diagram. This view is useful to see the spatial relationship between DNA modifications.

*Author:* Richard Seyler

*Archive:* iubio/ibmpc/helixvu*

**Hyper W**

Hyper is a program for the analysis of enzyme kinetic data under MS Windows. Enzyme kinetic data are subjected to non-linear regression and the results displayed in five standard graphical forms and printed.

*Author:* J S Eastery, jse@liverpool.ac.uk

*Archive:* iubio/ibmpc/hyper*

**HyperPCR M**

A HyperCard stack that calculates the optimal annealing temp. of a PCR reaction according to the algorithm of Rychlik.

*Author:* Brian Osborne, bosborne@violet.berkeley.edu

*Archive:* iubio/mac/hyperpcr*, ebi/mac/

**Intron Analyzer W**

There are basic differences in the base composition of introns from animals and plants, and this program will examine introns to find regularities. From a given list of introns you can explore by aligning them either at the 5' or at the 3' end, or study the adjoining exon-parts. The program will build a consensus-sequence which shows the most frequent base in each position, including a graphic plot.

*Author:* Michael Liss, LISS@alf1.ngate.uni-regensburg.de

*Archive:* iubio/ibmpc/intron-analyzer*, ebi/dos/intana*

**Lalnview M, W, O**

LalnView is a graphical program for visualizing local alignments between two sequences. Sequences are represented by colored rectangles to give an overall picture of their similarities. It is able to display sequence features (active site, domain, motif, propeptide, etc.) along the alignment. LalnView is a useful tool for analysing pairwise alignments and for making the link between sequence homology and what is known about its structure or function.

*Author:* Laurent Duret, duret@dim.hcuge.ch

*Home:* ftp://expasy.hcuge.ch/pub/lalnview

**Lines&Kinetics M**

A graphic way to calculate linear regressions with normal or logarithmic data, the doubling time of a microbial culture, and the kinetic parameters for an enzyme reaction.

*Author:* Manuel G. Claros, claros@uma.es

*Home:* http://www.ie.embnet.org/-embnet.news/vol5_1/kinetics.html

*Archive:* iubio/mac/lines-kinetics*

**Linkage-1 M**

designed to aid the geneticist in the detection and estimation of linkage in segregating progenies.

*Author:* Karl A. Suiter, ksuiter@acpub.duke.edu

*Home:* iubio/mac/linkage1*

**Lintr W, O**

These programs are for testing the molecular clock on a given topology of a phylogenetic tree and making linearized trees, using nucleotide or amino acid sequences.

*Author:* Naoko Takezaki, ntakezak@lab.nig.ac.jp

*Archive:* iubio/evolve/lintr/

**LoopDLoop M, W, O (Java)**

loopDloop is a tool for drawing and editing RNA secondary structures in molecular biology. A MacOS specific and Java version are available. Mulfold will generate RNA foldings for display by loopDloop. A related program, LoopViewer, lacks the editing features but is simpler to use.

*Author:* D. Gilbert, software@bio.indiana.edu

*Home:* iubio/loopdloop/

*Archive:* ebi/mac/loop*

**MACAW M, W**

Multiple Alignment Construction & Analysis Workbench (MACAW) is a program for locating, analyzing, and editing blocks of localized sequence similarity among multiple sequences and linking them into a multiple alignment. It includes sequence alignment search, editing and display. It is a very nice program according to many, allowing one to look for blocks of homology in sequences.

*Author:* Greg Schuler & Stephen Altschul, schuler@ncbi.nlm.nih.gov

*Home:* ftp://ncbi.nlm.nih.gov/pub/macaw/

*Archive:* iubio/ncbi/macaw/, ebi/dos/macaw*

**MacBoxshade M, O**

A program for creating good-looking printouts from multiple aligned protein or DNA sequences. The program does no alignment by itself, it uses files from a multiple alignment program. Output can be PostScript, EPSF, PICT, RTF or ASCII text. Identical and similar residues in the multiple-alignment are represented by different colors or shadings. There are many options of shading, sequence numbering, consensus output and so on.

*Author:* Michael D. Baron, michael.baron@bbsrc.ac.uk (macos), Kay Hofmann (original)

*Home:* ftp://ulrec3.unil.ch/pub/-boxshade/macboxshade

*Archive:* iubio/mac/macboxshade*

**MacPattern M**

MacPattern is a Macintosh application for protein pattern searches (using PROSITE) and block profile searches (using BLOCKS). MacPattern assists in finding putative functions for new protein sequences by supporting pattern searches using the PROSITE database, block searches using the BLOCKS database, and statistical analyses (maximal segment score analysis and Eguchi-Seto method).

*Author:* Rainer Fuchs, rainer_fuchs@glaxo.com

*Archive:* iubio/mac/macpattern*, ebi/mac

**MacPlasmap M**

If your study or research involves preparation of circular plasmid maps, you will find MacPlasmap an indispensable tool to have. It draws, stores, and prints high-quality circular plasmid maps with the data you specify.

*Author:* Jingdong Liu

*Archive:* iubio/mac/macplasmap*, ebi/mac/

**MacProt M**

MacProt is a of a set of programs for analyzing protein sequences for secondary structure, chain flexibility, hydropathy, helical wheels, and so on.

*Author:* Peter Markiewicz

*Archive:* iubio/mac/plota/, ebi/mac/plota_*

**MacStripe M**

MacStripe is a program for the prediction and analysis of potential coiled-coil regions in protein sequences. MacStripe is the ideal tool for anyone who wants to explore potential alpha-helical coiled coils in the sequence of their protein. With a full Macintosh interface, the results of analyses (raw data or publication quality plots) can easily be exported to other software. MacStripe uses the algorithm of Andrei Lupas's COILS2 for detailed and reliable coiled-coil predictions.

*Author:* Alex Knight, aek4@york.ac.uk

*Home:* http://www.york.ac.uk/-depts/biol/units/coils/coilcoil.html

*Archive:* iubio/mac/macstripe*

**MacT M**

MacT is a set of programs for the Macintosh to construct and evaluate unrooted trees derived from amino acid sequences using a distance matrix method.

*Author:* Angela Luettke & Rainer Fuchs

*Archive:* ebi/mac/mact_*, iubio/mac/mact.*

**Map Manager M, W**

Map Manager is a program which helps analyze the results of genetic mapping experiments using intercrosses with codominant markers, backcrosses or recombinant inbred strains in experimental plants or animals. It is a specialized database program which allows easy storage, retrieval, and display of information from such mapping experiments, and it also has tools for searching and for statistical analysis of the experimental results. These tools assist the user in determining linkage among loci and in determining the order of loci.

*Author:* Kenneth F. Manly, kmanly@-mcbio.med.buffalo.edu

*Home:* http,ftp://mcbio.med.buffalo.edu/

*Archive:* iubio/mac/map-manager*, ebi/mac/-mapmanager*

**MapMaker M, W, O**

MapMaker is a linkage analysis package designed to help construct primary linkage maps of markers segregating in experimental crosses. One version performs full multipoint linkage analysis for dominant, recessive, and co- dominant (e.g. RFLP-like) markers.

*Author:* Whitehead Inst. for Biomedi. Res., mapmaker@genome.wi.mit.edu

*Home:* ftp://genome.wi.mit.edu/-distribution/mapmaker3

*Archive:* iubio/mapmaker/

**Materials & Methods M**

A Hypercard stack for the storage & retrieval of laboratory procedures. The stack comes pre-loaded

with many commonly-used procedures used in molecular biology.

*Author:* James W. Brown, jwbrown@mbio.ncsu.edu

*Archive:* iubio/mac/mandm*, ebi/mac/-matmeth*

**Matilda W**

Matilda is a specialized DNA database management system that helps scientists extract the high-level information that they need for recombinant DNA experiments from a large sequence and genetic information database. It incorporates functional data and restriction map data. Sequence and functional data are extracted from sequence files and from additional information. As recombinant DNA clones are constructed, their descriptions are added to the database so that they can be used to describe clones constructed later.

*Author:* Isralewitz, B. & Shalloway, D.

*Archive:* iubio/ibmpc/matilda*

**MatInspector M, W, O**

MatInd is a simple but powerful method to derive a matrix description of a consensus from a number of short sequences on which the definition of an IUPAC code would be based. MatInspector is a program that uses a large library of predefined matrix descriptions of transcription factor binding sites to locate matches in nucleotide sequences of unlimited length. It assigns a quality rating to matches and thus allows a quality- based filtering and selection of matches.

*Author:* K.Quandt, et al, quandt@gsf.de

*Home:* ftp://ariane.gsf.de/pub/

*Archive:* ebi/mac/matind*, ebi/dos/matind*

**Memsat W, O**

MEMbrane protein Structure And Topology.

*Author:* David T. Jones, jones@bsm.bioc.ucl.ac.uk

*Home:* ftp://ftp.biochem.ucl.ac.uk/-pub/MEMSAT

**Metree W**

A package for inferring and testing minimum evolution trees. This package is intended to find the minimum evolution tree that has the smallest value of the sum of branch lengths for a set of sequences, identify a set of trees that are not significantly different from the ME tree, and print the trees in a publishable form.

*Author:* Andrey Rzhetsky and Masatoshi Nei, aur1@psuvm.psu.edu

*Archive:* iubio/ibmpc/metree*

**MitoProt M, O**

It supplies a series of parameters that permit theoretical evaluation on mitochondrial targeting sequences and the importability. MitoProt II provides the possibility to predict mitochondrial proteins harboring targeting sequences. Chloroplast proteins also can be studied.

*Author:* Manuel G. Claros, claros@uma.es, Pierre Vincens

*Home:* ftp://ftp.ens.fr/pub/molbio/, ftp://ftp.rediris.es/software/incoming/science/

*Archive:* iubio/mac/mitprot*, ebi/mac/

**Molwt W**

This program calculates molecular weight from an entered chemical formula, and gives concentrations in various units in response to an entered formula and concentration.

*Author:* John A. Kiernan, jkiernan@julian.uwo.ca

*Archive:* iubio/ibmpc/molwt*, ebi/dos/-molwt*

**Mufold (Mac) M**

A MacOS port of Michael Zucker's MFold software for prediction of RNA secondary structure by free energy minimization, including sub optimal folding with temperature dependence. See also PCFold.

*Author:* D. Gilbert (mac port), M. Zuker (MFold)

*Home:* iubio/mac/mulfold*, ftp://snark.wustl.edu/pub/ (MFold)

*Archive:* ebi/mac/

**NIH Image M**

Image can be used to measure the area, average density, center of gravity, and angle of orientation of a user defined region of interest. It also performs automated particle analysis and can be used to measure path lengths and angles.

*Author:* Wayne Rasband

*Home:* ftp://zippy.nimh.nih.gov/pub/nih-image/, http://rsb.info.nih.gov/nih-image/

**NJbafd W, O**

These programs are for constructing a neighbor-joining or UPGMA tree from allele frequencies of microsatellite DNA or other genetic markers, and computing heterozygosities and Gst.

*Author:* Naoko Takezaki, ntakezak@lab.nig.ac.jp

*Archive:* iubio/evolve/njbafd/

**NJPlot M, W, O**

NJPlot is a phylogenetic tree drawing program that handles files describing trees by the nested parentheses method (e.g. PHYLIP-built trees). Features: A graphical interface allows to re-root a tree anywhere and to swap branches. Bootstrap values are displayed next to internal branches. Branch lengths can be optionally displayed. Tree plots can be saved to a PostScript or PICT file.

*Author:* Manolo Gouy, mgouy@biomserv.univ-lyon1.fr

*Home:* ftp://biom3.univ-lyon1.fr/pub/mol_phylogeny/njplot

*archive:* ebi/mac/

**Noncode W**

A program that will read an ESEE file containing nucleic acid sequences and produce a distance matrix using the Kimura 2-parameter model.

*Author:* Eric L. Cabot, cabot@gcg.com

*Archive:* iubio/ibmpc/noncode*

**Numclone W**

Estimates the number of clones one has to screen from a genomic library in order to find a desired clone.

*Author:* John Nash, Nash@biologysx.lan.nrc.ca

*Archive:* iubio/ibmpc/numclone*

**Oligobase W**

This is a shareware program for Windows designed to organize and catalog oligonucleotides collection of a biological laboratory. It stores information about oligonucleotides, select subsets according to specified criteria, print out order forms, calculate molecular weight and melting temperature, and manipulate oligonucleotides.

*Author:* Igor Sidorenkov, sidorenk@rocketmail.com

*Home:* http://lochfort.com/oligobase

*Archive:* iubio/ibmpc/obase*

**OligoCR M**

OligoCR is a data management tool for organizing and cataloging oligonucleotide collection. It allows you to store information about the oligos, e.g., the category (PCR, sequencing), type of project the oligo is used for, its application, its description. To search your oligo database, just click the mouse. It has the capability to proofread your sequences.

*Author:* Yongming Sun, ysun@-hdklab.wustl.edu

*Home:* http://hdklab.wustl.edu/~ysun

*Archive:* iubio/mac/oligocr*

**OligoMutantMaker W**

OligoMutantMaker simplifies the designing and screening of oligonucleotide-directed single amino acid substitution experiments by searching for nucleotide sequences which introduce a restriction endonuclease recognition sequence into the codon substitution site of the mutant.

*Author:* Kevin Beadles et al.

*Archive:* iubio/ibmpc/oligo*, ebi/dos/oligo*

**Onix W**

A program for MS Windows, Onix allows users to examine proteins with known 3D structure from PDB. This program was designed for structure investigation of ligand binding site in proteins. ONIX is interactive software with a high-performance interface, fast 3D molecular graphics and analysis of water-accessible surface.

*Author:* A.S.Ivanov et al., ivanov@ibmh.msk.su

*Home:* ftp://org.chem.msu.su/pub/-software/Onix/

**P1 Clones M**

This is a simple HyperCard stack for keeping track of clones that were constructed with the bacteriophage P1 cloning system.

*Author:* Ken Abremski, sabremske@esvax.dnet.dupont.com

*Archive:* iubio/mac/p1clones*

**PAML M, W, O**

Phylogentic Analysis by Maximum Likelihood contains three main programs for model fitting and phylogenetic tree reconstruction using nucleotide or amino-acid sequence data.

*Author:* Ziheng Yang, z.yang@ucl.ac.uk

*Home:* ftp://abacus.gene.ucl.ac.uk/pub/paml

*Archive:* iubio/evolve/paml*

**PcFold _____ W**

A PC version of Michael Zuker's RNA-folding program which uses an energy minimization algorithm to predict stem and loop regions of RNA structures. See also MulFold, and home site for Unix versions.

*Author:* Michael Zuker et al., zuker@snark.wustl.edu

*Home:* ftp://snark.wustl.edu/pub/

*Archive:* iubio/ibmpc/pcfold*, ebi/dos/pcfold

**PHYLIP _____ M, W, O**

A PHYLogeny Inference Package of many programs for phylogenetic analysis, including parsimony, compatibility, distance matrix invariants ("evolutionary parsimony") and likelihood methods on various kinds of data.

*Author:* Joseph Felsenstein, joe@genetics.washington.edu

*Home:* http, ftp://evolution.genetics.-washington.edu/

*Archive:* iubio/evolve/phylip*, ftp://ftp.nig.ac.jp/pub/unix/phylip, ftp://ftp.bioss.sari.ac.uk/pub/-phylogeny/phylip

**Phylodendron _____ M, W, O (Java)**

Phylodendron is an application for drawing phylogenetic trees It reads data in New Hampshire (Newick) format. Options allow you to adorn and edit the tree.

*Author:* D. Gilbert

*Home:* iubio/java/apps/trees/

**Phyltest _____ W**

A program for testing phylogenetic hypothesis, with comparison of three alternative phylogenetic trees, estimation of average pairwise distances, and others.

*Author:* Sudhir Kumar, imeg@psuvm.psu.edu

*Archive:* iubio/ibmpc/phyltest*

**Plasmid Processor _____ W**

Plasmid Processor is a simple tool for plasmid presentation for scientific and educational purposes. It features both circular and linear DNA, user defined restriction sites, genes and multiple cloning site. In addition you can manipulate plasmid by inserting and deleting fragments. Created drawings can be copied to clipboard or saved to disk for later use. Printing from within program is also supported.

*Author:* T. Kivirauma, P. Oikari and J.Saarela, Dept. of Biochemistry & Biotechnology, U. of Kuopio, plasmid@uku.fi

*Home:* http://www.uku.fi/~kiviraum/-plasmid/plasmid.html

*Archive:* iubio/ibmpc/plasmid-processor*, ebi/dos/plasmid

**Plasmid-Maker _____ M**

Draws linear and circular plasmid maps, allows borders of various widths and fills of grays, arrows, and other options.

*Home:* http://yeamob.pci.chemie.uni-tuebingen.de/Archiv/-PlasmidMaker.html

*Archive:* iubio/mac/plasmid-maker*

*Author:* Kai-Uwe Froehlich, kaifr@uni-tuebingen.de

**Primer Design _____ W**

PrimerDesign is a DOS program to choose primer for PCR or oligonucleotide probes. It is tailored to check known sequences for repeats and unique sequences and subsequently to create primers according to this data. A lot of constrains are available to meet your conditions. It can handle up to 31500 base pairs. Additional features: unique sequences, repeats, restriction sites.

*Author:* Andreas Becker, Joerg Napiwotzki, becker@ps1515.chemie.uni-marburg.de

*Home:* ftp://ftp.chemie.uni-marburg.de/-pub/PrimerDesign

**Primer _____ M, W, O**

Primer is a computer program for automatically selecting PCR primers. It tests oligos for annealing temperature, complementarity to genomic repeat sequences, ability to form primer-dimer, and other criteria. Primer annealing temperature calculation is based on thermodynamic parameters.

*Author:* Steve Lincoln et al., primer@genome.wi.edu

*Home:* ftp://genome.wi.mit.edu/pub/-software/Primer2.2

*Archive:* iubio/primer/primer-wi*

**Primer-Master _____ W**

Automatically search and selection of optimal primers and primers pairs for various variants of

PCR; analysis of oligs supposed to be used as PCR primers or hybridization probes; editor for comfortable typing-in new nucleotide sequences;

*Author:* Proutski Vitali, Sokur Oleg, proutski@influenza.spb.su

*Archive:* ebi/dos

**Primers! M**

Primers! is a primer design shareware application, written by the author of Whitehead Institute Primer2. It allows users to interactively scroll through lists of forward and reverse primers to pick exactly the primer pair wanted.

*Author:* Richard Resnick, rjr@applepi.com

*Home:* http://www.applepi.com/

*Archive:* iubio/mac/primers*

**ProAnal W**

ProAnal is for analysis of multiple protein alignments, studying the structure-function and structure-activity relationships in protein/peptide families. The program uses aligned amino acid sequences with data of their activity and searches for correlations between data on activity and various physico-chemical characteristics of different regions in primary structures.

*Author:* Alexey Eroshkin, eroshkin@vector.nsk.su

*Archive:* iubio/ibmpc/proanal*, ebi/dos/

**ProAnalyst W**

ProAnalyst is for investigation of structural differences between proteins divided by functional, evolutionary or other criteria; structure-activity relationships investigation; searching motifs; protein-engineering experiments; and many other protein analysis functions.

*Author:* Vladimir Ivanisenko, Alexey Eroshkin, eroshkin@vector.nsk.su

*Archive:* iubio/ibmpc/panalyst*, ebi/dos/-proanalyst

**ProAnWin W**

Multiple sequence alignment, analysis of protein sequences and structures, structure-activity relationships, design of protein-engineering experiments. Threads multiple alignment onto known 3-dimensional structure; searches linear, spatial sites, conservative and variable in changes of specified physico-chemical properties; plots of different physico-chemical profiles for individual or a set of protein sequences; and many other functions.

*Author:* I.Pika et al., eroshkin@vector.nsk.su

*Archive:* iubio/ibmpc/paw*, ebi/dos/proanwin

**ProfileGraph W**

A graphical protein analysis tool.

*Author:* Kay Oliver Hofmann, khofmann@biomed.biolan.uni-koeln.de

*Archive:* iubio/ibmpc/prograph*, ebi/dos/pgraph*

**PromFind W**

PromFind is a DOS program that accepts a DNA sequence, and adds a feature table to annotate the location of putative promoter regions.

*Author:* Gordon B. Hutchinson, hutch@netshop.bc.ca

*Archive:* iubio/ibmpc/promfind*

**ProMSED W**

ProMSED, a Windows application for both automatic and manual DNA and protein sequence alignment, editing, comparison and analysis. Automatic alignment is based on ClustulV; manual alignment and visual analysis are facilitated by group and block operations and amino acid coloring reflecting their similarity.

*Author:* Anatoly Frolov, Alexey Eroshkin, eroshkin@vector.nsk.su

*Archive:* iubio/ibmpc/promsed*, ebi/dos/promsed/

**Prophet W, O**

Prophet offers advanced, easy-to-use software tools for data management and visualization, and statistical analysis - from simple descriptive statistics to multi-factor ANOVA, logistic regression, and non-linear modeling. It also offers tools for analyzing biological sequences, including multiple sequence alignment, translation, restriction enzyme and proteolytic cleavage analyses, PCR primer design, BLAST searches, remote database retrievals, and more.

*Author:* Prophet software group, BBN, prophet-info@bbn.com

*Home:* http://www-prophet.bbn.com/

**Protein Sequence Analysis W**

The program is a sequence editor with the capability to amino acid composition, hydrodynamic calculation, mass for various isotope labeling, isoelectric point, UV spectrum,

relative hydrophobicity, secondary structure prediction, and others.

*Archive:* iubio/ibmpc/prot-sa*

**Puzzle M, W, O**

Puzzle is a maximum likelihood analysis for nucleotide, amino acid, and two-state data. It reconstructs phylogenetic trees from molecular sequence data, and has a fast tree search that allows analysis of large data sets. Puzzle is PHYLIP compatible.

*Author:* Korbinian Strimmer, Arndt von Haeseler, strimmer@zi.biologie.uni-muenchen.de

*Home:* ftp://fx.zi.biologie.uni-muenchen.de/pub/puzzle

*Archive:* iubio/evolve/puzzle/, ebi/mac/puzzle, dos/puzzle, unix/puzzle

**RAMHA W**

Monte Carlo simulation of the random mutagenesis of synthetic cDNAs.

*Author:* David P. Siderovski, Siderovski@Galen.-OCI.UToronto.CA

*Archive:* iubio/ibmpc/ramha*

**RasMol M, W, O**

RasMol is a molecular modelling program for the visualization of proteins and nucleic acids. It reads protein databank (PDB) files and interactively renders them in a variety of formats, including wire, stick, stick_and_ball, CPK, and ribbon.

*Author:* R. Sayle, ros@dcs.ed.ac.uk

*Home:* ftp://ftp.dcs.ed.ac.uk/pub/rasmol/

*Archive:* ftp://kekule.osc.edu/pub/-chemistry/software/X-WINDOWS/rasmol*, ebi/mac/rasmol*, software/dos/raswin*,

**Rbinding W**

Calculates the number of binding sites and the affinities of cell surface receptors for ligands (Scatchard analysis).

*Author:* Nico van Belzen & Joop van Zoelen, belzen@pa1.fgg.eur.nl

*Archive:* iubio/ibmpc/rbindin*

**ReadSeq M, W, O**

A program for converting among several biosequence file formats.

*Author:* Don Gilbert, software@bio.indiana.edu

*Home:* iubio/readseq/

*Archive:* ebi/mac/readseq*

**ReAlign W**

A program that realigns a DNA alignment according to a peptide alignment, thereby improving the alignment in areas not too well conserved.

*Author:* Rasmus Wernersson, RWer@novo.dk

*Archive:* iubio/ibmpc/realign*

**RepFind W**

RepFind (promoter find) is a MS DOS program to identify common repetitive elements in DNA sequence. It is also able to identify and mask vector sequence.

*Author:* Gordon B. Hutchinson, hutch@netshop.bc.ca

*Archive:* iubio/ibmpc/repfind*, ebi/dos/repfind*

**Restdata W**

Restriction data and phylogenetic analysis, computes the numbers of nucleotide substitutions per site for pairs of DNA sequences; constructs phylogenetic trees by using the neighbor-joining method.

*Author:* Tatsuya Ota, imeg@psuvm.psu.edu

*Archive:* iubio/ibmpc/restdata*

**RestSite W**

Several programs for analyzing restriction site or fragment data for use in molecular systematics studies.

*Author:* Joyce C. Miller

*Archive:* iubio/ibmpc/restsite*

**RNA Dotplot M**

RNA Dotplot is a simple utility to print a dot matrix of the potential base pairing interactions in an RNA sequence.

*Author:* David S. McPheeters, mcpheeters@biochemistry.cwru.edu

*Archive:* iubio/mac/rna-dotplot*

**Rnadraw W**

Rnadraw offers RNA optimal structure / basepair-probability matrix / heat curve calculation on Intel computers, providing a consistent user interface

with many possibilities to view, print, import/export and edit calculation results.

*Author:* Ole Matzura, ole@mango.mef.ki.se

*Home:* ftp://broccoli.mfn.ki.se/pub/rnadraw

**RNA_d2 W**

RNA_d2 is a user-friendly program developed for interactively generating aesthetic and non-overlapping drawings of RNA secondary structures. It allows easy untangling and editing of RNA molecules > 1000 nucleotides long.

*Author:* J. Perochon-Dorisse et al., rnad2@ibcg.biotoul.fr

*Home:* ftp://hpsrv.biotoul.fr/rna

**Sagittarius_Dna W**

A package for exon/intron structure revealing on the base of protein k-tuples statistic.

*Author:* Victor B. Strelets, strelets@bio.indiana.edu

*Archive:* iubio/ibmpc/sag-exo*, ebi/dos/sag-exo

**Sagittarius_Pir W**

A highly compact databank variant of original PIR database designed to assist individuals in utilization of sequence database information without huge storage space requests. Includes fast homology searches and selection of sequences by fields (name, source, keyword, etc.), or (non)perfect homology with user-defined short sequence.

*Author:* Victor B. Strelets, strelets@bio.indiana.edu

*Archive:* iubio/ibmpc/sag-pir*

**Sagittarius_Seqanalref W**

A dialog shell for storage and manipulation of reference information. This particular variant is oriented on SEQANALREF databases compiled by A.Bairoch.

*Author:* Victor B. Strelets, strelets@bio.indiana.edu

*Archive:* iubio/ibmpc/seqanalr*, ebi/dos/sag-sar*

**Send W**

A program for computing the standard errors of nucleotide diversity and divergence using the algorithm of Nei and Jin.

*Author:* Li Jin

*Archive:* iubio/ibmpc/send*

**Sendbs W, O**

A program that computes average nucleotide substitutions within and between populations using the algorithm of Nei and Jin. It computes standard errors with a bootstrap method that differs from Nei and Jin's.

*Author:* Naoko Takezaki, ntakezak@lab.nig.ac.jp

*Archive:* iubio/evolve/sendbs*

**Seq-Eudora-Blast M**

Macintosh AppleScript applications that automate your BLAST searches. Drag-and-dropped sequence files to the BLAST server at ncbi.nlm.nih.gov using Eudora mail.

*Author:* Brian Osborne, bosborne@nature.berkeley.edu

*Home:* http://pgebaker4.pw.usda.-gov/bio/bio.html

*Archive:*: iubio/mac/seq-eudora-blast*

**Seqaid II W**

Seqaid II is a MS-DOS program for DNA and protein sequence analysis. Functions include editing, modified Needleman-Wunsch alignment, dot matrix comparison, fragment sizer, base composition, translations, protein structure and hydropathicity, restriction site search, and locating potential exons by codon bias.

*Author:* Donald Roufa & D.D. Rhoads

*Archive:* iubio/ibmpc/sequaid*, ebi/dos/sqaid*

**SeqApp M**

A Macintosh biosequence editor, analyzer, and network handyman. See SeqPup.

*Author:* D. Gilbert, seqapp@bio.indiana.edu

*Home:* iubio/seqapp/

**SeqPup M, W, O (Java)**

The successor to SeqApp, SeqPup is a biological sequence editor and analysis program. It includes links to network services and external analysis programs. Features include multiple sequence alignment and editing, support for several file formats, sequence feature editing, manipulation and marking, translate dna/protein, consensus, reverse/complement, and distance methods, pretty print of alignments and sequences with boxed and shaded regions, Internet searches, use of external analysis programs, including ClustalW multiple alignment, CAP contig assembly, tacg restriction map, and a remote client-server interface. The current version runs on any os supporting Java, an

older version runs on MacOS, MSWin and some Unix.

*Author:* D. Gilbert, seqpup@bio.indiana.edu

*Home:* iubio/seqpup/

**SeqSimPresenter _____ M**

SeqSimPresenter converts a set of aligned sequences to shaded bars of which correspond to the degree of similarity. It presents large alignments in a compact form and allows a fast recognition of the amount, extension, and distribution of conserved regions.

*Author:* cbkfr01@mailserv.zdv.uni-tuebingen.de

*Archive:* iubio/mac/seqsimpresent*, ebi/mac/

**Sequin _____ M, W, O**

Sequin is a stand-alone software tool developed by the NCBI for submitting entries to the GenBank, EMBL, or DDBJ sequence databases. It is capable of handling simple submissions which contain a single short mRNA sequence, and complex submissions containing long sequences, multiple annotations, segmented sets of DNA, or phylogenetic and population studies.

*Author:* Jonathan Kans, Colombe Chappey, info@ncbi.nlm.nih.gov

*Home:* ftp://ncbi.nlm.nih.gov/sequin/, http://www.ncbi.nlm.nih.gov/Sequin

**SeqVu _____ M**

An alignment editor with analysis options that allows you to work quickly and simply with multiple sequences. It is ideal for manually correcting alignments produced using software such as ClustalV.

*Author:* James Gardner, j.gardner@garvan.unsw.edu.au

*Home:* ftp://gimr.garvan.unsw.edu.au/pub/

*Archive:* iubio/mac/seqvu*

**Shm _____ W, O**

Shm was designed to provide assistance in the analysis of somatic (point) mutations induced in the immunoglobulin genes of B-lymphocytes. It builds clonal trees by parsimony and displays them along with somatic mutations.

*Author:* Laurentiu COCEA, cocea@necker.fr

*Archive:* iubio/ibmpc/shm*

**Sigma _____ M, O**

System for Integrated Genome Map Assembly graphical genome map editor. As a viewer, Sigma puts full color maps of the genome in the users hands to display, browse, manipulate, and print. It is capable of giving the user a perspective on an entire chromosome map, as well as an arbitrarily detailed view. Features allow the user to find specific parts of a map. Sigma allows the user to integrate data from a variety of sources. A convenient user interface makes data entry easy.

*Author:* Theoretical Biology and Biophysics Group at LANL, sigma@ncgr.org

*Home:* http://www.ncgr.org/sigma/-home.html

*Archive:* ebi/linkage_and_mapping/SIGMA

**Silmut _____ W**

Silmut helps you to identify regions in a sequence which can be altered to introduce restriction enzyme sites and other sequences by silent mutations.

*Author:* Raj Shankarappa, bsh@med.pitt.edu, K. Vijayananda, vijay@litsun.epfl.ch

*Archive:* iubio/ibmpc/silmut*

**Sim2 _____ M, W, O**

This program build local alignments of two sequences, each of which may be hundreds of kilobases long.

*Author:* Chao K-M et al., zjing@sunset.nlm.nih.gov

*Home:* ftp://ncbi.nlm.nih.gov/pub/sim2

**Sites _____ M, W, O**

Sites is a program for the analysis of comparative DNA sequences. It is primarily intended for data with multiple closely related sequences.

*Author:* Jody Hey

*Archive:* iubio/evolve/sites/

**SixcutterFreq _____ M**

The Hypercard stack calculates the frequency of the various six-cutter restriction enzymes in a few genomes, including bacteriophage lambda, Mus musculus, wheat, E. coli, Saccharomyces, Homo sapiens. The algorithm is based on the frequency of dinucleotide pairs.

*Author:* Brian Osborne, bosborne@nature.berkeley.edu

*Archive:* iubio/mac/sixcutterfreq*

**Sneath ST _____ W**

Statistical programs to screen a matrix of molecular sequences for atypical sequence

comparisons, and to simulate the addition of constant sites to randomly-placed differences in a molecular sequence comparison.

*Author:* P.H.A. Sneath, mjs@le.ac.uk

*Archive:* iubio/ibmpc/sneathst*

**Solupred M, W**

This spreadsheet allows one to predict the solubility of recombinant proteins in E. coli based on the amino acid content.

*Author:* Roger Harrison & Dan Diaz, BL275@cleveland.freenet.edu

*Archive:* iubio/mac/solupred-mac*, ibmpc/solupred*

**SorFind W**

SorFind is a DOS program that adds to DNA sequence files a feature table to annotate the location of putative coding exons.

*Author:* Gordon B. Hutchinson, hutch@netshop.bc.ca

*Archive:* iubio/ibmpc/sorfind*, ebi/dos/sorfin*

**Spectrum M, W**

Spectrum is a Macintosh and MS Windows program to read in phylogenetic data in Nexus format, and display the bipartition spectra corresponding to the data. It can also be used to find the tree whose expected spectrum is closest to the observed spectrum. It outputs spectra in Microsoft Excel and other formats.

*Author:* Michael Charleston, Roderic Page, m.a.charleston@bio.gla.ac.uk

*Home:* http://taxonomy.zoology.gla.ac.-uk/mike/spectrum/

**Spombe-Strain M**

S.pombe strain collection is a Hypercard stack for cataloguing the genotypes of S.pombe yeast strains. It is adapted from Kai-Uwe Frölich's yeast strain.

*Author:* Doug Drummond ddrummon@fs2.scg.man.ac.uk

*Archive:* iubio/mac/spombe-strain*

**ssu_rRNA M**

This Hypercard stack contains the entire Ribosomal Database Project sequence release 1. The sequences are accessible via a series of linked phylogenetic trees, a list, or by name.

*Author:* James W. Brown, jwbrown@mbio.ncsu.edu

*Archive:* iubio/mac/ssu-rrna*

**Swiss-PdbViewer M, W**

Swiss-PdbViewer is an application that can display PDB files. Several proteins can be analyzed and can be "piled-up" in 3D space. Differences can be calculated on selected amino acids of the aligned proteins, for comparison of active sites. It can also measure distances, angles, torsions angles between atoms as well as add/remove amino acids from the view. It includes many other features.

*Author:* Nicolas Guex, Manuel Peitsch, ng45767@ggr.co.uk

*Home:* ftp://expasy.hcuge.ch/pub/-PDBViewers/Prot3Dviewer, http://www.expasy.ch/spdbv/-mainpage.html, http://www.pdb-.bnl.gov/expasy/spdbv/-mainpage.htm

**tacg M, W, O**

A program for restriction enzyme and other analyses of DNA. This is a command-line program that can be used from MacOS, Wintel through others such as SeqPup.

*Author:* Harry Mangalam, mangalam@uci.edu

*Home:* http://hornet.bio.uci.edu/-~hjm/projects/tacg/

*Archive:* iubio/restrict-enz/tacg*

**TFPGA W**

TFPGA (Tools for population genetic analyses) is a Windows program for the analysis of allozyme and molecular population genetic data. The program calculates simple descriptive statistics, genetic distances, and F-statistics. It also performs tests for Hardy-Weinberg equilibrium, population differentiation and performs UPGMA clustering and Mantel Tests.

*Author:* Mark P Miller, mpm2@nauvax.ucc.nau.edu

*Home:* http://dana.ucc.nau.edu/~mpm2

**TopPredII M**

Prediction of transmembrane segments in integral membrane proteins, and the putative topologies.

*Author:* Claros M.G & von Heijne G., claros@cica.es, gvh@cbs.ki.se

*Archive:* ebi/mac/, iubio/mac/toppred*

**TOPS W, O**

Program to automatically generate and edit protein topology cartoons. These cartoons are two

dimensional representations of the secondary structure of proteins.

*Author:* Tom Flores, flores@ebi.ac.uk

*Home:* ebi/pub/contrib/TOPS

**Tree Draw Deck** **M**

A Hypercard stack which draws phylogenetic trees. It is derived from Drawgram and Drawtree of Phylip by J. Felsenstein

*Author:* D. Gilbert, software@bio.indiana.edu

*Home:* iubio/mac/treedraw*

*Archive:* ebi/mac/treedraw*

**Treecon** **W**

A package for the construction of phylogenetic trees. Its advantages include menu-driven, easy-to-use interface, quick, handles large datasets, large set of distance-measure methods and distance-matrix based tree construction tools, sophisticated options like subset resampling and test of outgroup influence, additional tools like production of partial alignments and indication of informative positions.

*Author:* Yves Van de Peer, yvdp@reks.uia.ac.be

*Home:* ftp://uiam3.uia.ac.be/

*Archive:* iubio/ibmpc/treecon*

**TreeView** **M, W**

A program for drawing phylogenies on MacOS and MS Windows. The program reads NEXUS, PHYLIP, Clustal W and similar tree formats.

*Author:* Roderic D M Page, r.page@bio.gla.ac.uk

*Home:* http://taxonomy.zoology.-gla.ac.uk/rod/treeview.html

*Archive:* iubio/mac/treeview*, iubio/ibmpc/treeview*

**Vised** **W**

A visual sequence editor/display software for Windows, including effective and easy-to-use interface. Features include edit up to 200 sequences of 18,000 bases; powerful pattern search function; support for many sequence file formats; supports extraction of sequences from library files; figure preparation; boxed output of sequence identities; imports MACAW alignments; protein sequence prediction in one or all six frames.

*Author:* Ken Peters, kpeters@qb.island.net

*Archive:* iubio/ibmpc/vised*

**Visual BLAST and FastA** **W**

These programs are designed for interactive analysis of full BLAST and FASTA output files containing protein sequence alignments. They implement analytical tools which automate detailed analysis of BLAST and FASTA outputs, and include tools for multiple alignment analysis.

*Author:* Patrick Durand et al., durand@lmcp.jussieu.fr

*Home:* http://www.lmcp.jussieu.fr/~durand/

**WinDot** **W**

A dotplot program for MS Windows.

*Author:* Ramin Nakisa, ramin@ic.ac.uk

*Archive:* iubio/ibmpc/windot, ebi/dos/windot

**WinMGM** **W**

Visualisation and manipulation tools for proteins, nucleic acids and organic molecules, including manipulation of molecules represented as CPK, stick and ball, ribbons and cylinders, and colorations by atomic type, atoms of a selected area, of an active site, and others. Several other features are included.

*Author:* Mehdi Rahman, Robert Brasseur, mehdirah@fsagx.ac.be

*Home:* http://www.fsagx.ac.be/-info_faculte/info_dep/info_bp/-mehdi/winmgm/winmgmen.htm

**WinSeq** **W, O**

The ReadSeq program for converting among biosequence file formats, for MS Windows.

*Author:* Ramin Nakisa, ramin@ic.ac.uk

*Archive:* iubio/ibmpc/winseq, ebi/dos/winseq, see iubio/readseq/ for others

**WPDB** **W**

WPDB (the Protein Data Bank through MS Windows) is a package with a compressed version of PDB and a set of tools to query features of a single structure or perform a comparative analysis on multiple structures with emphasis on sequence alignment and structure superposition.

*Author:* Ilya N. Shindyalov, Philip E. Bourne, bourne@sdsc.edu

*Home:* ftp://ftp.sdsc.edu/pub/sdsc/-biology/WPDB/

**Yeast Strain** **M**

This is a Hypercard stack for cataloguing the genotypes of S. cerevisiae yeast strains.

*Author:* Kai-Uwe Froehlich, cbkfr01@mailserv.zdv.uni-tuebingen.de

*Archive:* iubio/mac/yeaststrain*, ebi/mac/

# Index