@AbstractSeqPup is a biological sequence editor and analysis program for Macintosh computers. It includes links to network services and external analysis programs.Features include   multiple sequence alignment editor   single sequence editor window   read and write several sequence file formats   consensus,reverse,complement,degap operations   automatic preference saving   internet send mail   internet sequence analysis services   and more   **NOTICE**: This is an early, unfinished version of the program.   Expect it to fail in various ways. This release is made available to those who wish to test and comment on its future development. This release will expire on a date indicated in the About dialog.SeqPup is being written by Don Gilbert using DCLAP, a C++ Class Application framework, and founded on the NCBI Toolkit, especially it's Vibrant user-interface section written by Jonathan Kans.   DCLAP is based loosely on the MacApp extensible programming framework from Apple Computer.You can obtain updates of this release thru anonymous ftp to ftp.bio.indiana.edu, in folder /molbio/seqpup, as seqpup.hqx (Macintosh), seqpup.zip (MS Windows), seqpup-sun-sparc.tar.Z (SunOS Sparcstation), and perhaps other binaries. There may be additional distribution software, data or information in this /molbio/seqpup folder. See the Readme files in it for details. The file SeqPup.Help is a plain text file which may be read from your favorite wordprocessor or from the Edit program in the Apps folder. If you have problems getting SeqPup to launch, and cannot read this help from SeqPup, please read it with another application to help solve the problem.Comments, bug reports and suggestions for new features (see below) may be addressed via e-mail to SeqPup@Bio.Indiana.EduHistorySeqPup was started Sept. 1990 as MacApp sequence editor/analysis platform on which analysis programs from other authors, typically command line w/ weak user interfaces, could be easily incorporated into a useable Mac interface.@SeqPup HelpSeqPup provides a biological sequence editor and related functions, including links to network services and external analysis programs.
****

 ****   NOTICE   ****   **** This is an early, unfinished version of the   **** program. Expect it to fail in various ways.   **** This release made available to those who wish   **** to test and comment on its future development. ****   ***This program is s. a long-term research/development proj. It has already gone thru several changes since its start in September 1990. Don't expect it to be "complete" for a while (years) yet. It is publicly available now as it has been useful to some as is.Comments, bug reports and suggestions for new features (see below) may be addressed via e-mail to SeqPup@Bio.Indiana.EduWith any bug reports, I would appreciate as much detail as is reasonable without putting you off from making the report. Include description of computerfounded on the NCBI Toolkit, especially it's Vibrant user-interface section written by Jonathan Kans.   DCLAP is based loosely on the MacApp extensible programming framework from Apple Computer.You can obtain updates of this release thru anonymous ftp to ftp.bio.indiana.edu, in folder /molbio/seqpup, as seqpup.hqx (mac), seqpup.zip (mswin), seqpup-sun-sparc.tar.Z (sunos), and perhaps other binaries. There may be additional distribution software, data or information in this /molbio/seqpup folder. See the Readme files in it for details. If you obtain SeqApp thru InterneAny external applications that may distributed with SeqPup are copyrighted by their respective authors and subject to distribution provisions as described by those authors. @Installing[--- obsolete ------ needs rewriting for SeqPup -]@ViewsThere are now four main types of views or displays in SeqPup:A multiple-sequence view which is the primary display when you open a sequence document; the single sequence editting view; various print views which result from an analysis, like the Restriction map; and dialog views where you control some function.Many of these views have dialog controls -- push buttons, check boxes, radio controls and edittable text items -- to let you fine-tune a view to fit your preference. Many of these views also will remember your last preferences.

When a view has edittable text items, including the sequence entry views, most usual Mac undo/cut/copy/paste features will work, as will font, size and style controls. Two or more views of the same data are possible. Some of these are truly views of the same data -- changes made in one view are reflected in another. Other views are static pictures taken of the data at the time the analysis was performed -- later changes to the data do not affect that picture.@@Aligned multi-sequence viewThe main view into a sequence document is the multiple sequence editor window, which lists sequence names to the left and sequence bases as one line that can be scrolled thru. Bases can be colored (now only nucleic colorings) or black. Sequence can be editted here, especially to align them, and subranges and subgroupings can be selected for further operations or analysis. Entire sequence(s) can be cut/copied/pasted by selecting the left name(s). Mouse-down selects one. Shift-mouse down selects many in group, Command-mouse down selects many unconnected. Double click name to open single sequence view. Select name, then grab and move up or down to relocate.Select the lock/unlock button at the view top to lock/unlock text editting in the sequence line. With lock on (no editting) you can use shift and command mouse to select a subrange of sequences to operate on.Bases can be slid to left and right, like beads on an abacus, when the edit lock is On (now default). Select a base or group of bases (over one or several sequences), using mouse, shift+mouse, option+mouse, command+mouse. Then grab selected bases with mouse (mouse up, then mouse down on selection), and slide to left or right. Indels "-" or spacing on ends "." will be added and squeezed out as needed to slide the bases. See also the "Degap" menu selection to remove all gaps thus entered from a sequence.@@Single sequence viewFor entering/editting a single sequence, this view displays one sequence with more info and control. Edit the name here (later other documentation). Bring out this view by double-clicking sequence name in align view, or choosing Edit from Sequence menu.@@Print viewsVarious analyses provide non-editable displays. These are usually save-able as PICT format for editting in your favorite MacDraw program, or print-able.@Data filesSeqPup uses plain TEXT type files for its primary sequence data. These files can be exchanged without modification with many other sequence analysis programs. SeqPup automatically determines the sequence format of a data file when openning it. You have an choice of several formats to save it as. As of this writing, the GenBank format is prefered (see bugs).There is a "seqpup.cnf" (mac) / "seqpup.ini" (mswin) / ".seqpuprc" (unix) file which stores various user options like window positions, mail address, child tasks. Also look for "dclap..." prefs file. This automatically generated by SeqPup. It goes in a System Folder:Preferences: folder (mac) .

@FeaturesThe following topics describe main features found in the SeqPup menus.@@File **New** will create an align view of sequence data. New Text will create a plain text document, which is the format of the sequence data files also. **Open** will open an exising file.   The default choice will open a file of sequences into a new window. You can choose "Sequence, append", or hold down the SHIFT key, to open a sequence file and append it to an existing alignment window.Other **Open** options include opening a plain text file, a file of phylogeny trees in Newick format (see Phylip documentation), or a Gopher document.**Save**, Save as, Save a copy in, all will save the current document to disk files. Revert will restore the open align view to the last version saved to disk.**Save selection**, Saves only highlighted sequences to a new disk file. Doesn't affect save status of current full alignment document.**Print** setup, print will print the current view.**Help** brings up a view to page thru the help file.**Preferences** will set some user preferences. @@Editing**Undo**, **cut**, **copy**, **paste**, **clear**, **select all** -- these standard mac commands will operate on text as well as on sequences in (hopefully) intuitive, usual ways.**Find**, Find same, Find "selection" will search for strings in text.**Replace**, replace same will replace target strings (not yet enabled).@@Sequence manipulations**New sequence** -- append a new, blank sequence to the sequence document.**Edit** -- open single sequence editting view for selected items.**Reverse**, **Complement**, **Rev-complement** -- Reverse, complement or reverse+complement a sequence. Works on one or more sequences, and the selected subrange.**Rna-Dna,Dna-Rna** -- Convert dna to rna (t->u) and vice versa. Works on one or more sequences, and the selected subrange.

Degap -- remove alignment gaps "~".   Works on one or more sequences, and the selected subrange. Gaps of "-" are locked and not affected by Degap.   Works on one or more sequences, and the selected subrange.**Lock Indel & Unlock Indel** -- Convert from unlocked gaps "~", to locked gaps "-". Unlocked gaps will disappear and appear as needed as you slide bases left and right. Locked gaps are not affected by sliding nor by Degap.   Works on one or more sequences, and the selected subrange.**Consensus** -- generate a consensus sequence of the selected sequences.**Translate** -- translate to/from amino acid. Relies on Codon.Table data.**Pretty print** -- a prettier view of a single or aligned sequences. Use these views to print your sequences. Printing from the editing display will not be supported fully, and may not print all of your sequence(s).**Restriction map** -- Restriction enzyme cut points of selected sequence. Also protein translation options.**Dotty plot** -- provide a dot plot comparison of two sequences. **Nucleic, amino codes** -- These provide both reminders of the base codes, and a way to select colors to assocate with each code (new in v 1.9a). See below for some discussion of the two "aa-color" documents that now ship with SeqPup. @@InternetThe Internet features of SeqPup let you interchange ideas and data with people and biocomputing services around the world. If your Mac is connected already to the Internet, you probably are familiar with electronic mail and some of its uses. SeqPup includes a selection of network access features in the developing area of networked biocomputing. You will find access to me, at least to get comments and bug reports to me, very easy. There is a feature to send and receive e-mail, as well as mail links to customized e-mail services. These include searching for sequence similarity via BLAST and FastA programs on the Genbank/Intelligenetics computers, fetching sequences, data and software from Genbank and EMBL.There is now an feature called Gopher, which gives you access to a wide range of information services now developing on the Internet. Gopher is something like Telnet or FTP (file transfer), but also different. It includes some of the keyword searching features of WAIS (Wide Area Information Services). There are currently several biology gopher services found around the globe. These include fast and up-to-date keyword searches of GenBank, EMBL, PIR and other important biology databanks.@@@Mail PreferencesThe mail prefs dialog asks for your return e-mail address, your preferred SMTP mail host, and your POP mail address and password. These addresses may be similar.

Return e-mail address: This is where another person should send mail so it will reach you.   Example: Bob.Jones@University.Edu   or: bjones@MainComputer.University.Edu   SMTP Mail host: This is the internet address of the computer thru which SeqPup will send out mail to the rest of the world.   Example: MainComputer.University.Edu   @@@Send MailSend an electronic mail message. You must enter an address to send to, and have entered your return address in the mail preferences dialog.@Mail-based Search and FetchVarious network resources provide biocomputing services thru e-mail. These include retreiving sequence entries from the various databanks (GenBank, EMBL, PIR), fetching help documents, and searching for sequences in the databanks that match your query sequences.@BugsSeqPup version -1 has plenty of bugs and missing features, including: no Undo (this is a real bite to those used to it)   mostly no cut/copy/paste/clearlimited printing of documents or views   mostly no align-view manipulations (move,cut/copy,edit in place, shift, ...)   no pretty print viewsno restriction maps no dot plots   no ...   problems w/ window display & keeping track of active window (x,mswin)   I'll be adding back many of these features from the Macintosh SeqApp as time permits.SeqApp version 1.9a involved major amounts of rewriting of the underlying program code, without as many improvements as I would have liked.   This major rewrite was prompted by the move from Apple's extensible macintosh application framework, MacApp version 2 to version 3. The move was needed because MacApp 2 is essentially obsolete. New general application features are found in MacApp version 3, a few of which you may notice. Also, this is preliminary to translating the code from the Object Pascal language into the C++ language. The future of MacApp is to move to a computer-make independent application framework called Bedrock, which may be available in 1994 for Mac, Windows, and possibly Unix systems. It is my hope to move SeqPup onto this, but I make no promises or predictions at this time.

There are still several known bugs, and probably some unknown bugs lurking in this release.   A prime rule in using this software is to be aware of its incomplete nature, and **save early and often**. Please also send in detailed descriptions of bugs you find.   If the bug seems to be erratic, as many are, try to find some circumstances which will cause it to appear repeatable and explain these to me.   Your reports will help improve this program for you and others. E-mail bug reports to

SeqPup@Bio.Indiana.Edu@Future FeaturesThe future of MacApp on which SeqApp is based, is to move to a platform independent application framework called Bedrock, which may be available in 1994 for Mac, Windows, and possibly Unix systems. It is my hope to move SeqPup onto this, but I make no promises or predictions at this time.[Winter 1994: The future is here -- BedRock is dead, but DClap is here instead to provide a platform independent framework.]Here is a list of things which may be added to SeqPup in the future, depending on your interest.Sequence documentation handling. Currently no provisions for documentation per sequence. This will at least change to a window for any comments and saving it into files (where file format permits). Possibly I will put effort into dealing with the features, references, etc., in a fashion along lines of Genbank/EMBL documentation structure &/or Authorin documentation. Your comments on the importance of this are desired.Feature table parsing -- pull out subsequences from Gen/EMBL feature info.Align, single sequence pretty print -- header, page numbering user prefs should be added.Restriction map -- Could use some speed-up. Some would like graphic map (i.e., one line or circle w/ cut points per zyme).NCBI ASN.1 file format. -- this will be available soon after the 1.9a 157 release, as the newest Readseq code which reads/writes ASN.1, is fully rolled into SeqPup.Open reading frame mapping, other single sequence analyses (what?)Simple protein analysis routines, better protein handling (colors are now nuc specific). -- colors are user selectable now in 1.9a.   I expect to put most/all of the features of DNA Strider into SeqPup -- let me know your favorite ones.Links to child tasks: these depend as much on agreement of the authors of various routines as on user interest. Among routines that I already have a handle on -- MulFold (Zuker), Phylip phylogeny analysis (Felsenstein), tree draw program (Felsenstein + my hacks a la Tree Draw desk).

Links to NETWORKED child tasks (e.g., on fast compute servers).   Perhaps thru Gopher, or otherwise.@HistorySeqPup/SeqApp was started Sept. 1990 as MacApp sequence editor/analysis platform on which analysis programs from other authors, typically command line with weak user interfaces, could be easily incorporated into a useable Mac interface.SeqPup, March 1994, version -1. First public release of SeqPup, the platform-independent version derived from the earlier Macintosh SeqApp. SeqPup will eventually include most or all of SeqApp's features, and new features will be added. SeqApp, 20 June 93, version 1.9a162 -- a minor update, with these bug fixes since last release: -- fixed base number index in align view to correctly index the bases. Also fixed base number index scrolling link to sequence scrolling. -- fixed paste of align sequence into blank window to correctly update full width of display.12+ June 93, version 1.9a157+ -- a semi-major update, and time extension release with various enhancements and corrections. These include -- lock/unlock indels (alignment gaps). Useful when sliding bases around during hand alignment, to keep alignment fixed in some sections. -- color amino (and nucleic) acids of your choice.   -- added support for more sequence file formats: MSF, PAUP, PIR. SeqApp now relies on the current Readseq code for sequence reading & writing. -- save selection option to save subset of bases to file. -- addition the useful contig assembly program CAP, written by Xiaoqiu Huang. -- major revision of preference saving method (less buggy, I hope) -- major revision of the underlying application framework, due to moving from MacApp 2 to MacApp 3. -- fixed a bug that caused loss of data when alignment with a selection was saved to disk.5 Oct 92, version 1.8a152+ -- a semi-major update with various enhancements and corrections. These include - corrections to the main alignment display, - improvements to the help system, - major changes to the sequence print-out options, -- including addition of a dotplot display (curtesy of DottyPlot), -- a phylogeny tree display (curtesy of TreeDraw Deck & J. Felsenstein's DrawTree), -- improved Pretty Print, which now has a single sequence form and a better aligned sequence form, -- improved Restriction map display, - addition and updating of several e-mail service links, -- including Blast Search and Genbank Fetch via NCBI,-- BLOCKS, Genmark, and Pythia services,- updated Internet gopher client (equal to GopherApp),- editable Child Tasks dialogs- addition of links to Phylip applications as Child Tasks- addition of Phylip interleaved format as sequence output option

11 June 92, version 1.6a35 is primarily a bug fix release. Several of the disasterous bugs have been squashed. This version now works on the Mac SE model, except for sendmail. No new features have been added. 7Jun92, v. 1.5a?? -- fixed several of the causes of mysterious bombs (mostly uninitialized handles), link b/n multiseq and 1-seq views is better now, folded in GopherApp updates, death date moved to Jan 93, 25Mar92, v1.5a32 (or later). First release to general public. Includes Internet Gopher client. Also released subset as GopherApp for non-biologists.4Mar92, v 1.4a38 -- added base sliding in align view. Bases now slide something like beads on an abucus. Select a section with mouse, then grab section and shift left or right. Gaps are inserted/removed as needed. For use as contig aligner, still needs equivalent of GCG GelOverlap to automatically find contig/fragment overlaps.Also added "Degap" menu item, to remove "." and "-". Fixed several small bugs including Align pretty print which again should display.2Mar92, v 1.4a19 -- fixed several annoying bugs, see SeqApp.Help, section on bugs for their resolution. These include Complement/Reverse/Dna2Rna/ Translation which should work now in align view; Consensus menu item; entering sequence in align window now doesn't freeze after 30+ bases; pearson/fasta format reading; ...10Feb92, v 1.4a6 -- fix for Mac System 6; add Internet service dialogs for Univ. Houston gene-server, Geneid @ BU, Grail @ ORNL; correct About Clustalv attribution.5Feb92, v 1.4a4 -- limited release to network resource managers, clustalv authors, testers.Vers 1.4, Dec91 - Feb92. Dropped multi-sequence picker window, made multi-align window the primary view (no need for both; extra confusion for users). added pretty print, restriction map, sequence conversions. Generalized "call clustal" to Hypercard-like, System 7 aware menu for calling external tasks. Fleshed out internet e-mail objects, added help objects, window menu, nucleic/amino help windows. Many major/minor revisions to all aspects to clean out bugs.   Preliminary release to a limited set of testers (1.4a?)Vers. 1.3, Sept - Dec91. Modified clustalv for use as external app (commandline file, background task, ...). Added basic Internet e-mail routines call clustal routine (preliminary child task) Many major/minor revisions to all aspects to clean out bugs. Jun91-Aug91: overwork at other tasks kept SeqApp on back burner.Mar91-Jun91: not much work on SeqApp, fleshed out TCP methods (UTCP, USMTP, UPOP).Feb 1991, vers 1.2? made available to Indiana University biologists and NCBI biocomputists.Vers. 1.1, Oct 1990, multiple sequence picker and multiple sequence alignement window, including colored bases, added to deal with alignment and common multi-sequence file formats.

Version 1, Sep 1990. Single sequence edit window + TextEdit window, from MacApp skeleton/example source + readseq.