Microcomputers and Phylogenetic Analysis

WILLIAM L. FINK

The rise of empiricism in systematic biology has been one happy result of the changes that have shocked that discipline during the past 20 years. Emphasis on quality of data and data analysis has replaced the reliance on authority that marked systematic biology for much of its history. This transformation is largely due to the widespread application of phylogenetic methods (*1*). Phylogenetics is based on two assumptions. First, species are connected genealogically, through ancestor-descendant relationships. These genealogies can be represented as branching trees or graphs, called phylogenies or cladograms. Second, we assume that the properties of species change during the course of evolution and that some of these changes are passed on to descendant species. These transformed traits serve to mark the latter as having evolved from the ancestral species that developed a particular trait. Much of phylogenetics involves the discovery and analysis of these traits, or characters, with the aim of detecting the genealogical connections among taxa.

One area that has seen significant improvements recently is numerical phylogenetics. There are several approaches to computation of phylogenetic trees, and some of these algorithms have their origin in other fields, such as applied mathematics and graph theory. One example of this is the *NP*-complete question of finding the shortest distance between some number of points (Steiner's problem or the "traveling salesman problem") (*2*). In phylogenetic terms, this is equivalent to finding the evolutionary tree with the fewest character changes for a data set. The number of possible trees grows very rapidly with increasing numbers of taxa, and for large data sets exact solutions are not possible with current computational techniques.

The microcomputer revolution has also affected systematics, and some very sophisticated phylogenetic algorithms are being written for microcomputers, thus providing powerful tools to a wide audience. This review will examine some of the more popular and widely available programs that are designed explicitly to compute phylogenetic trees. The review is restricted to programs that analyze discrete (character) data. Continuous and distance data present more serious challenges to a reviewer (and a user). Continuous data, often mensural, consist of taxic descriptions which partially overlap; coding these data in a nonarbitrary way is problematic. Distance data consist of numerical matrices representing the similarities or differences between pairs of taxa; these numbers are derived from the transformation of observational data, such as DNA-DNA hybridization melting points or allozyme states. Although there are many distance programs, especially in molecularly oriented laboratory settings, only a few are available for microcomputers; most are "in-house" programs customized for particular mainframe computers and their operating systems. Distance clustering remains controversial after a series of papers on the theoretical merits of the algorithms and the data themselves (*3*).

Parsimony, Probability, and Compatibility

There are three approaches to dealing with character data in phylogenetic analysis: parsimony, probability, and compatibility. The theoretical assumptions and justifications of these are debated abundantly in the literature (*4*). It must be emphasized that any of these programs should be used only after the biologist is thoroughly familiar with the arguments and is willing to accept the assumptions required by the algorithm of choice. There is a great deal of sociology underlying papers on these alternatives, and the reader may find the going rough. Even in the instructional materials for some of the programs reviewed here there is some reason to "read between the lines" to evaluate statements being made.

Basically, parsimony algorithms seek trees that require the fewest character changes, thus minimizing the need to invoke convergence or reversal (homoplasy) in explaining the data. All characters, whether homoplasious or not, are considered evidence bearing on phylogenetic hypotheses. Probability methods base their tree-building strategies on some initial assumptions regarding evolutionary rate changes. These algorithms are usually recommended for molecular data, since such data are thought to fit the models better than morphological data. So far, probability models are rather primitive, often with assumptions that are oversimplified for computational efficiency. I did not test any of the available probability algorithms in this review (of the programs evaluated only PHYLIP includes them). Compatibility methods attempt to find the largest sets of characters that show no homoplasy and then build trees based on those characters only. Characters that show homoplasy in the context of a tree are not considered to be evidence bearing on that tree. Some users advocate that homoplasious characters be "fitted" by a secondary analysis to the tree (or trees) formed by the largest cliques, but neither the methods nor the rationale for doing so have been detailed fully (*5*). Clique methods have also been considered as tools for understanding character distributions (*6*).

Evaluation of phylogenetics programs can be difficult, in that there is no "truth" against which to compare their results. Parsimony programs can be compared against each other, because their goals are to find the shortest tree per data set. Compatibility programs can also be tested against each other, in the sense that they should always find the same cliques. But because of these differing assumptions and goals, any comparisons between the results of the clustering methods must be based on evaluation of the methods themselves, a task not attempted here. In addition, this review does not encompass a detailed "horse race" of central processing unit seconds used to find trees. Although numerous data sets were run, only three were selected to indicate relative speed and results, as representative of matrix size classes. One must be aware that PHYLIP and PAUP both undergo frequent modification, and the results reported herein are subject to change.

Program Specifications

Table 1 describes the software specifications and hardware requirements for the programs reviewed. PAUP [by David Swofford of the Illinois Natural History Survey (*7*)] is a package that is strictly for parsimony analysis. PHYLIP [by Joseph Felsenstein of the University of Washington (*8*)] performs parsimony, probability, and compatibility analysis. CLINCH [by Kent Fiala, of the State University of New York at Stony Brook (*9*)] is a program for doing compati-

Department of Biology and Museum of Zoology, University of Michigan, Ann Arbor, MI 48109.

Table 1. Software specifications and hardware requirements of programs reviewed.

	PAUP	PHYLIP	CLINCH	MacClade
Type	Parsimony	Parsimony, probability, compatibility	Compatibility	Character analysis, parsimony
Version tested	4.21	2.8*	6.2	1.0
Language	Compiled Fortran-77, Assembler	Source code Pascal†	Compiled Fortran-77‡	Compiled Pascal
Host computer§	IBM PC and compatibles	Many	MS-DOS machines	Apple-Macintosh
Software required	DOS 2.0 or higher	System-dependent‖	DOS 2.0 or higher	Supplied
RAM in kilobytes needed	256 or 512¶	NR#	NR#	512
Math co-processor	Required	Optional	Optional	Inapplicable
Disk drive requirements	2 (hard disk recommended)	1	1	1
Mode of operation	Interactive or batch	Batch**	Batch	Interactive
Documentation	Extensive	Extensive	Moderate	Moderate
Error diagnostics	Extensive	None	Some	Some
Mainframe version available	Yes	Yes	Yes	No
Cost	$50	Free††	Free††	$6.50

*MIX program was version 2.9. †Available in compiled Pascal, see (*11*). ‡Source code distributed. §Reviewer's computations were done on a Zenith 161 microcomputer (an 8088-based IBM compatible) with clock speed of 4.77 megahertz, 640 kilobytes of random access memory (RAM), a 20-megabyte hard disk, and an 8087 math co-processor. ‖IBM version (*11*) requires PC-DOS. ¶Two versions available, see Table 2. #No recommendation by author of the program. **Simple DOS shell available for interactive mode, see (*11*). ††User must supply media (tapes, floppy disks).

bility analysis. MacClade [by Wayne Maddison of the Museum of Comparative Zoology at Harvard (*10*)] is mainly a graphics program for visualization of character distributions on phylogenetic trees, although it does have a primitive parsimony algorithm. A word processor or editor is needed for data entry with all four programs. The first three programs are available for mainframe computers. The price of either the mainframe or microversion of PAUP is the same; site licenses are available for the latter. If Swofford has not installed PAUP on a mainframe that is similar to the user's, he prefers to do it himself, at a cost that covers his expenses. Such installation usually includes a seminar for instruction.

PHYLIP is distributed as uncompiled source code, and the user must obtain a Pascal compiler. The instructions include hints on what program changes are needed for successful running with several popular compilers. In most cases the user will have to do some minor programming, or have it done. Recoding is necessary to make the programs run interactively, and may have to be done to get output to a permanent file. PHYLIP is available compiled for the IBM PC and compatibles, with a simple DOS shell that makes the program easier for the novice to use (*11*). Details of the distribution and availability of the generic package can be obtained from Felsenstein (*8*).

CLINCH, in its compiled form, is limited to 50 taxa and 50 characters, a size too small for some serious users. The code included in the package can be reconfigured if one has access to an appropriate compiler (Microsoft Fortran-77, version 3.2, was used by Fiala). A recompilation by Julian Humphries (using the Lahey F-77 compiler, version 2) to increase the number of taxa revealed a code error that Fiala will correct in future releases (*12*). This program is memory intensive; how large a matrix could be run on a microcomputer is not clear.

Support for users can be an important part of a software purchase. Swofford (PAUP), Felsenstein (PHYLIP), and Maddison (MacClade) have newsletters that are sent to registered users to inform them of updates, bug fixes, and enhancements. Swofford and Felsenstein usually are available for consultation, but they prefer to be contacted via electronic mail [see (*7*) and (*8*)]. They also prefer to have the user read the manuals and other instructional materials before contacting them!

Discussion

The authors of these programs have deliberately chosen different strategies regarding portability, ease of use, and preferred algorithms. The following descriptions highlight strengths and weaknesses of each. Among the features that I consider important are: correct results (according to the assumptions of the method), ease of use, quality of output, speed of execution, and availability of several clustering options. None of the programs offers all of these desirable features, and readers will have to determine what features are best suited to their research programs. Operating features of these programs are summarized in Table 2. In Table 3 results are summarized for running three representative data sets using various programs from PAUP, PHYLIP, and CLINCH.

Both PAUP and PHYLIP have facilities for analyzing DNA sequence data, and PHYLIP has a program for analysis of protein sequence data. There are some significant differences in the ways in which discrete "morphological" data and molecular data are analyzed. With the latter, the program may need to place restrictions on the kinds of substitutions that can be made, there are special needs for coding, and there must be capabilities for noting deleted sequences, for removal of redundant or noninformative sites, and for ordering the transformations. The issue of hypothesizing homologies further complicates the picture. I did not do sequence data analyses for this review.

Parsimony Analysis—PAUP and PHYLIP. The most impressive package currently available is PAUP (Phylogenetic Analysis Using Parsimony). The program is fast and has great flexibility in its clustering strategies. It finds multiple solutions, allows for ordered or unordered data, and has some very sophisticated output for data analysis. Documentation for PAUP is extensive, and includes a well-written instruction manual with numerous examples. The package includes a program to compute consensus trees that summarize the similarities of multiple-tree solutions. In addition, there is the ability to generate publication-quality plots on commonly available plotters (through a program primarily written by Chris Meacham; this plotting program is also available in PHYLIP, below). Of all the features I listed above as desirable, all that is lacking in PAUP is choice of other algorithms—this is strictly a parsimony program. The micro package is available only for IBM PC's (including the AT and XT) and their compatibles. This specialization accounts for some of the interactive features. There are also "pop-up" help windows with most of the options listed; this is a handy feature for new users, but as it is memory-resident experienced users with large data sets probably will prefer to have the extra memory. The program can be used for either standard

Table 2. Operating features of programs reviewed.

	PAUP	PHYLIP	CLINCH	MacClade
Maximum number of taxa/characters				
As distributed	30/100 (256-kbyte version)* 50/330 (512-kbyte version)	50/200*	50/50	80/200
Possible	Configurable by author	?	?	Same
Missing data allowed	Yes	In some programs	Yes	Yes
User tree input	Yes	Yes	Yes	Yes
Character weighting possible	Yes	In most programs	No	Yes
Unordered characters possible	Yes	In DNA-PARS	No	Yes
Multiple trees found	Yes	In PENNY and CLIQUE	Yes	Inapplicable
Exact solutions available	Yes†	Yes†	Yes	Inapplicable
Consensus trees	Yes	Yes	No	No

*Some programs/options limited to fewer taxa. †Usually fewer than 20 taxa in PAUP; fewer than 10 in PHYLIP's "Branch and Bound" (PENNY).

character data or molecular sequence data. Symbols in the matrix can be defined by the user. Swofford also distributes on request a program (REDSEQ) that processes sequence data, eliminating uninformative positions and recoding the data for input into PAUP.

PAUP offers flexibility in search strategy, with options that guarantee shortest solutions and heuristic algorithms that are usually much faster than the former but provide no guaranteed results. "Alltrees" is a brute-force program that produces all possible solutions, and "branch and bound" is a more sophisticated algorithm for finding shortest possible trees for larger data sets. Both of these methods get "guaranteed" results but are restricted to limited data sets (nine or fewer taxa allowed for "alltrees," and usually fewer than 20 for "branch and bound"). The "cleanness" of the data will affect running time, and the user will have to determine whether either of these options is practical. The other search strategies are heuristic, trying to find the shortest trees by various methods of "branch-swapping" (either locally or globally), a process in which alternative placements of taxa are tried. Most of these should be run using the "mulpars" option, which specifies that multiple trees should be sought. Other options include alternative addition sequences of taxa, specification of numbers of trees to be held in memory while building trees, and deletion and restoration of taxa or characters. None of the heuristic strategies is guaranteed to find the shortest solution. In most cases, the default settings are usually the best for finding shortest trees, but some data sets will require some alterations of these options. It is always wise to try a series of alternative strategies. For example, in analyzing data set 2, the combination of mulpars and global branch-swapping found two trees of length 224 (there are at least 40 of this length). Changes in the addition sequence and number of trees held in memory found three trees at length of 221. No other program was able to find trees of that length with this data set. Clearly, running just one or two options would have resulted in a suboptimal solution. Since even large data sets run in PAUP in just a few minutes on a microcomputer, one can be reasonably sure of the results in a relatively short amount of time.

One of the strongest features of PAUP is the range of output available. Character changes may be listed by taxon, indicating supporting characters for each taxon and branch, or by character, indicating on which branch (or branches) each character changes state. Other options include various character optimizations (alternative branch placements for characters that cannot be placed unambiguously in a tree) and figures indicating the smallest and largest numbers of characters supporting each taxon and branch under the range of optimizations available in the package (one might be leery about expounding on the evolution of a group when there are character optimizations that show no data supporting that group). For large data sets, use of some of these options can result in voluminous output. In these cases it

Table 3. Benchmark tests of phylogenetic programs on three data sets using reviewer's computer. Data set 1 has 12 taxa, 32 characters; data set 2 has 26 taxa, 111 characters; data set 3 has 27 taxa and 323 characters. Sets 1 and 3 have some characters coded as missing. See (*13*). These are typical representatives of several runs of each data set with each program.

Program	Options	Elapsed time (hr:min:sec)	Number of trees	Tree length
	Data set 1			
PAUP	Local swap	0:00:53	9	46
	Global swap	0:01:10	9	46
	Branch and bound	0:00:33	9	46
PHYLIP				
MIX	Local swap	0:03:47	1	46
	Global swap	0:11:50	1	46
PENNY (branch and bound)	Inapplicable	5:11:36	9	46
METRO	Inapplicable	0:17:30	1*	46
CLIQUE†	Inapplicable	0:00:15	1	21‡
CLINCH†	Inapplicable	0:00:19	1	21‡
	Data set 2			
PAUP	Local swap	0:02:23	2	224
	Global swap	0:05:10	2	224
	Addseq-Simple	0:03:44	3	221
PHYLIP				
MIX	Local swap	1:15:52	1	224
	Global swap	6:58:10	1	222
METRO	Inapplicable	1:19:00	1	224
CLIQUE	Inapplicable	2:39:37	6	59§
CLINCH	Inapplicable	0:28:40	6	59§
	Data set 3			
PAUP	Local swap	0:06:45	8	574
	Global swap	0:11:24	8	574
PHYLIP‖	Inapplicable	N/A	N/A	N/A
CLINCH¶	Inapplicable	N/A	N/A	N/A

*Two topologies were found after several runs. †The single character coded as missing in this matrix was recoded to 0 so that CLIQUE results could be compared to CLINCH. ‡A single largest clique of 21 characters was found. §Six maximal cliques of 59 characters were found. ‖Unable to dimension programs to take this large data set; this may be a problem with the compiler (Turbo Pascal, version 3.0). ¶Unable to compile this program to take this large data set because of compiler or machine limitations.

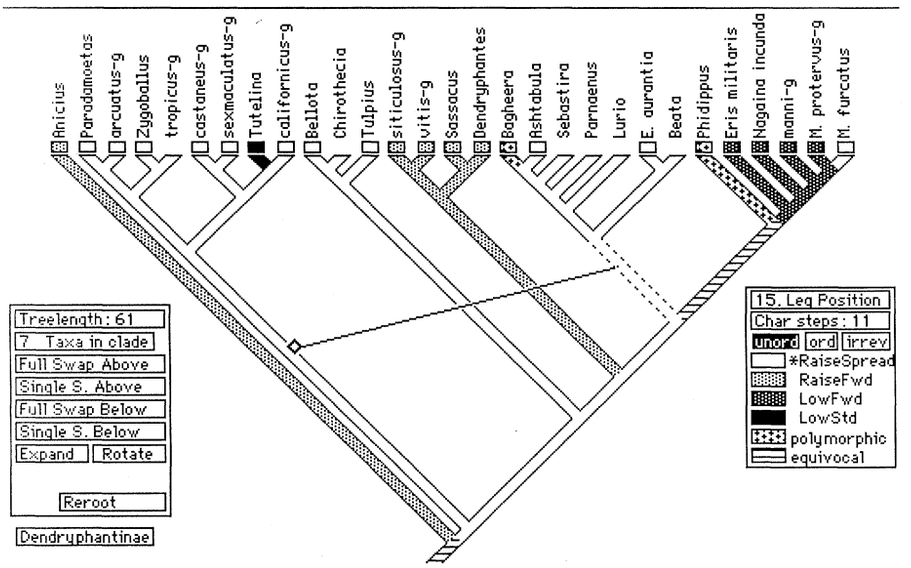

Fig. 1. MacClade program for character analysis, shown on the screen of a Macintosh microcomputer. Character states are indicated by patterns on the cladogram. A branch of the tree on the right-hand side has been grabbed and is about to be dropped (and thus relocated) onto a branch on the left. The organisms whose relationships are represented are jumping spiders. The character under consideration is first leg position during courtship.

may be wise to use the mainframe version, if only to have access to very high speed printers.

Both PAUP and PHYLIP (to be treated below) permit input of tree topologies. This allows the user to enter trees from the literature, for example, to see their length and character distributions in the context of the user's data matrix. This is a handy feature, especially now that more systematists are publishing data in a form that can be coded for computer analysis.

PHYLIP includes programs for plotting output, computing consensus trees, and finding trees with the use of parsimony, probability, and clique algorithms. In addition there are several programs that incorporate a "bootstrap" technique to estimate confidence limits around a phylogeny. This technique samples the data set to construct a new matrix and then uses the general parsimony programs (MIX, DOLLOP, DNAPARS) to generate trees. Felsenstein lists criteria for statistical evaluation of a group's existence. There is also a "branch and bound" algorithm for character as well as DNA data, although it is possible under some conditions for the program to finish without finding the most parsimonious tree (or trees). The probability programs are intended for use with molecular data. I did not use these programs, but the user may find that these programs are rather slow, as are most of the programs in PHYLIP.

Output from PHYLIP is simple and does not include the extensive character lists and optimization procedures of PAUP. Optionally, in some programs characters of hypothetical ancestors are printed out, and optimization can be done by hand.

The availability of several methods in one package is attractive, allowing the user to examine the results of different assumptions. The assumptions that Felsenstein considers each program to make about evolutionary process are outlined. Unfortunately he cites only his own work, so a naïve user might not be exposed to alternative explanations of these methods.

PHYLIP includes some novel approaches to parsimony tree-building strategies, such as METRO, a program that incorporates a random element into tree choice, thereby, it is hoped, preventing the program from getting "stuck" in a local minimum when there are other areas of character space that would yield shorter trees. My results with METRO were not promising, since it never got a shorter tree than PAUP found and often did not find a tree as short. There is a random-number generator that must be "seeded" before the run, and it is quite possible that, had I reseeded it enough times, the program would have performed as well as PAUP. Unfortunately, METRO is very slow, and multiple runs can be very time-consuming.

MIX is a program that allows the user to choose whether a character can reverse states (Wagner parsimony) or not (Camin-Sokal parsimony). By changing parameters, this program can produce trees that allow character reversal, no character reversal (=clique methods), or any combination of the two. One can also choose a global branch-swap option. This program also does not find multiple trees, and Felsenstein recommends at least ten runs per matrix, with changes made in the taxon entry sequence. I followed these instructions, but my experience indicates that with some data this will be insufficient, and the time needed for thorough searching could be enormous. For example, data set 2, with the use of the global swap option, ran for nearly 6 hours and 45 minutes before finding a solution, a tree that was one step longer than the three trees PAUP found (several other runs found only longer trees). The molecular parsimony programs use basically the same algorithms as the "standard" parsimony programs and can be expected to produce comparable results in terms of speed and quality of solution.

There are some serious limitations to PHYLIP as a research tool for parsimony analysis. The first is that only one of the programs (PENNY, which is limited to very small data sets) is able to find multiple trees. The second is that there are data sets for which the parsimony programs do not get trees as short as PAUP does. In some cases, this may be a function of the number of times the user is willing to run data through a given procedure, but the slowness of the programs is a hindrance to multiple runs, especially when the potential number of alternate topologies is large. It is fairly common to find numerous alternate topologies with PAUP, and a user of PHYLIP would have trouble deciding when to stop trying. Interpretation of the output from the bootstrap programs would also be suspect, since the search algorithms do not always find the shortest trees. Felsenstein has stated that multiple tree capability will be added in a future release. Until that is done and the efficiency and effectiveness of the algorithms (including the compatibility algorithms; see below) are improved, there is little but price to make this package attractive. I would be wary of published results of an analysis from it, and it is really too slow for use as a teaching tool.

Compatibility Analysis—CLINCH and CLIQUE. CLINCH was developed on a mainframe during the era of card readers, and that heritage shows, as it is a "batch" program without the bells and whistles of PAUP. Input consists of a series of batch commands, including a character-state tree for each character (if all characters are binary, no such trees are required). Default output consists of a listing of options in effect, the input matrix, the compatibility matrix (indicating whether a character is compatible with other characters), the total number of characters a particular character is compatible with, a list of cliques (up to a

value set optionally), a distribution table of clique sizes, and "interpretation" of some of the cliques. The "interpretation" is the listing of the set of taxa (including hypotheticals) defined by that clique (this is a written description of a Hasse diagram). Optionally, the set description (the Hasse diagram itself) and a cladogram can also be output. Other options allow for input of missing data and user-defined cliques, and secondary analysis, in which only cliques containing a specified set of characters will be found.

The CLIQUE program in PHYLIP works fairly rapidly on small data sets and can handle reasonably large data sets (Table 2). An option allows the user to specify a minimum clique size for search and printout. Unfortunately, the version tested cannot handle missing data, and Felsenstein suggests using his MIX program with options that effectively make it a compatibility program. But MIX could not find a range of trees equivalent to the multiple cliques that CLIQUE can find. I was unable to run data set 3 with either of these programs, apparently because of a dimensioning problem with the compiler (Turbo Pascal 3.0). On the small data set CLINCH and PHYLIP's CLIQUE are about equal in speed, but on the larger matrix CLINCH is considerably faster.

Graphics-Based Character Analysis. Although it is not really a program intended for rigorous searching for phylogenetic trees, MacClade is a remarkable program that points the direction for future developments in this field. It takes advantage of the graphics capabilities of the Apple Macintosh to allow visualization of phylogenetic trees with their character distributions (Fig. 1). Although the program has a relatively primitive tree-building algorithm in it, most users will want to port a data matrix to the Apple (or enter it as described in the manual), together with a tree topology obtained from another program. Once that is done, however, the fun begins, as the user can do an impressive number of manipulations, including moving tree branches about with the mouse, changing the reversibility or weight of a character, and rerooting the tree, all with nearly immediate graphic representation of how the distribution of any specified character is changed under the new conditions. If a branch is moved, for example, the length of the new tree is calculated and the distribution of the character is altered as appropriate. Graphics patterns represent different character states. The states of seven additional characters for the terminal taxa can be shown on the screen as well. The program is an excellent exploratory device, allowing one to see immediately how given topologies would affect hypotheses of character evolution. The next version of MacClade, to be released before the end of the year, will be more powerful and flexible and will feature ability to share data files with PAUP and PHYLIP. The emerging generation of scientific workstations will be ideal for programs such as this, and more sophisticated, graphics-based character analysis will become commonplace.

Summary

The programs discussed above show how microcomputers have added to the arsenal of systematic biologists. This is a rapidly developing field, and there are no doubt major changes on the horizon. Swofford is working on a new version of PAUP that will have some of the interactive features of MacClade (and will not require a math coprocessor) and there are efforts under way to make PAUP available on Macintosh. PHYLIP has undergone a steady evolution since its release, and Felsenstein has plans to continue that policy. A microcomputer descendant of the large mainframe program PHYSYS, authored by James S. Farris, is supposed to be forthcoming before the end of the year.

Just as this review was being completed, J. Rohlf and R. Sokal of the State University of New York at Stony Brook unveiled a beta-test microcomputer version of their phenetic program package, NT-SYS. Inasmuch as it was not in release form and does not include algorithms specifically designed to do phylogenetic analysis, it has not been included here (although some phenetic techniques, such as UPGMA, produce results similar to parsimony trees under certain assumptions).

REFERENCES AND NOTES

1. N. Eldredge and J. Cracraft, *Phylogenetic Patterns and the Evolutionary Process* (Columbia Univ. Press, New York, 1980); W. Hennig, *Phylogenetic Systematics* (Univ. of Illinois Press, Urbana, 1966); E. O. Wiley, *Phylogenetics: Theory and Practice of Phylogenetic Systematics* (Wiley, New York, 1983).
2. R. L. Graham and L. R. Foulds, *Math. Biosci.* **60**, 133 (1982); W. H. E. Day, *J. Theor. Biol.* **103**, 429 (1983).
3. J. S. Farris, in *Advances in Cladistics*, V. A. Funk and D. R. Brooks, Eds. (New York Botanical Garden, Bronx, NY, 1981), p. 3; *Cladistics* **2**, 144 (1986); J. Felsenstein, *Evolution* **38**, 16 (1984); *Cladistics* **2**, 130 (1986).
4. G. F. Estabrook, *Syst. Bot.* **3**, 146 (1978); J. S. Farris, vol. 2 of *Advances in Cladistics*, N. I. Platnick and V. R. Funk, Eds. (Columbia Univ. Press, New York, 1983), p. 1; J. Felsenstein, *Quart. Rev. Biol.* **57**, 379 (1982); *Annu. Rev. Syst. Ecol.* **14**, 313 (1983); E. Sober, *ibid.*, p. 335.
5. G. F. Estabrook, J. G. Strauch, K. L. Fiala, *Syst. Zool.* **26**, 269 (1977).
6. G. F. Estabrook, *Cladistics: Perspectives on the Reconstruction of Evolutionary History*, T. Duncan and T. F. Stuessy, Eds. (Columbia Univ. Press, New York, 1984), p. 135; C. A. Meacham, *ibid.*, p. 152.
7. David L. Swofford, Illinois Natural History Survey, 607 East Peabody Drive, Champaign, IL 61820. BITNET address: DAVESWOF@UIUCVMD.
8. Joe Felsenstein, Department of Genetics, SK-50, University of Washington, Seattle, WA 98195. USENET address: uw-beaver!entropy!uw-evolution!joe. Programs available through electronic mail.
9. Kent Fiala, Department of Ecology and Evolution, State University of New York, Stony Brook, NY 11794.
10. Wayne Maddison, Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138.
11. George D. F. Wilson, A-002, Scripps Institution of Oceanography, La Jolla, CA 92093.
12. Dr. Julian Humphries, Department of Anatomy, University of Chicago, 1025 East 57 Street, Chicago, IL 60637.
13. Data set 1 is a modification of the data published by W. L. Fink and S. V. Fink, *Copeia* **1986**, 494 (1986). Data set 2 is from B. Chernoff, *Proc. Acad. Nat. Sci. Phila.*, in press. Data set 3 is from W. L. Fink, *Misc. Publ. Mus. Zool. Univ. Mich. 171* (1985). All are available on request.

14. I thank Dr. Julian Humphries for technical help and suggestions.