MacClade

Analysis of Phylogeny and Character Evolution

Version 3

Wayne P. Maddison
University of Arizona

David R. Maddison
Harvard University

Sinauer Associates, Inc. • *Publishers*
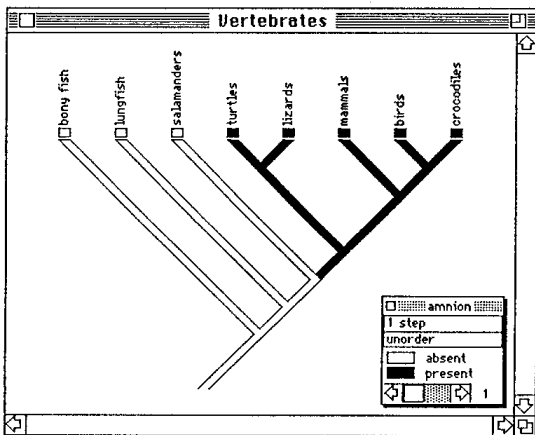
SUNDERLAND, MASSACHUSETTS, U.S.A.

1. INTRODUCTION

MacClade

This book is both a manual for the computer program MacClade, describing its features and potential uses, as well as a portrayal of a phylogenetic approach to studying diversity and evolution. It is relatively easy to see the diversity of living organisms, but it has proved more difficult to see that diversity in terms of its history; the slow development of a thoroughly phylogenetic perspective in biology attests to this challenge. Together this book and program present methods for analyzing and exploring phylogenetic hypotheses, including hypotheses about character evolution. MacClade is one attempt to give our mind's eye phylogenetic lenses, to help us think about and see lineages and evolution.

Some of the chapters and sections of this book (especially Chapters 3–6) discuss phylogenetic principles in general, and indeed the book could be read solely for its presentation of phylogenetic theory. Much of the rest of the book (especially Chapters 7–20) is devoted to MacClade itself. In this introductory chapter, we give a brief overview of MacClade's capabilities, followed by a summary of the chapters in this book, as well as information on citing MacClade, earlier versions of the program, future versions of MacClade, reporting bugs, and acquiring technical data about the program.

MacClade provides an interactive environment for exploring phylogeny. In MacClade's tree window, hypothesized phylogenetic trees or cladograms can be manipulated and character evolution visualized upon them. To manipulate the tree, tools are provided to move branches, reroot clades, create polytomies, and automatically search for more parsimonious trees. Character evolution is reconstructed on the tree and indicated by "painting" the branches. Alternative reconstructions of character evolution can be explored. Summaries of changes in all characters can be depicted on the tree. As trees are manipulated, MacClade updates statistics such as treelength and the results are depicted in graphics and charts.

Examining character evolution on a tree
in MacClade's tree window
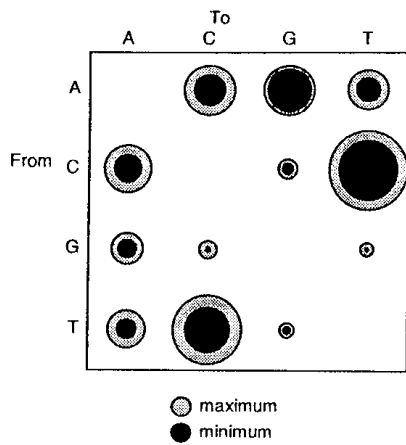
A chart showing the relative frequencies
of reconstructed changes between bases
on a tree, with area of circle proportional
to frequency of change

MacClade's data editor

MacClade provides charts summarizing various aspects of character evolution on one or more trees, as well as charts comparing two or more trees. For example, the charts can show the number of trees of each length, the number of characters on the tree with different consistency indices, and so on. It has a number of charts specifically designed for DNA/RNA sequence data, including one showing the number of changes on the tree at first, second, and third codon positions, and a chart of the relative frequencies of various transitions and transversions.

In MacClade's data editor, systematic and comparative data are entered and edited. The editor has numerous features for manipulating rows, columns, and blocks of data, and for recoding data. Assumptions about how character states transform into one another can be applied.

We will not detail MacClade's additional capabilities here. At the start of each of the Chapters 9 through 20 is a brief summary of the features described in the chapter. By browsing through these, you will get a more complete overview of MacClade's features.

If we had to admit a grand purpose to all of these features, it might be "to help biologists explore the relationships between data and hypotheses in phylogenetic biology". We envisage MacClade's use by biologists of many backgrounds. For example, suppose:

1. A systematist is working on the phylogenetic relationships of snail species. She enters data for 100 morphological characters and 20 taxa in MacClade's data editor, saves the file, and reads it into Swofford's PAUP. PAUP is used to find parsimonious trees, but to her surprise, the resulting trees separate two species she had thought closely related. By moving back to MacClade, examining alternative trees by hand, and using the charting functions, she discovers that the unexpected result is due to the influence of two characters of the nervous system. This not only provokes her to examine those characters in more detail, but also to gather molecular data. The molecular data set, on 60 taxa and 1500 nucleotides, is too large to get exact solutions from PAUP. PAUP's heuristic algorithms, combined with her ability to suggest and examine alternative trees in MacClade, convince her that the most-parsimonious trees for both data sets agree that the two species are not sister species. Although she had no intention of looking at fossil species, she realizes that this result suggests that their peculiar shared morphology might be primitive and old. By using the stratigraphic character type in MacClade, she discovers the minimum age of the ancestor in which the morphology was apparently derived was Paleocene, much older than the group of predators against which the morphology was thought to be a defense.

2. A population biologist has been studying the adaptive advantages of larval dispersal strategies. Although his past studies have fo-

cused on measuring risks and fecundities at two study sites in the Gulf of California, he realizes that his hypotheses could be tested by seeing if it accurately predicts a species' strategy according to its ecological position. After talking to a phylogenetic biologist, he reluctantly admits that his question is actually one of phylogenetic correlation between ecology and strategy. He finds a collaborator working on the phylogeny of the group, and together they use MacClade to map the evolution of larval dispersal strategies on the phylogenetic tree. He discovers that the correlation between strategy and ecology is not nearly as strong as he would have hoped, but he notices on MacClade's character tracing that the two groups on the phylogenetic tree with a special dispersal strategy also have the most species. His study shifts to an examination of the influence of dispersal strategy on speciation and extinction. After accumulating known phylogenies from numerous groups and exploring them with MacClade, a clear correlation emerges between speciation success and the special dispersal strategy. Itching to get back to his study sites, he discovers that his newly acquired familiarity with trees can be applied to his populations to examine patterns of dispersal and gene flow using reconstructed gene trees. He uses MacClade's random tree and random character generation to formulate null models to test his population hypotheses.

3. A molecular biologist studying cell surface proteins in a herbivorous insect suspects that a particular domain might be involved in binding secondary compounds made by the host plants. A direct chemical approach to demonstrate binding proves difficult, but the genes coding for this protein have been sequenced in several dozen related species. Using MacClade to help reconstruct the phylogeny of the species and to chart how evolutionary changes of amino acids are distributed along the length of the protein, she finds that this domain is indeed the most rapidly evolving. Furthermore, by comparing the phylogenies of the insect and the plant host using Page's COMPONENT, she makes a convincing case that one of the insect groups arose with a shift to a new plant host that had powerful and diverse secondary compounds. Not only did MacClade reconstruct many amino acid changes along that branch of the insect phylogenetic tree, but it also showed that changes in other amino acids were concentrated in the clade of insects living on the noxious hosts. With such evidence that the domain was evolving in response to host secondary compounds, the molecular biologist faced with renewed vigor the direct chemical approach to examining the binding.

4. An evolutionary biologist teaching a basic evolution course wants to introduce students to phylogenetic reconstruction. In a live demonstration MacClade is used first to help students picture phylogenetic trees in their minds' eyes. Then, the concept of fit between a tree and a character is shown via MacClade's character tracing.

Students are given an example data set, and let loose. Soon they are moving branches around with increasing rapidity, treating the search for parsimonious trees like a video game, competing against one another for shorter trees. In the process they discover that grouping by derived similarity will gain them shorter trees. One of them notices that one character disagrees with most of the remaining characters on whatever tree they can find. Indeed, on the parsimonious trees this character shows many cases of convergence. Another character is found that is similarly convergent, and students discuss what evolutionary processes might lead to the two characters having correlated evolutionary changes.

These scenarios illustrate the varied uses of a phylogenetic perspective, and of MacClade in making it accessible. We hope that whether or not you use MacClade, you will find in reading this book a new appreciation for the importance of phylogenetic studies.

Learning how to use MacClade

MacClade is a large and complex program. If you are just getting started with MacClade, we advise you to follow through the examples in Chapter 2, "A Tutorial Overview of MacClade". This will guide you step by step through simple data files, and thereby introduce you to some of the basic functions of MacClade.

Chapters 7 through 20 of this book will provide more detailed guidance. Although we hope that many of MacClade's features will be obvious from the program itself, some important ones will not. So that you may discover these features, and thoroughly understand MacClade's analyses, we strongly recommend that you at least glance through Chapters 7 through 20.

Part II of the book, comprising Chapters 3 through 6, concerns phylogenetic theory. "A Phylogenetic Perspective" outlines how phylogenetic analysis, such as that provided by MacClade, can be of value in answering various biological questions. "Introduction to Phylogenetic Inference" discusses methods of inference of phylogenies, and of patterns and processes of character evolution. "Reconstructing Character Evolution Using Parsimony" describes the assumptions and methods MacClade uses in reconstructing character evolution. A chapter by Daniel Fisher, "Stratigraphic Parsimony", briefly describes the principles of stratocladistics. We recommend that you read the relevant sections of these chapters before using MacClade to study character evolution.

Part III contains fourteen chapters (7 through 20) that provide detailed information on using MacClade.

Chapters 7 through 9 introduce the basics of using MacClade and of file management. "**Using MacClade on your Macintosh**" gives an overview of hardware and software requirements, installation, memory management, the help facility, and setting preferences. "**Managing Data Files**" tells how to create, open, and save data files. "**Importing and Exporting Text Files**" describes how to share files with HENNIG86, PHYLIP, and other programs.

Detailed information about manipulating data, assumptions, and trees is provided in Chapters 10 through 12. "**The Data Editor**" tells how to use MacClade's data editor to enter and edit data. "**Assumptions About Characters**" details the various assumptions used in MacClade's calculations, such as weighting of characters and how states transform one to another. "**Trees and Tree Manipulation**" introduces MacClade's tree window, and how trees can be manipulated, stored, and retrieved.

Chapters 13 through 18 focus on analysis of character evolution and the relationship between the character data and the tree. In "**Tracing Character Evolution**", methods for reconstructing the ancestral states (and thereby pathways of character evolution) are discussed. In "**Basic Tree and Character Statistics**", simple statistics of treelength and indices such as the consistency index are outlined. MacClade's charting facility is described in "**Charting Tree and Character Statistics**". Use of continuous-valued characters is explained in "**Continuous Characters**". The next chapter discusses "**Patterns of Correlated Character Evolution**", particularly the concentrated-changes test for phylogenetic character correlation. MacClade's randomization facilities, which can be used in various statistical tests, are discussed in "**Generating Random Data and Random Trees**".

Several of MacClade's features specific to molecular data are described under "**Using MacClade with Molecular Data**". In this chapter we also give some examples of MacClade's applications to molecular data.

The chapter "**Recording Your Work: Printing, Graphics, and Text**" gives strategies for keeping records. Printing facilities are described there.

The remaining sections of the book include an appendix on file format, an appendix on using MacClade and PAUP together, and Indices.

Conventions

Throughout this book, menu items and dialog boxes will be shown in **boldface** type. Buttons in dialog boxes will be shown in the font Helvetica. Important comments will be flagged with labels, as follows:

WARNING: *Important warnings are indicated in this fashion.*

NOTE: *Notes, shortcuts or tips are indicated in this fashion.*

EXAMPLE: *Exercises using the Example files included with MacClade are indicated in this fashion.*

How to cite MacClade

The computer program MacClade is not just a calculating tool. As with any published paper, it contains new ideas, implicit suggestions, and new methods. MacClade is distinct from works published on paper in that new ideas and methods presented are in a form that can be immediately used by researchers and students; it responds actively to a user instead of passively to a reader. Use of ideas in this book or in MacClade must be acknowledged as with any other published paper. (This may seem obvious; we mention it only as history suggests that it is not.) MacClade should also be cited in your Materials and Methods as a source of the calculations or output used in your papers.

Despite the scientific content of MacClade and other phylogenetic programs such as HENNIG86 (Farris, 1988), PAUP (Swofford, 1991) and PHYLIP (Felsenstein, 1991), some journal editors may object to placing a citation to a computer program in the Literature Cited. However, as the 21st century approaches, we think this attitude will become rarer. We recommend the following citation, which can be used both for the program and this accompanying book:

Maddison, W. P. and D. R. Maddison, 1992. MacClade: Analysis of phylogeny and character evolution. Version 3.0. Sinauer Associates, Sunderland, Massachusetts.

Earlier versions of MacClade

Previous versions of MacClade are numbered 1.0, 1.01, and 2.1. Any other version of MacClade, such as 2.65, 2.87, 2.97, 2.99, 3 Beta 9.9, were pre-release test versions that we showed at meetings or distributed to a few colleagues to test for us. (The easiest way to determine if a copy is a test version is to choose **About MacClade** under the menu. If "Release Version" appears in the dialog box, or the version number is 1.0, 1.01, or 2.1, it is not a test version.) Test versions were distributed *for testing only*, under the condition that no results be published based on those versions, as they have bugs in many parts of the program, including the algorithms that give basic results such as tracing of character evolution and treelengths. Unfortunately, copies of the test versions have been circulated, and some people may have test copies without realizing it.

Thus we repeat: DO NOT USE RESULTS FROM ANY TEST VERSION IN PUBLICATIONS. Except when we have indicated otherwise, results from test versions should not be trusted. We expect that any analy-

ses done with test versions will be redone with the officially released version 3 of MacClade.

Differences between MacClade 3 and 2.1

MacClade version 3 is different in many ways from version 2.1. Most striking, perhaps, is the spreadsheet editor for entering data, which allows an intuitive means of row and column editing, copy/pasting, adding footnotes to any cell in the matrix, and so on. The raw text files of MacClade 3 follow a different format than those of earlier versions. This new format, called NEXUS, is also the format used by version 3 of PAUP. MacClade 3 automatically saves files in the correct format. This allows you to enter data without worrying about details of where semicolons and periods and the like need to be placed in the NEXUS file.

MacClade 3 can read data files in the formats used by versions 1.0, 1.01, and 2.1 (but it may not understand the formats used by test versions of MacClade).

The 80 taxa by 200 character limit on earlier versions of MacClade has been increased significantly (to 1,500 taxa and 16,000 characters).

Editing of assumptions has been made much easier, in part through a graphic dialog box for editing user-defined character types as matrices or character state trees. Trees can now have polytomies, and can be edited with a more powerful, more graphic set of tools. Trees no longer have to include all taxa in the data matrix, characters can be traced in color, the old Show Changes summary mode has been expanded and unlocked. A charting facility has been added that generates various tables or bar charts to summarize characters or trees. Publication-quality trees can be produced with choice of fonts, angle of taxon names, tree shapes, and so on. Many other features have changed; if you are familiar with version 2.1, and version 3 behaves unexpectedly, then check the rest of this section, and in the remainder of this book.

Version 3, in contrast with version 2.1, maintains the data file, including the user-defined types, type sets, weight sets, trees, and character names, in the computer's memory (RAM) while you move between the data editor and tree window. In version 2.1 the editor and tree window were like separate programs, and in transferring to the tree window, the program would reread the file from the disk.

MacClade no longer has built-in fonts with male/female symbols. Please write to your favorite typeface company to lobby for biological symbols.

Differences in calculations or assumptions

MacClade 3 will give results different from version 2.1 in some circumstances.

MacClade 3 treats Dollo characters differently from version 2.1; if two states of a Dollo character might otherwise be placed at the root with equal parsimony, MacClade 3 will place both states at the root, whereas version 2.1 will place only the lower-numbered state there.

User-defined type characters with polymorphisms of more than two states are treated differently by MacClade 3, which will use the lowest number of steps from one state to any other state as the number of steps required within a terminal taxon should the first state be reconstructed at the taxon's base.

The type called "reversible" no longer exists in version 3. This type was used instead of ordered and unordered for binary characters in MacClade 2.1 because there is no useful distinction between ordered and unordered types for binary characters. However, the reversible type can be dispensed with if one realizes that calling a binary character either ordered or unordered gives the same results. In version 3 it is expected that you will use unordered or ordered for binary characters.

Future plans for MacClade

Our plans and your suggestions

Among the planned improvements for future versions of MacClade are:
— Modular and extensible program format (see below)
— Full incorporation of continuous-valued characters into the program, with editing in a spreadsheet data editor
— Better tools for sequence alignment
— More ways to display the tree in the tree window
— Better interface for charting facilities and tree printing
— Likelihood and other statistical techniques
— A better editor for character state symbols and names
— Better interface for getting and storing trees and tree files
— Full System 7 support

We would *very* much like suggestions from you, whether they concern user-interface features (e.g., tree appearance on screen), or new sorts of calculations (e.g., a new correlation test, new charts).

Adding your own calculations

One of our long-term goals is to make MacClade capable of being expanded by the user. We cannot hope to incorporate all of the phyloge-

netic calculations that have been and will be developed, but we can build means by which new calculations can be added as modules to MacClade. In the future this may involve the use of System 7's Apple Events to pass information from MacClade to stand-alone modules.

At present, MacClade has some primitive tools that allow access to some of MacClade's memory structures. These allow you to write programs that run under Multifinder or System 7 alongside MacClade and do calculations using the same tree and data that MacClade is using. For details of how to do this, please contact us.

Reporting bugs

We have extensively checked and tested the main algorithms in MacClade in an attempt to ensure that no faulty results will be generated; in various parts of this book we have described our checking procedures. We hope that no such bugs remain, but there will remain, we are sure, a number of bugs having to do with screen display and occasional crashes. If you think you have found a bug in MacClade, please report it to us.

To correct a bug, we need the following information:

1. A description of the problem. This includes a list of *all* the steps that led to the bug. If at all possible, try to repeat the problem yourself until you know the exact sequence of steps that reliably lead to the bug. Without knowing an exact sequence of events that reproduces the bug, there is a good chance we cannot find the bug and fix it.

2. We also may need an *exact* copy of the data files or tree files involved, as many bugs appear only with specific files. Therefore, a copy of the files should be saved. (If we do borrow your files, they will be kept strictly confidential, and will be discarded once the bug is fixed.)

3. Your name and address (e-mail if possible).

4. Exact version number of MacClade used (see **About MacClade** in the menu).

5. Model of Macintosh.

6. Amount of memory (RAM) in your Macintosh.

7. Version number of system (see **About Finder** or **About this Macintosh** in the menu when in Finder).

8. If not using System 7, were you using Multifinder?

9. If using Multifinder or System 7, how much memory was allocated to MacClade (determined by selecting MacClade in the Finder and choosing **Get Info**)?

10. Any other possibly relevant information about data file format (such as standard, extended, DNA, etc.), number of taxa and characters, type of printer, and so on.

To report a bug, send the above information to one of us, Attention: Bug Report (our addresses are listed in the Preface).

You may find that you are not sure whether you have discovered a bug or simply don't understand how MacClade is supposed to function. We would like to hear from you even if the problem is only that MacClade's intended behavior was not made clear to you. We suggest that you first reread the appropriate sections of this book to see how MacClade was intended to behave.

Technical information and availability of source code

MacClade 3 was developed and compiled using Symantec's Think Pascal™ version 4.01, except for some sections in assembly language.

If you would like more technical information about MacClade, some details about the functioning of MacClade's algorithms are given in Chapter 5 and the various other chapters concerning them. If you want further information on the algorithms or other technical details, please contact the authors. The source code for those parts of the program that deal directly with the calculations of results are available from us, so that they may be examined by the scientific community.

2. A Tutorial Overview of MacClade

As an introduction to MacClade and some of its capabilities, we offer the following tutorial. It will guide you step-by-step through some simple data files. (We will assume basic familiarity with the operation of a Macintosh; if you need help with basic operations, see your Macintosh manual.) Those already familiar with MacClade may want to skip this chapter.

As you work through the tutorial, remember that MacClade 3 has four main windows. The *Data Editor* is an editor in which the taxa are named, and their states in various characters are entered. In the *Tree Window* trees are built and modified, and their relationship to the character data explored. You can move between the data editor and tree window using items in the Display menu. MacClade cannot show both windows at the same time; thus at any given time MacClade is either in data editing mode or tree analysis mode. The *Character Status Window*, requested in the Display menu, lists the characters in the data file, and various assumptions about each character (its weight, assumptions about transformations from one state to another, and whether the character is currently included or excluded). This window not only lists these assumptions, but also allows you to change them. The character status window also lists statistics relating the characters to the current tree (e.g., their consistency indices) when the tree window is in use. When the tree window is on the screen, the *Chart Window* can be requested from the Chart menu to show various summary statistics about the characters and their evolution according to the tree. By exploring each of these windows you will introduce yourself to many of MacClade's features.

Tree manipulation

Vertebrates

Open the MacClade Examples folder. Double-click on the Vertebrates data file to start MacClade. The data file will take you straight to the tree window, and you will be presented with a phylogeny on the screen, as in the following figure. (If you are presented with a data matrix instead, the file has been altered — try selecting **Go To Tree Window** from the **Display** menu.)
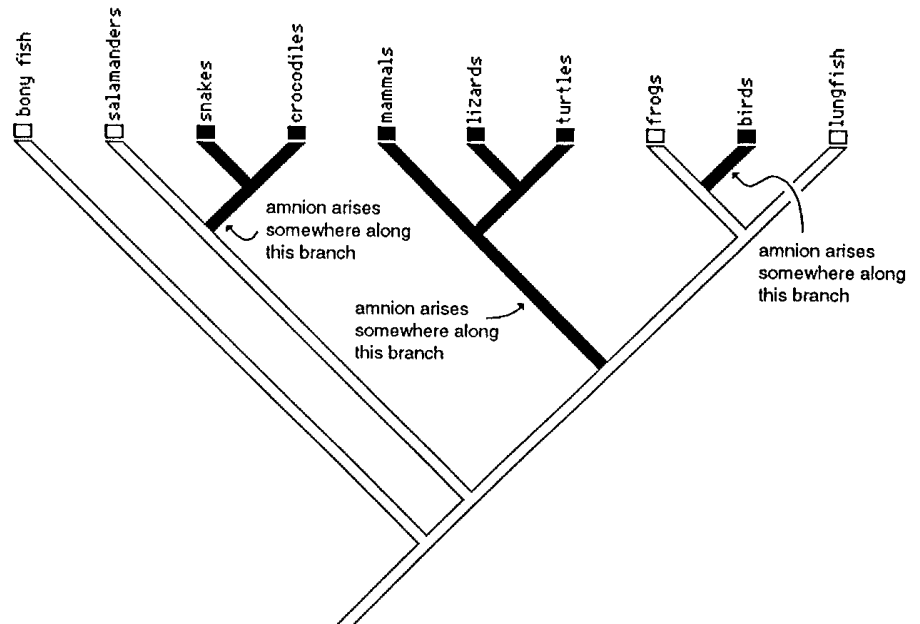
13
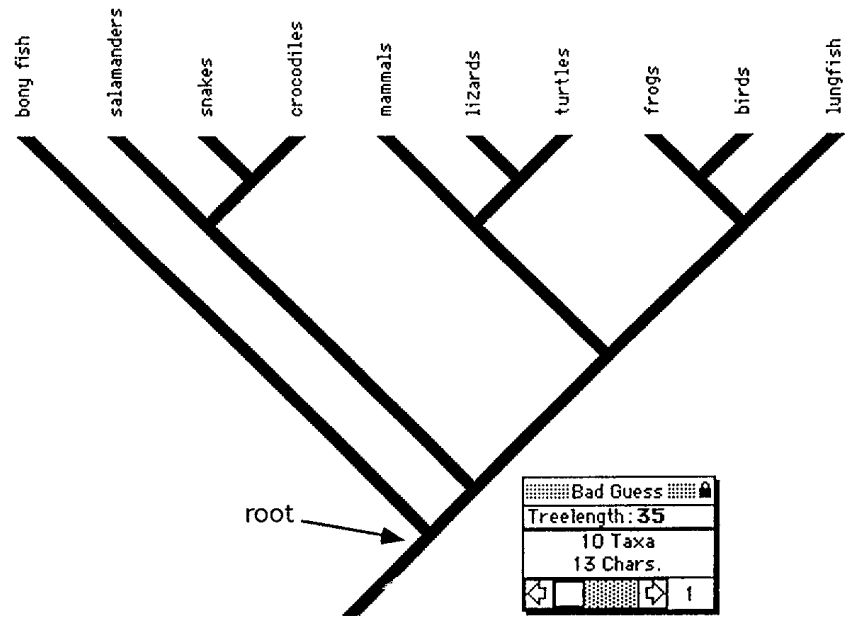
More accurately, what you see is one hypothesis about how the ancestral vertebrate lineage split and evolved, leading up to the major living vertebrate groups. As its name ("Bad Guess") implies, it is almost certainly wrong. The branch at the bottom of the tree, representing the ancestor of the taxa, is called the **root** of the tree.

To see how the traits of vertebrates have evolved according to this hypothesis, choose **Trace Character** from the **Trace** menu. You should now see the evolution of the amnion traced onto the screen.
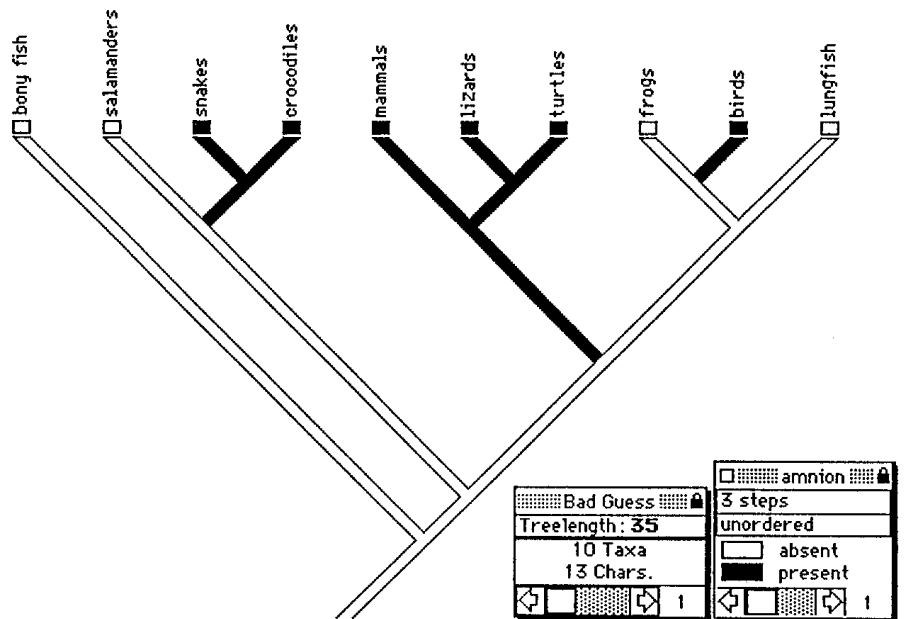
Look at the legend in the lower right of the tree window. Note that the character amnion has two states: "absent" or "present". According to this phylogeny, the amnion (portions of the tree shaded in black) arose three times. This is also indicated in the legend at the far right, where it indicates that there were "3 steps" of evolution in this character. (If you have a color monitor, then the states might not be symbolized by white and black, but by colors, most likely yellow and blue.) These three arisals are shown in the diagram below:

bony fish salamanders snakes crocodiles mammals lizards turtles frogs birds lungfish

amnion arises somewhere along this branch

amnion arises somewhere along this branch

amnion arises somewhere along this branch

That the amnion arose three times independently (and thus that the amnions of snakes, mammals, and birds are **analogous**) seems unreasonable, as it is a complex explanation of the observations, and a simpler explanation is available. It seems more reasonable to presume that the amnion arose only once, in an ancestral tetrapod vertebrate, and that it was passed on to descendants of that ancestor (and thus that the amnions of snakes, mammals, and birds are **homologous**, having been derived from a common ancestor).

N O T E: *Here MacClade shows the amnion arising at the start of a branch. MacClade does this for convenience only. The evolution of the trait might have happened anywhere along the history of the branch.*

Now, let's alter the phylogenetic hypothesis slightly. Move the snakes + crocodiles over so that they now are next to the mammals + lizards + turtles. To do this, move the mouse's arrow pointer over to the branch just below the snakes + crocodiles, click down on the mouse, and, while holding the mouse button down, drag the diamond over to the branch just below mammals + lizards + turtles:
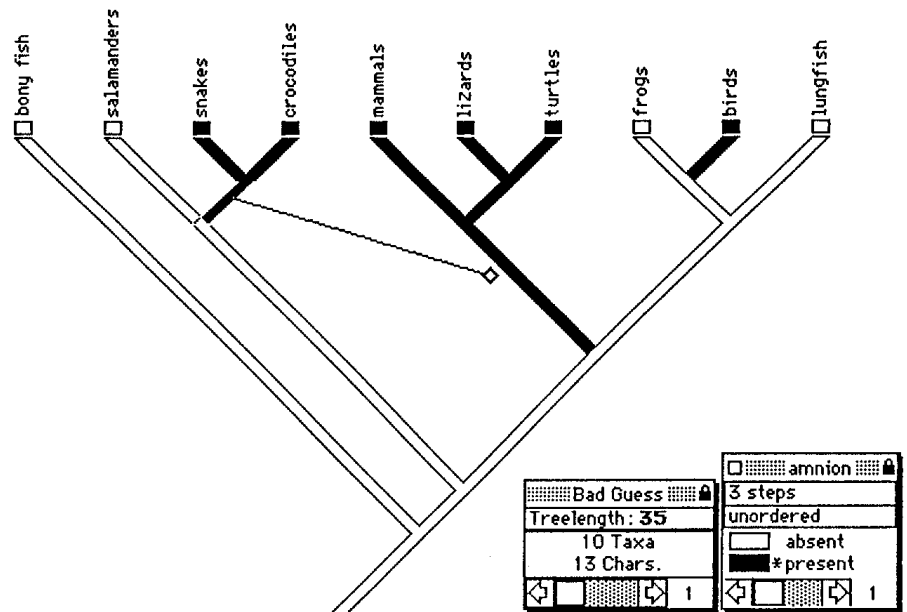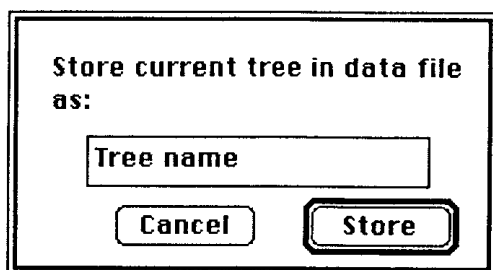
When the mouse is over the latter branch (this you can tell as the branch will change shades), let go of the mouse button. The snakes and crocodiles should have moved over beside the mammals + lizards + turtles.
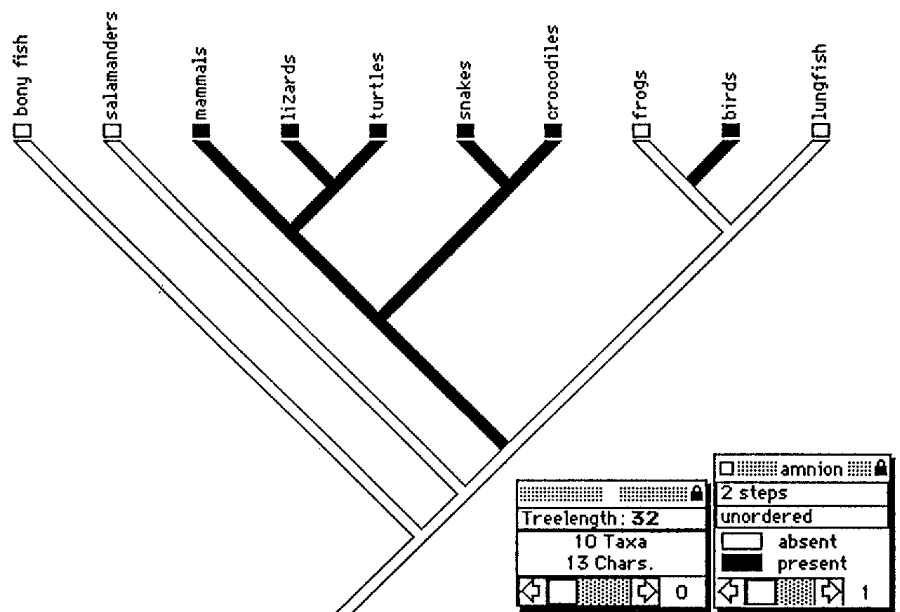
According to this new hypothesis, the amnion arose only twice; that's better, as it is a simpler explanation.

In the second box from the right is a line that reads "Treelength: 32". The **treelength** is a statistic that gives an indication as to how well the tree fits the data. The longer the treelength, the worse the fit; the shorter the treelength, the better the fit. The treelength is simply the sum, over all characters, of the number of evolutionary steps in each character on the tree. For the tree on the screen, the number of steps in the first character is 2; the number of steps in the second character is 3 (to scroll to see the second character, click once on the lower, right-hand arrow in the character legend at the far right); in the third character, 5; and so on. The treelength is thus $2 + 3 + 5 + \cdots = 32$. Shorter treelengths mean that, in general, the tree fits the characters well, that is, that the tree does not require one to presume multiple independent arisals of traits.
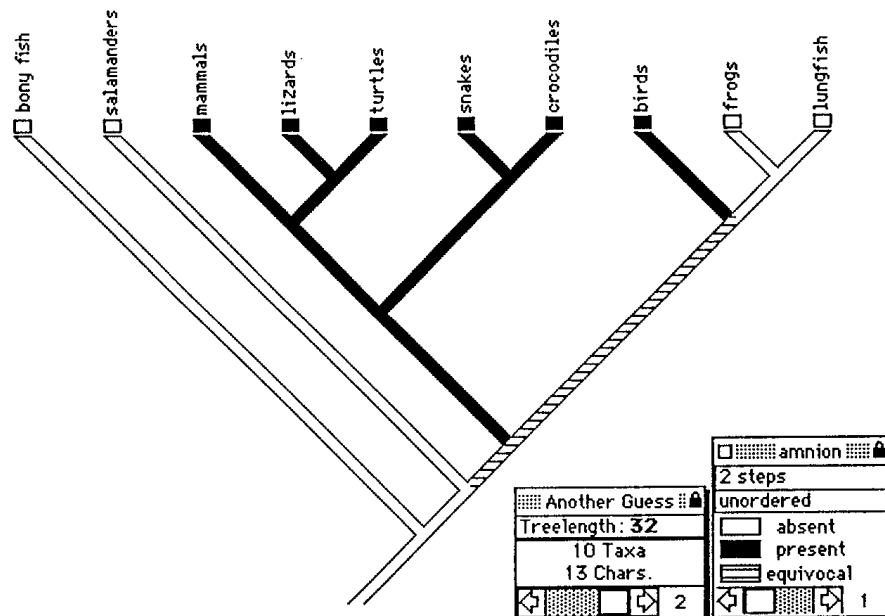
Once you have found a tree you would like to save, store it in the file with the **Store Tree** command of the **Trees** menu. In the dialog box that is presented to you, enter a name for the tree and press **Store**. You can re-examine the stored tree when you want using the **Trees** command. (You may want to save the modified file, with its new stored tree, to disk so that you can see your tree the next time you open the Vertebrates file. You would do this using the **Save File** command of the **File** menu. However, if you do save the file to disk, it will no longer be in its original condition, making it more difficult for others to follow this tutorial.)

To see other options for tree manipulation, select the **Tools** menu. You will notice the menu is a palette with various small icons, which represent tools for your use. Choose the tool that looks like this: ⌐┓ , then click this tool on a branch in the tree. The tree will be rerooted where you touched it (unless you touched the tree at or near the root, in which case rerooting wouldn't make any difference).

Try some other tree manipulation tools. ⫟ will ask MacClade to try rearranging branches in the clade above the branch touched, in order to search for alternative trees that are more parsimonious according to the data set. ↘⟋ will destroy the branch touched and so yield a polytomy from a dichotomy. Play with the other tools to see what they do. MacClade lists the function of the tool in the small window at the bottom of the palette. See Chapter 12 for a more detailed description of these tools.

Once you have finished exploring the tree manipulation tools, use the **Trees** dialog box in the **Tree** menu to choose the "Another Guess" tree. Trace character 1 onto the tree by scrolling to it using the arrows in the character legend. You should see the following:

Store current tree in data file as:

Tree name

Cancel   Store

Dialog box for naming and storing trees

Of course, if you were really serious about examining the evolution of this character, you would do it with a phylogenetic tree that was well supported by available evidence (clearly this one is not!). But this example is just for practice, so bear with us.
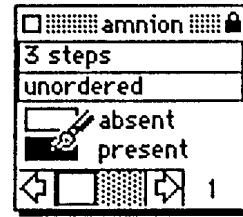
For the branches with horizontal striping (the "equivocal" pattern), the evidence is ambiguous as to how the characteristics evolved. The ancestors along these branches might have had amnions, but they might not have; either possibility is equally parsimonious. When you click on a striped branch (do this now), MacClade indicates this uncertainty by placing a question mark in the legend box on the right beside each state that could be along that branch.

In this case, there are two possible patterns of character evolution; that is, two patterns that equally simply explain the distribution of the amnion in living vertebrates. They are:

1. The amnion arose independently in the mammals-to-crocodiles clade and the birds;

2. The amnion arose only once, but was later lost in frogs and lungfish.

To see these two possibilities, you can use one of the several tools MacClade provides for manipulating and examining the tracing. These are stored in the palette in the **Tools** menu. Select the paintbrush tool

(✐). Click the paintbrush tool on the box in the legend to the left of "present", as shown in the following figure.

```
┌─────────────────────────┐
│ □ ▒▒▒▒ amnion ▒▒▒ 🔒 │
├─────────────────────────┤
│ 3 steps                 │
├─────────────────────────┤
│ unordered               │
├─────────────────────────┤
│ ☐⁄ absent              │
│ ■ present              │
├─────────────────────────┤
│ ◁ ☐ ▒▒ ▷  1           │
└─────────────────────────┘
```

The paintbrush should turn black (✎), indicating that it "contains" your selected state. (If you have a color monitor, and colors are used to indicate the states, then the paintbrush will turn the color of the state; otherwise, a filled paintbrush will always be black, an empty one white.) Now click on one of the equivocal branches. In doing this, you are fixing the state of the branch; you are telling MacClade: "Show me the evolution of the amnion assuming that organisms along the branch I touched had an amnion". You should see the amnion arising once and being lost once. Now, undo what you just did by choosing **Unfix All** from the **Trace** menu. Next, click the paintbrush tool on the box in the legend to the left of "absent". Now, when you click the paintbrush on the same branch, you should see the amnion arising twice.

There are many other features to help you analyze character evolution which we will introduce in other chapters (see especially Chapter 13).

Entering data

Now let's examine the data that you have been analyzing. Go to the **Display** menu and choose **Go to Data Editor**. This will take you to a data editor that should look something like this:

		1 amnion	2 appendages	3 body covering	4 thermoreg.	
1	ray finned fish	absent	fins	derm.scales	poikilotherm	⬆
2	frogs	absent	legs only	smooth	poikilotherm	
3	turtles	present	legs only	epid.scales	poikilotherm	
4	lungfish	absent	fins	derm.scales	poikilotherm	
5	salamanders	absent	legs only	smooth	poikilotherm	
6	crocodiles	present	legs only	epid.scales	poikilotherm	
7	lizards	present	legs only	epid.scales	poikilotherm	
8	birds	present	legs&wings	feathers	homeotherm	
9	mammals	present	s only/legs&wi	hair	homeotherm	
10	snakes	present	?	epid.scales	poikilotherm	
11						
12						
13						⬇

Uertebrates

On the left of each row are the names of the taxa. Along the top of each column are the names of the characters, and in the matrix itself are the

names of the character states possessed by each of the taxa for each of the characters. If you want to change the data, click on the cell of the matrix you want to edit, and type in the new data. For instance, click on the cell in the first column (amnion) beside "salamanders" and type "present" if you want to claim that salamanders have an amnion.

Now let us make a new data file. In the **File** menu choose **Close File.** If MacClade asks you "Do you want to save changes before closing?", respond "No". MacClade will then present you with a list of files you might open. Instead of opening any of them, choose the New button to start a new data file. You will be presented with a blank data matrix with one taxon and one character:
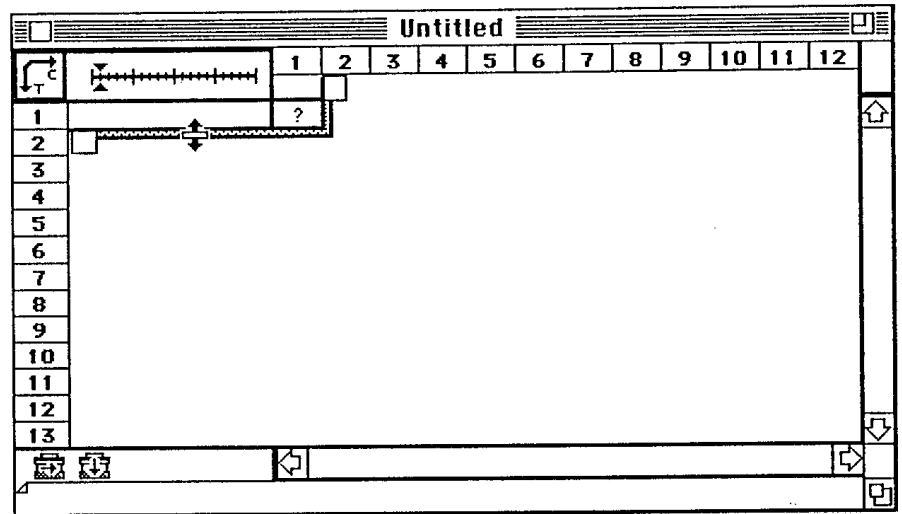
The thick gray border with a box at each end represents the edge of the matrix. To enlarge the matrix so as to add taxa, place the cursor over the bottom edge of this heavy border, as shown above (note the ✛ cursor), click and drag down. When you let go, MacClade will add more taxa. To make new characters, place the cursor over the right-hand heavy border, click and drag to the right. Add at least 5 characters, and at least 3 taxa.

Now fill in taxon names in the wide column at the left hand side. To enter the name of the first taxon, click on the cell beside the number 1, and type in the name. Continue for the other taxon names. Now enter character state data in the body of the matrix. Enter character state data as 0's, 1's, 2's, and so on. For now, we will name the states 0, 1, 2, and so on. If you want to use names like "absent", "present", "red", "blue", and so on in the matrix, refer to Chapter 10 to learn how to prepare the data file for entering character state data by full names, as we did in the Vertebrates example, or for using other symbols such as ACGT. Chapter 10 also explains many features of the editor not included in this Tutorial.

Specifying assumptions

| ▼ Character▼	Type▼	Weight▼	States
☑ 1.	unordered	1	2
☑ 2.	unordered	1	2
☑ 3.	unordered	1	2
☑ 4.	unordered	1	2
☑ 5.	unordered	1	2
☑ 6.	unordered	1	2
☑ 7.	unordered	1	2

Selecting character 3

Once you have typed some data into the new matrix, you can set the assumptions used for the characters. Choose **Character Status** from the **Display** menu. In this window are listed the characters and information about them. To practice changing assumptions, click on the row for character 3, then go to the **Change Type** submenu in the **Assume** menu and select **Ordered**. This will change the type of character 3 from unordered to ordered. (The difference between an unordered and ordered character is described in Chapter 5.) Now try excluding character 5 from subsequent analyses. To do this, select character 5 in this window. Then choose **Exclude** from the **Change Inc Exc** submenu in the **Assume** menu. The check mark on the left will change to an x and the character type will be listed as "(excluded)". Character 5 is now excluded from any tree analysis.

Now try going to the tree window (using the **Go To Tree Window** menu in the **Display** menu) and playing around with trees. (Don't forget to give names to all of the taxa first!)

Charting results
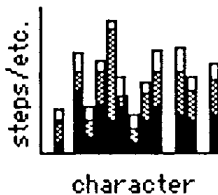
Lower-middle icon of
Character Steps/etc. dialog box
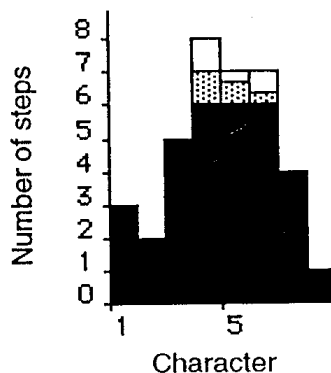
Lower-left corner of the chart

We can now briefly explore some of MacClade's charting facilities.

Close the current file, and open the file Amblygnathus in the Morphological Examples folder. This file contains morphological data on a group of ground beetles (Ball and Maddison, 1987). Go to the tree window, and choose **Character Steps/etc.** from the **Chart** menu. In the dialog box that appears, there are six large icons; touch on the lower-middle icon.

When you press **Chart**, MacClade will present you with a chart summarizing the number of steps in each character over the 21 most-parsimonious trees that are stored in the data file. The 28 characters will appear along the horizontal axis, the number of steps along the vertical axis.

In the chart that appears, most of the bars are solidly filled (with black or a color, depending on your monitor), indicating that those characters have the same number of steps on all 21 trees examined. For seven of the characters (4, 5, 6, 19, 21, 22, and 27), however, part of each bar is solid, part is grayed, and the top is white. The number of steps in these characters varies over the 21 trees. The solid part of the bar indicates the minimal number of steps across trees, the gray part the average number, and the white part the maximum number. For example, the chart indicates that character 4 has between 6–8 steps, with an average number of 7 steps.