Opinion TRENDS in Ecology and Evolution Vol.20 No.3 March 2005

Do early branching lineages signify ancestral traits?

Michael D. Crisp and Lyn G. Cook

School of Botany and Zoology, The Australian National University, Canberra, ACT 0200, Australia

A reverence for ancestors that has pre-occupied humans since time immemorial persists to the present. Reconstructing ancestry is the focus of many biological studies but failure to distinguish between present-day descendants and long-dead ancestors has led to incorrect interpretation of phylogenetic trees. This has resulted in erroneous reconstruction of traits such as morphology and ancestral areas. Misinterpretation becomes evident when authors use the terms 'basal' or 'early diverging' to refer to extant taxa. Here, we discuss the correct interpretation of trees and methods for reconstructing the ancestral features of organisms using recently developed statistical models. These models can be inaccurate unless they use information that is independent of phylogenies, such as genetics, molecular and developmental biology, functional morphology, geological and climatic processes, and the fossil record.

Introduction

Over the past ten years, there has been an increase in the number of studies using phylogenies to tackle questions of evolutionary biology, such as: what were the characteristics of extinct ancestors [1–4]?; how do lineages co-evolve [5,6]?; and when, where and why did lineages radiate [7–10]?. For example, phylogenetic analysis has shown that, unlikely as it seems, compound eyes have evolved many times in different arthropod lineages (i.e. nearly identical complex structures originated independently) [4]. However, there has also been an increasing trend of incorrect interpretation of phylogenetic trees [11]. This is not a trivial problem because it has led to false inferences of ancestry, such as 'To catch a glimpse of the very beginnings of the mammalian dynasty, we must travel to Australia. I'm looking for one of the most ancient of all mammals. It's so ancient, it shares at least one characteristic with the reptiles.' [David Attenborough, *The Life of Mammals*, Episode 1 (BBC, 2003)], which is referring to monotremes (the Australasian platypus and echidnas) (Box 1). Perhaps this points to problems in how phylogenetics are taught and assimilated at the undergraduate and graduate levels, and even beyond, when researchers come into phylogenetics from another field, such as molecular biology or ecology. Our aim here is to point out examples where interpretation of trees has been erroneous, to explain the problem and to show how conclusions are changed when phylogenies are interpreted correctly.

Interpreting a phylogenetic tree

Each branch point of a fully bifurcating tree results in two lineages; these are sister groups because they originate simultaneously from their shared ancestral lineage. If both lineages survive to the present, they are the same age and all extant species in both are equally removed in evolutionary time from their nearest common ancestor, the shared ancestral lineage or stem ('sister group rule') [12]. Once two lineages have separated, each evolves new characters independently of the other and, with time, each will show a mixture of plesiomorphic (see Glossary) and apomorphic character states. Therefore, extant species in both lineages resemble, to varying degrees, their common ancestor. Consequently, whereas character states can be relatively ancestral (plesiomorphic) or derived (apomorphic), these concepts are nonsensical when applied to whole organisms.

Glossary

Apomorphic: the more derived (more recently evolved) of two character states (contrast with plesiomorphic).
Dispersal: the establishment of a population in a new area, across what is usually a barrier to gene flow; normally assumed to be a rare, stochastic event ('jump dispersal'), and therefore not creating a congruent pattern in multiple lineages (in contrast to vicariance). However, dispersal can be coordinated in direction, driven by prevailing forces such as wind or water currents, though not necessarily simultaneous. Concerted dispersal is coordinated in time and space.
Hypothetical ancestors: in deep evolutionary time ancestry is always hypothetical because it cannot be verified by observation. Internal branches of a phylogeny represent inferred (i.e. hypothetical) ancestors. Terminal branches represent extant (observed) taxa and are not ancestors.
Ladderized: an unbalanced tree drawn showing the species-poor sister group on the same side at every node (Figure 1c; main text). The same tree appears different when drawn unladderized (Figure 1d; main text).
Node: a branch-point in a phylogeny, inferred to be a speciation event giving rise to sister groups.
Optimization: use of a model (examples in Table 1) to provide a 'best fit' of character states to the tree in order to infer ancestral states at internal nodes.
Plesiomorphic: the more ancestral (earlier evolved) of two character states (contrast with apomorphic).
Tree balance: a balanced tree has an equal number of species in every sister group pair (Figure 1b; main text). An unbalanced tree has sister groups differing in species richness at many nodes (Figure 1c, d; main text). Unbalanced trees are typically drawn in a ladderized configuration (Figure 1c).
Vicariance: the splitting of an area (e.g. by climatic or geological events). An expected consequence is the splitting of populations of multiple species contemporaneously, leading to near simultaneous allopatric speciation in multiple lineages. A pair of areas separated by a single vicariance event is expected to have multiple pairs of sister groups of the same age.

Corresponding author: Crisp, M.D. (mike.crisp@anu.edu.au).
Available online 13 December 2004

 0169-5347/$ - see front matter doi:10.1016/j.tree.2004.11.010

Box 1. The monotreme example explained

The 'discovery' of the platypus by Europeans 200 years ago provoked outrage over what was thought to be a hoax and monotremes have remained an enigma ever since. The orthodox view is that monotremes are survivors of ancient Mesozoic ancestors and are not closely related to other mammals [30]. Moreover, they have been considered to be a living museum of ancestral 'reptilian' features, such as egg laying. With only three living species, they are a classic example of naïve misinterpretation of evolutionary history. However, recent studies of monotremes have overturned both fallacies.

Molecular phylogenetics has shown that the three major lineages of living mammals (monotremes, marsupials and placentals) diverged so closely in time (no more than 20 million years apart, probably between 180 and 140 million years ago) that it has been difficult to determine which two are more closely related [30,31]. The relationship is so close that some analyses (e.g. with the use of mitochondrial genes [32]) support the 'Marsupionta' hypothesis, in which monotremes and marsupials are sister groups, with placentals more distantly related. However, these results were probably influenced by uncorrected biases in the data. Analyses that corrected these biases retrieved the conventional 'Therian' phylogeny, which groups placentals with marsupials, with monotremes being sister to these [31].

If a mammalian phylogeny had been constructed early in the Cretaceous, at a time when the monotremes were relatively diverse [30], the naïve perception of the phylogeny might have been reversed, with the therians appearing 'basal' (*cf.* Figure 1e–g main text). Discussion of the evolution of features might then have focussed on the reptilian features of placentals and the advanced features of monotremes. The reality is that mammals are a monophyletic group and, therefore, are all equidistant from their common ancestor. All mammals, including monotremes, are a 'patchwork of reptile, mammal and unique features' [33]. Monotremes share at least six uniquely derived morphological, chromosomal and developmental features and, for each of these, their sister group, the therians, have the plesiomorphic state found in the common mammalian ancestor [30,31,33]. For example, monotremes have nine or ten sex chromosomes that form complex chains at meiosis [34]. This feature is otherwise unknown in vertebrates.

An internal node of a tree can be referred to as being 'basal' or 'derived' in relation to another node, because internal nodes are hypothetical ancestors. A node is 'basal' if it is closer to the root than to the 'derived' node along any given path through the tree between the root and a terminal branch (Figure 1a).

A common misinterpretation of trees

Commonly, one lineage of a sister group pair has fewer extant taxa than does the other and we refer to such cases as the species-poor sister group and the species-rich sister group, respectively. A misinterpretation of trees occurs when the species-poor sister group is thought of as 'basal' or 'early diverging' with respect to its species-rich sister. Conversely, the species-rich sister group is misconceived as 'derived' or 'advanced' with respect to its species-poor sister. Both interpretations are flawed because the sister groups diverge simultaneously from their nearest common ancestor.

This misreading of phylogenies is encouraged by asymmetrical trees (tree balance) [13], because it could appear that one or more lineages are diverging from a main stem. The illusion can be even greater in unbalanced trees (also called 'pectinate' or 'Hennigian comb' trees), especially if the trees are ladderized [14] and misread as though there were a phylogenetic progression in the terminal (extant) taxa, across the top of the tree. However, there are two sister groups across every node in a fully bifurcating tree and, therefore, each node results in two lineages diverging from each other. This applies regardless of whether the tree is balanced (Figure 1b) or not (Figures 1c,d), and whether it is drawn as ladderized (Figure 1c) or not (Figure 1d). Therefore, the misinterpretation results from reading trees in a one-sided way.

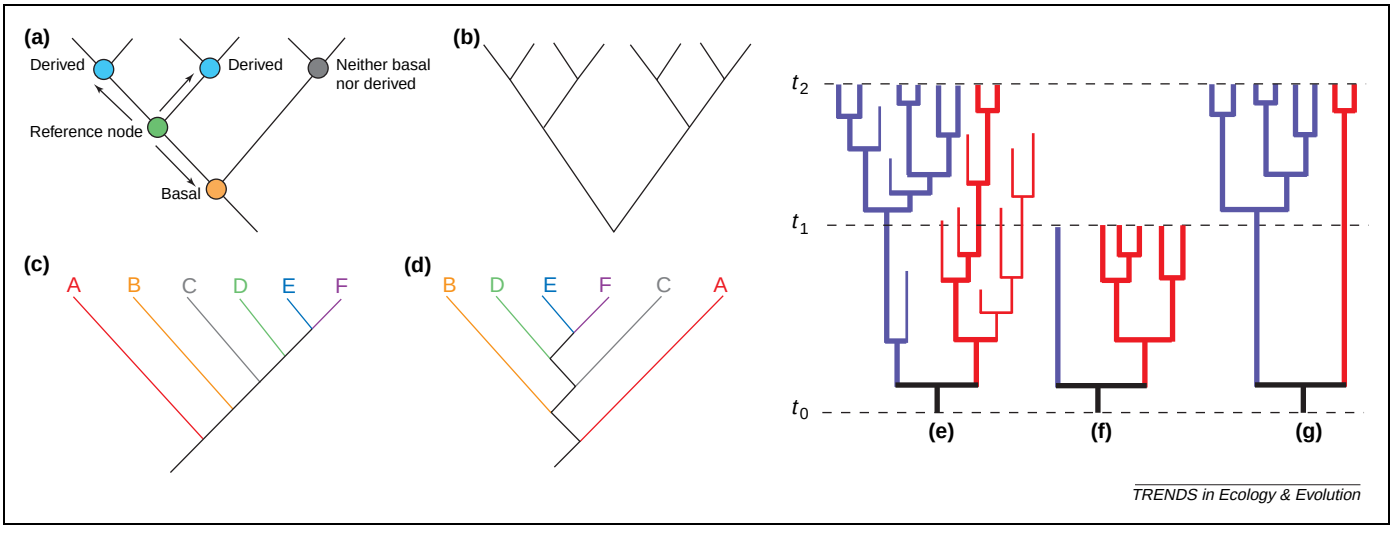

Figure 1. Interpretation of trees in relation to their shape. **(a)** Correct use of 'basal' or 'derived' when referring to a tree. Internal nodes of a tree can be referred to as 'basal' or 'derived', but only with respect to another node (the 'reference node') that is situated along the same path from the root to a terminal of the tree. A node situated on a different root-terminal path is neither basal nor derived with respect to the reference node. **(b–d)** Tree balance. (b) Balanced tree: there is the same number of species on either side of every node. (c) and (d) Unbalanced tree: there are unequal numbers of species on either side of each node. (c) Tree drawn as ladderized by showing the species-poor sister group on the left at every node. This presentation is prone to the naïve interpretation that there is a phylogenetic progression from left ('basal' taxa) to right ('derived' taxa). (d) The same unbalanced tree redrawn by swapping sister groups at alternate nodes. This presentation could be less prone to misinterpretation than the ladderized version. (**e–g**) Hypothetical evolutionary tree showing that tree balance can switch between lineages through time in two sister groups (red and blue). (e) Both sister groups originate at the speciation event at time t_0 and therefore are the same age. All lineages arising within both taxa are shown until the present (t_2). Thin lines indicate lineages that go extinct. (f,g) Phylogeny of the two sister groups at t_1 and t_2 respectively (omitting extinct lineages) showing reversal of tree balance.

Box 2. 'Basal' angiosperms

Relationships at the base of the angiosperms (flowering plants) have received much attention over a long period of time (reviewed in [35,36]), driven by a quest to reconstruct the morphology and biology of the common angiosperm ancestor (e.g. [2,37–39]). Difficulty in achieving this, classically referred to as 'Darwin's abominable mystery', still persists. Reconstruction of ancestral traits of angiosperms has been problematic because of poor phylogenetic resolution at deeper levels, a lack of unequivocal transitional forms in the fossil record, obscure homologies between angiosperms and other seed plants, and the great diversity of traits among extant angiosperms [35]. Some studies have inferred ancestral states objectively by mapping and optimizing characters across whole trees [37,38,40,41]. However, many authors have misinterpreted a species-poor sister group as being 'ancestral' and, as phylogenies have changed (Figure I), so too has the misinterpretation. Groups that were thought successively to be 'basal' are the magnoliids, especially Magnoliaceae (up to 1980), Chloranthaceae (1980s), *Ceratophyllum* (1993) (Figure Ia), and the ANITA grade (*Amborella*, Nymphaeales, Illiciales, Trimeniaceae and *Austrobaileya*) (from 1999) (Figure Ib) [2]. For example:

- Calycanthales were hypothesized to be the 'archetype' of the angiosperms [42].
- Support for *Ceratophyllum* as the sister group was inferred from 'primitive' features identified in the flowers of these plants [43].
- The 'ancestral' DNA content was first inferred from *Ceratophyllum* [44] and later from the ANITA grade [45].
- *Amborella* was stated to represent the first stage of angiosperm evolution [46].
- *Amborella*, inferred as the 'most ancient angiosperm', was examined for chloroplast genes characteristic of gymnosperms (a more distant sister group) [47].
- When rice (*Oryza*, representing the grasses) was placed as the sister group, it was inferred to have greater affinity to *Pinus* than to other angiosperms [43].

These inferences are equivalent to considering monotremes 'reptilian' because of shared egg-laying (Box 1).

Ironically, many of the features identified as being 'basal' in sister groups are likely to be independently derived and highly specialized, such as the morphological and physiological adaptations in *Ceratophyllum* and Nymphaeales to their aquatic habitat [38,43], absence of vessels in the wood of Winteraceae [37,48] and floral development in *Amborella* [49].

Some studies have used an appropriate sampling strategy for their purpose, albeit by targeting 'basal' angiosperms (e.g. [38,40,41,50]). This can be suitable if multiple sister groups are sampled across both sides of the deeper nodes of a phylogeny, rather than focussing on the species-poor sister group alone. Two-sided comparisons are more likely to sample the maximum phylogenetic diversity in traits.

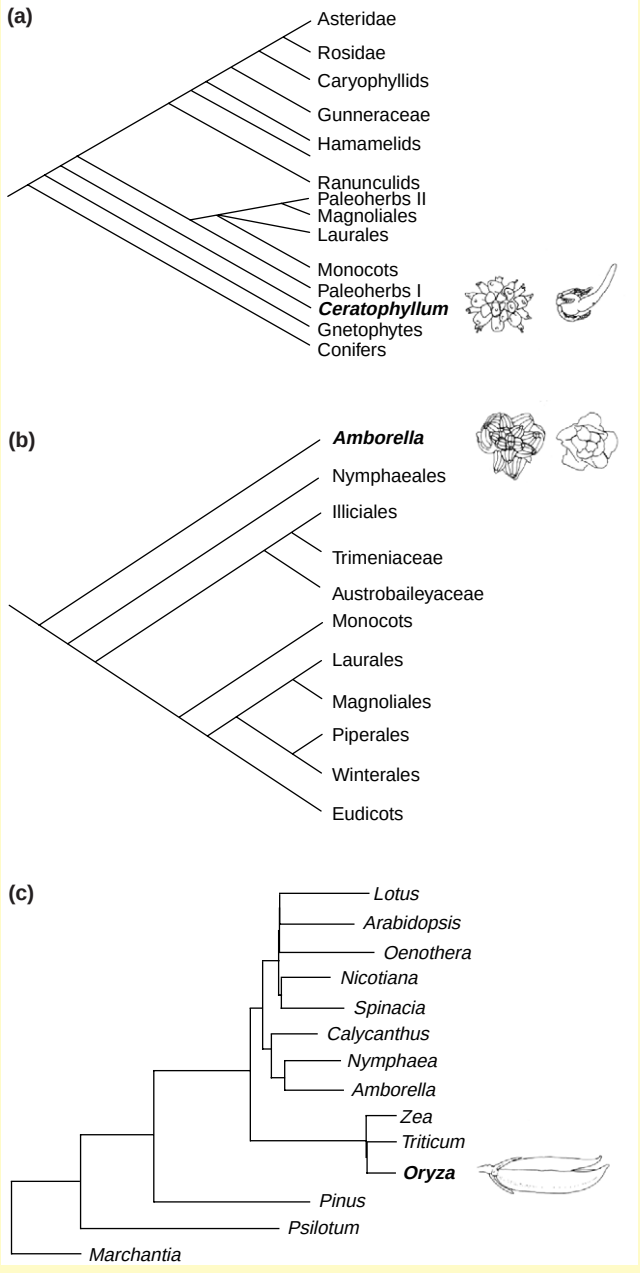

Figure I. Changing sister group reconstructions in angiosperms through time. **(a)** *Ceratophyllum*, **(b)** *Amborella* and **(c)** Grasses. Reproduced, with permission, from [46,51,52], respectively.

One-sided reading of trees can also cause paradoxes. Two sister groups can alternate in appearing 'basal' if the timing of speciation and extinction events differs between the two lineages. For example, in Figure 1e–g, two sister groups undergo the same number of speciation events following divergence from their common ancestor. The major difference between the groups is that one group radiates later (seven speciation events after t_1) than does the second group (five speciation events before t_1) (Figure 1e). If a phylogeny is constructed showing relationships among taxa extant at t_1, the species-poor group (one species) could be misinterpreted as being 'basal' relative to the species-rich group (five species) (Figure 1f). However, at t_2, this misinterpretation would be reversed, with the species-poor group (two species; previously the species-rich group) appearing 'basal' relative to the species-rich group (six species; previously the species-poor group) (Figure 1g). This reversal is paradoxical only in terms of a false assumption that one extant sister is older than the other.

The misinterpretation of one sister as being 'basal' and, therefore, having more ancestral traits compared with the 'derived' sister, can result in an unjustified reconstruction

Box 3. An example of ancestral area reconstruction by misinterpretation

In a study of African violets, *Saintpaulia* spp. [53], an ancestral area was inferred by misinterpretation, although no unambiguous reconstruction is possible from the tree. The African violets are essentially restricted to three isolated mountain ranges, with each range having different endemic violet species. In a molecular phylogeny, each of the three main lineages corresponds to one mountain range (Figure Ia) and the single species in the Uluguru Mts is sister group to the rest. Based on this information alone, Uluguru was inferred as the ancestral area of *Saintpaulia* and the source of dispersal and evolution of the other lineages into the other two mountain ranges (Figure Ib). However, several reconstructions are equally likely, as costed by the total number of events (speciation, dispersal and vicariance). Figures Ib–d illustrate three of six equally probable dispersal reconstructions (all require two dispersals with subsequent speciation). A vicariance explanation (not shown) has the same cost. First, area A splits away from united area B+C, then B and C split into separate areas (two vicariance events with subsequent speciation). Another equally costly scenario (not shown) would be a dispersal from A to B+C, followed by a vicariance between B and C.

Therefore, independent information, such as factors affecting the dispersal of the plants, should be used in conjunction with the tree. Also, dated fossils or known age of mountain ranges (if young relative to the speciation events) could be used with molecular modelling to test hypotheses. For example, if the speciation event separating *S. goetzeana* from the rest of *Saintpaulia* was found to be older than the Uluguru Mts, this range could be ruled out as the ancestral area.

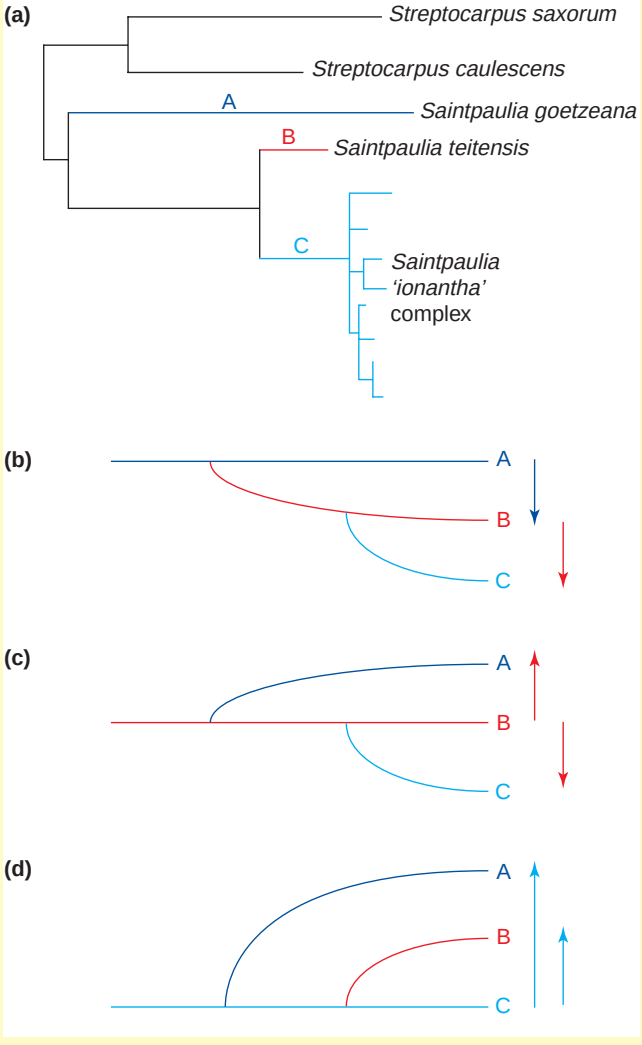

Figure I. Alternative biogeographical reconstructions of *Saintpaulia*. **(a)** A phylogeny of African violets *Saintpaulia* spp. indicating the areas in which they occur (A, Uluguru Mts, B, Teita Hills, C, Usambara Mts). A naïve ('basal') interpretation of this tree would consider the species-poor sister group (blue) to be ancestral and, thus, its area (A), as the centre of origin for the entire lineage. Likewise, taxa in the species-poor area B (red) would be considered to be ancestral to the taxa in area C (violet). **(b)** The same tree simplified and showing the naïve reconstruction of origin and dispersal. **(c)** The same tree results if area B is ancestral, with independent dispersals from this source to areas A and C. **(d)** A third reconstruction, in which area C is ancestral. (a) Reproduced, with permission, from [53].

of ancestral traits that otherwise might have been ambiguous if valid character-state reconstruction methods had been used; that is, traits found exclusively in the species-poor sister group are mistakenly assumed to be 'primitive' and to represent traits of the common ancestor of both sisters. This is the fallacy underlying the quote above about monotremes and is explained in Box 1. However, the intuition that character states occurring in multiple lineages are likely to be ancestral does have some validity (Box 2).

Examples of misinterpretation of trees

Reconstructing the ancestral angiosperm is a popular area of research and the literature is replete with simplistic misinterpretation of phylogenies. Progress in phylogenetic methods and new sources of data have resulted in different species-poor groups being hypothesized successively as being the sister group to the rest of the angiosperms. With misinterpretation, each of these sister groups has been successively assumed to represent the ancestral angiosperm and research programs have been repeatedly refocused to investigate the assumed ancestral features of each (Box 2). Thus, misinterpretation of trees can mislead the setting of research objectives.

Simplistic interpretation of phylogenies is common in biogeography, a complex field that involves questions and methods that are too diverse to be reviewed here (but see [15]). For example, one approach attempts to reconstruct the history of continental break-up by comparisons of phylogenies of multiple taxa [9]. Another approach attempts to trace the history of a single lineage back in time and space to its ancestral area [16]. Several authors have naïvely assumed that a species-poor sister group occurs in the ancestral area (Hennig's Progression Rule [17]). Such an inference can give a false sense of decisiveness in the data, even though no solution is possible from knowing the tree and distribution of terminal taxa alone (Box 3).

How to use phylogenies to reconstruct ancestral states

The reconstruction of ancestral states presents numerous challenges because of the inherent limitations of the

Table 1. Methods and models for inferring ancestral areas

Method or model	Type	Properties	Deficiencies	Examples of appropriate use	Refs
Parsimony	P[a]	Minimizes discrete-state changes over the tree; state changes can be weighted equally (Fitch) or differentially using step matrices	Cannot indicate probability of estimates; does not use branch lengths, thus underestimates change when it is frequent relative to speciation	Double fertilization, long thought to be a unique defining character of angiosperms, might have originated independently in gymnosperms, or earlier, in the first seed plants	[14,19, 41]
Dispersal vicariance analysis (DIVA)	P	Uses cost matrices to estimate ancestral areas; differentially costs vicariance and dispersal	Bias towards sympatric speciation and against early dispersal	Ancestral area reconstructions for oaks were more consistent with fossil record using DIVA than using Fitch parsimony or strict vicariance	[16,24]
Markov continuous time transition model	ML[b]	Estimates rate of discrete-state changes, allowing asymmetry	Evolutionary assumptions, e.g. that rate of change is constant throughout the tree, can be unrealistic (see punctuated models)	Overturned the parsimony estimate of origin of ruminant digestion in artiodactyls by taking rate of change into account	[21][c]
Punctuated evolution model	ML	Assumes evolutionary change occurs only at speciation	Evolutionary assumption is equivalent to having equal branch lengths and likely to be unrealistic, as in parsimony	For *Psychotria* plants in Hawaii, ancestral area inferences differ from both parsimony and ML models using differential branch lengths	[3,29]
Generalized least squares (=Squared change 'parsimony')	ML	Minimizes sum of squared changes over the tree; Brownian motion model of evolution; uses branch lengths	Does not indicate probability of estimates; non-directional version cannot reconstruct ancestral values outside range of descendant values	Size (area) of the first lower molar in mammals was reconstructed accurately, with verification from the fossil record	[18,28, 54]
Maximum likelihood with General time reversible model	ML	Based on Markov model, adapted from nucleotide modelling; allows differential rates and symmetry	Can fail if model unrealistic	Estimated ancestral areas, and rates and direction of dispersal of plants among Hawaiian islands	[3,55]
Stochastic (Bayesian inference), e.g. using Markov continuous time transition model	B[d]	Models multiple parameters including tree topology; posterior probabilities given for estimates	Can fail if model unrealistic	In contrast to parsimony, Bayesian analysis found multiple credible histories of gains and losses of horned soldiers in aphids	[26]

[a]Parsimony.
[b]Maximum likelihood.
[c]Implemented in Mesquite http://mesquiteproject.org.
[d]Bayesian inference.

available data (e.g. [14,18–20]). However, in spite of extensive reviews [19–21] and empirical comparisons of methods [3,22–24], ongoing misinterpretation of phylogenies suggests that basic methods require reiterating (Table 1). Assuming that appropriate methods have been used to estimate the phylogeny [25], character states are mapped onto the terminals of a rooted tree and are optimized to the internal nodes using parsimony, maximum likelihood (ML), or Bayesian inference with ML. There are several explicit optimization methods available for reconstructing ancestors using trees (Table 1). All methods are assumption based, even if authors and users claim that they are not. However, they differ markedly in which processes they consider, and different methods using the same data can give very different results, as shown by the comparative studies mentioned below.

Parsimony-optimization prefers the reconstruction that implies the fewest character-state changes over the phylogenetic tree [14]. Parsimony methods can explicitly model processes to a limited degree (Table 1). ML modelling of the evolution of traits on trees has several advantages over parsimony [19–21] (Table 1). Whereas parsimony reconstructs a single state at each node, ML can indicate probability of alternative states. If change is highly probable relative to branch length, then a basic assumption of parsimony is violated and the rate of evolutionary change is likely to be underestimated; however, ML uses branch lengths to model the rate of change along each branch. A Bayesian approach to ML modelling (Table 1) is inherently statistical and takes account of uncertainty in tree topology, branch lengths and parameter estimation, in contrast to the other methods [26,27]. Furthermore, there is potential for developing *a priori* hypotheses about evolutionary events and testing them using the phylogeny; for example, a test that two taxa share the same character state by descent [26].

Limitations of a model-based approach

Tree-based inference of ancestral traits has come into widespread use because sharing of traits among related organisms has a strong phylogenetic component. Unfortunately, reconstructing the evolution of traits on trees is

problematic because, in most cases, the estimate cannot be validated against independent knowledge. Ancient evolutionary events leave little trace apart from the changes to DNA and morphological traits that are used to construct phylogenies. Therefore, models have been developed that make simple assumptions about the rate and mode of evolution and are applied over whole phylogenies. However, every lineage evolves independently of its sister group, and the rate and mode can change independently in each. The few studies that have compared known histories with reconstructions using standard models have found that most models make serious errors [22,23,28]. For traits with high rates of change, tree-based modelling should probably not be used, because the evolution of such traits is better explained at the population level [29]. Simple parsimony is consistently poorer than most other models, but no model succeeds in all circumstances. For example, both the parsimony and generalized least squares methods (Table 1) failed to reconstruct a known (laboratory-controlled) history of rapid and directional phenotypic change in a bacteriophage lineage, even when the known character states of the ancestor were included in the analysis [23]. In this study, the independent contrasts method, which reconstructs ancestors only as a by-product of estimating correlations between two traits, succeeded where the other methods failed. Yet, in another study, of planktonic Foraminifera, using a comprehensive fossil record to test the tree-based reconstructions, a two-parameter ML model outperformed four other models, including independent contrasts [22].

When there is knowledge of evolutionary processes in the traits of interest, and this is independent of the phylogeny, it might be possible to develop a realistic model that successfully reconstructs ancestors [19,22,23,27,28]. Such information could come from genetics, molecular and developmental biology, functional morphology, geological and climatic processes, and the fossil record. This approach has been used to reconstruct dispersal of plants in the Hawaiian islands by including, in a ML model, the known history of islands that originated linearly from a volcanic hot spot [3].

Summary

Intuitive interpretation of ancestry from trees is likely to lead to errors, especially the common fallacy that a species-poor lineage is more 'ancestral' or 'diverges earlier' than does its species-rich sister group. Errors occur when trees are read in a one-sided way, which is more commonly done when trees branch asymmetrically. If ancestral features are to be inferred from a phylogeny, a method that optimizes character states over the whole tree should be used.

There are several explicit optimization methods available for reconstructing ancestors using trees (Table 1). However, they differ markedly in which processes they consider, and different methods using the same data can give different results, as shown by the comparative studies mentioned above. Reconstructions using current methods that use data from the tree alone are more ambiguous than is generally appreciated. Unrealistic models are likely to fail and, therefore, models should use all the relevant independent information, as well as that from the tree. Bayesian inference can simultaneously model a wide range of variables and, moreover, it estimates the uncertainty of the results [27]. However, complex models are likely to be applicable only to the specific lineages and situations for which they were developed.

Acknowledgements

We thank Simon Blomberg, Peter Cranston, Cliff Cunningham, Penny Gullan, David Morris, Lindy Orthia, Chris Simon, the Coopers and Cladistics discussion group and two anonymous referees for useful suggestions about the article.

References

1 Dawkins, R. (2004) *The Ancestor's Tale: A Pilgrimage to the Dawn of Life*, Weidenfeld & Nicolson
2 Endress, P.K. (2001) The flowers in extant basal angiosperms and inferences on ancestral flowers. *Int. J. Plant Sci.* 162, 1111–1140
3 Nepokroeff, M. *et al*. (2003) Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understorey tree genus *Psychotria* (Rubiaceae): a comparison of parsimony and likelihood approaches. *Syst. Biol.* 52, 820–838
4 Oakley, T.H. and Cunningham, C.W. (2002) Molecular phylogenetic evidence for the independent evolutionary origin of an arthropod compound eye. *Proc. Natl. Acad. Sci. U. S. A.* 99, 1426–1430
5 Page, R.D.M., ed. (2003) *Tangled Trees: Phylogeny, Cospeciation and Coevolution*, The University of Chicago Press
6 Lutzoni, F. *et al*. (2001) Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* 411, 937–940
7 Johannesson, K. (2001) Parallel speciation: a key to sympatric divergence. *Trends Ecol. Evol.* 16, 148–153
8 Mayhew, P.J. (2002) Shifts in hexapod diversification and what Haldane could have said. *Proc. R. Soc. Lond. Ser. B* 269, 969–974
9 Sanmartín, I. and Ronquist, F. (2004) Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst. Biol.* 53, 216–243
10 Cockburn, A. (2003) Cooperative breeding in oscine passerines: does sociality inhibit speciation? *Proc. R. Soc. Lond. Ser. B* 270, 2207–2214
11 Krell, F-T. and Cranston, P.S. (2004) Which side of the tree is more basal? *Syst. Entomol.* 29, 279–281
12 Page, R.D.M. and Holmes, E.C. (1998) *Molecular Evolution: A Phylogenetic Approach*, Blackwell Science
13 Agapow, P.M. and Purvis, A. (2002) Power of eight tree shape statistics to detect nonrandom diversification: a comparison by simulation of two models of cladogenesis. *Syst. Biol.* 51, 866–872
14 Maddison, D.R. and Maddison, W.P. (2003) *MacClade 4: Analysis of Phylogeny and Character Evolution*, Sinauer Associates
15 Crisci, J.V. *et al*. (2003) *Historical Biogeography: An Introduction*, Harvard University Press
16 Ronquist, F. (1997) Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46, 195–203
17 Humphries, C.J. (2001) Vicariance biogeography. In *Encyclopaedia of Biodiversity* (Vol. 5) (Levin, S.A., ed.), pp. 767–779, Academic Press
18 Schluter, D. *et al*. (1997) Likelihood of ancestor states in adaptive radiation. *Evolution* 51, 1699–1711
19 Cunningham, C.W. *et al*. (1998) Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol. Evol.* 13, 361–366
20 Omland, K.E. (1999) The assumptions and challenges of ancestral state reconstructions. *Syst. Biol.* 48, 604–611
21 Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature* 401, 877–884
22 Webster, A.J. and Purvis, A. (2002) Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proc. R. Soc. Lond. Ser. B* 269, 143–149
23 Oakley, T.H. and Cunningham, C.W. (2000) Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. *Evolution* 54, 397–405
24 Manos, P.S. and Stanford, A.M. (2001) The historical biogeography of Fagaceae: tracking the tertiary history of temperate and subtropical forests of the Northern Hemisphere. *Int. J. Plant Sci.* 162, S77–S93

25 Holder, M. and Lewis, P.O. (2003) Phylogeny estimation: traditional and Bayesian approaches. *Nat. Rev. Genet.* 4, 275–284
26 Huelsenbeck, J.P. *et al*. (2003) Stochastic mapping of morphological characters. *Syst. Biol.* 52, 131–158
27 Ronquist, F. (2004) Bayesian inference of character evolution. *Trends Ecol. Evol.* 19, 475–481
28 Polly, P.D. (2001) Paleontology and the comparative method: ancestral node reconstructions versus observed node values. *Am. Nat.* 157, 596–609
29 Losos, J.B. (1999) Commentaries – Uncertainty in the reconstruction of ancestral character states and limitations on the use of phylogenetic comparative methods. *Anim. Behav.* 58, 1319–1324
30 Musser, A.M. (2003) Review of the monotreme fossil record and comparison of palaeontological and molecular data. *Comp. Biochem. Physiol. A.* 136, 927–942
31 Phillips, M.J. and Penny, D. (2003) The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol. Phylogenet. Evol.* 28, 171–185
32 Janke, A. *et al*. (2002) Phylogenetic analysis of 18S rRNA and the mitochondrial genomes of the wombat, *Vombatus ursinus*, and the spiny anteater, *Tachyglossus aculeatus*: increased support for the Marsupionta hypothesis. *J. Mol. Evol.* 54, 71–80
33 Grützner, F. *et al*. (2003) The monotreme genome: a patchwork of reptile, mammal and unique features? *Comp. Biochem. Physiol. A.* 136, 867–881
34 Grützner, F. *et al*. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature:* doi:10.1038/nature03021
35 Crane, P.R. *et al*. (1995) The origin and early diversification of angiosperms. *Nature* 374, 27–33
36 Judd, W.S. *et al*. (2002) *Plant Systematics: A Phylogenetic Approach*, Sinauer Associates
37 Doyle, J.A. and Endress, P.K. (2000) Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. *Int. J. Plant Sci.* 161, S121–S153
38 Feild, T.S. *et al*. (2003) The ancestral ecology of angiosperms: emerging perspectives from basal lineages. *Int. J. Plant Sci.* 164, S129–S142
39 Mathews, S. *et al*. (2003) Adaptive evolution in the photosensory domain of Phytochrome A in early angiosperms. *Mol. Biol. Evol.* 20, 1087–1097
40 Floyd, S.K. and Friedman, W.E. (2000) Evolution of endosperm developmental patterns among basal flowering plants. *Int. J. Plant Sci.* 161, S57–S81
41 Friedman, W.E. and Floyd, S.K. (2001) Perspective: the origin of flowering plants and their reproductive biology – A tale of two phylogenies. *Evolution* 55, 217–231
42 Loconte, H. and Stevenson, D.W. (1991) Cladistics of the Magnoliidae. *Cladistics* 7, 267–296
43 Endress, P.K. (1994) Evolutionary aspects of the floral structure in *Ceratophyllum*. *Plant Syst. Evol.* 8 (Suppl.), 175–183
44 Leitch, I.J. *et al*. (1998) Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. *Ann. Bot.* 82 (Suppl. A), 85–94
45 Leitch, I.J. and Hanson, L. (2002) DNA C-values in seven families fill phylogenetic gaps in the basal angiosperms. *Bot. J. Linn. Soc.* 140, 175–179
46 Qiu, Y.L. *et al*. (1999) The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402, 404–407
47 Goremykin, V.V. *et al*. (2003) Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Mol. Biol. Evol.* 20, 1499–1505
48 Feild, T.S. *et al*. (2002) Hardly a relict: freezing and the evolution of vesselless wood in Winteraceae. *Evolution* 56, 464–478
49 Posluszny, U. and Tomlinson, P.B. (2003) Aspects of inflorescence and floral development in the putative basal angiosperm *Amborella trichopoda* (Amborellaceae). *Can. J. Bot.* 81, 28–39
50 Endress, P.K. and Igersheim, A. (2000) Gynoecium structure and evolution in basal angiosperms. *Int. J. Plant Sci.* 161, S211–S223
51 Chase, M.W. *et al*. (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. *Ann. Mo. Bot. Gard.* 80, 528–580
52 Goremykin, V.V. *et al*. (2004) The chloroplast genome of *Nymphaea alba*, whole genome analyses and the problem of identifying the most basal angiosperm. *Mol. Biol. Evol.* 21, 1445–1454
53 Möller, M. and Cronk, Q.C.B. (1997) Phylogeny and disjunct distribution: evolution of *Saintpaulia* (Gesneriaceae). *Proc. R. Soc. Lond. Ser. B* 264, 1827–1836
54 Martins, E.P. and Hansen, T.F. (1997) Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* 149, 646–667
55 Rodriguez, F. *et al*. (1990) The general stochastic-model of nucleotide substitution. *J. Theor. Biol.* 142, 485–501