Palaeont. afr., **38**, 27-32 (2002)

CHARACTER STATE TRANSFORMATIONS AND THE FIT OF PHYLOGENIES TO THE FOSSIL RECORD

by

Kenneth D. Angielczyk

Department of Integrative Biology and Museum of Paleontology, 1101 Valley Life Sciences Building,
University of California, Berkeley, CA 94720-4780 USA
e-mail: etranger@socrates.berkeley.edu

ABSTRACT

There is only one true history of life, and the biostratigraphic record and the phylogenetic relationships of organisms provide the most important information regarding this history. Ideally, the historical signal preserved in each of the data sets should be the same, and several methods have been proposed to compare the fit of phylogenies to the fossil record. All of these techniques use stratigraphic data associated with taxa, but our ability to recognize taxa and reconstruct their phylogenetic relationships ultimately is based on patterns of character state distributions that we observe. This raises the question of whether character states can be used to measure the fit of a phylogeny to the fossil record. Here I argue that we can, if the *order of appearance of character states* is considered. Optimization of character states on a phylogeny results in a predicted order of appearance of character states; derived states must arise after basal states. This order can be compared to that predicted by the fossil record. Although a number of factors can affect the frequency at which derived character states are sampled before basal states in the fossil record, conflicts between the two data sets should be relatively rare. Phylogenies that imply a large number of character state transformations that are inconsistent with the fossil record may need to be reconsidered before the fossil record is criticized.

KEYWORDS: Phylogeny, Stratigraphy, Character-Based Measures of Fit, Taxon-Based Measures of Fit

INTRODUCTION

The history of life cannot be observed directly. However, several sources of historical information allow inferences about evolutionary history. The phylogenetic relationships of organisms provide insight into the order of appearance of clades and the evolution of their distinctive suites of features. Phylogenetic studies have long played a central role in palaeontological research. The fossil record also preserves information about the relative and absolute ages of clades and their evolutionary histories, and documents organisms that otherwise would be unknown. Because there is only one true history, the signal preserved in each data set should be the same and predictions based on one can be tested with observations from the other.

A variety of methods compares the fit of a phylogeny to the stratigraphic record (e.g., Benton & Storrs 1994; Gauthier *et al.* 1988; Huelsenbeck 1994; Norell & Novacek 1992a, 1992b; Siddall 1998; Wills 1999). Others use stratigraphic information directly to construct phylogenetic trees (e.g., Clyde & Fisher 1997; Fisher 1988, 1991, 1992, 1994, 1997; Gingerich 1979; Huelsenbeck & Rannala 1997; Wagner 1995). All of these methods use stratigraphic data associated with taxa, and proceed from the premise that the order of appearance of taxa on a phylogeny and in the stratigraphic record ideally should be the same. However, when a cladogram is combined with a reconstruction of how a character or characters might have evolved, predictions about the order of appearance of character states are made. The predicted order can be compared to the order observed in the fossil record, providing a measure of how well a particular phylogeny (i.e., a cladogram and a hypothesis of character evolution) fits the fossil record. In this study, I will explore the concept that the order of character state transformations can be congruent or incongruent with the fossil record. Elsewhere I present a method that uses this idea to measure the fit between phylogenetic trees and the fossil record (Angielczyk 2002).

CHARACTERS, TAXA, AND THE STRATIGRAPHIC RECORD

When a phylogeny is to be compared to the fossil record, the included Operational Taxonomic Units (OTUs) traditionally have been used to make the comparison. The OTUs are an integral part of the data matrix that is analyzed, and usually the pattern of relationships among the OTUs is of primary interest. Also, palaeontologists are accustomed to thinking about stratigraphic data in terms of taxa (e.g., the first appearance, range, and last appearance of a genus). The methods available to measure the fit of a phylogeny to stratigraphy reflect these patterns of thought. For example, the Spearman Rank Correlation (SRC, Gauthier *et al.*, 1988; Norell & Novacek 1992a, 1992b) compares the order of appearance of OTUs on a cladogram to that in the fossil record. The Relative Completeness Index (RCI, Benton & Storrs 1994) compares the amount of time represented by the ranges of the OTUs to the length of time represented by gaps between those OTUs that is implied by a particular topology. The Manhattan Stratigraphic Metric (MSM, Siddall 1998) records the pair-wise distances between OTUs as a matrix, optimizes that matrix on a given topology, and compares the resulting length value to the minimum possible length for the distance matrix. These and other stratigraphic metrics provide some means to assess the degree of fit between a phylogenetic tree and the fossil record. Many of these techniques have been criticized on methodological grounds (e.g., Benton & Storrs, 1994; Hitchin & Benton 1997a, 1997b; Pol & Norell 2001; Siddall, 1996, 1997, 1998; Wagner, 2000a, 2000b; Wagner & Sidor 2000; Wills 1999), but regardless of those potential problems all of the methods use the same stratigraphic data set, that associated with taxa.

Although they play an equally fundamental role in phylogenetic analysis, characters have been overlooked in studies comparing phylogenies to the fossil record. Most phylogenetic systematists agree that taxa should be monophyletic groups (i.e., clades). However, we cannot directly observe clades or the phylogenetic relationships they are based upon. Fortunately, because organisms inherit many of their features from their ancestors, the characters of organisms can be used to reconstruct their phylogenetic relationships. Once we have constructed a hypothesis about the relationships of the organisms of interest based on the characters we have observed, we can apply names to monophyletic groups. If new specimens are collected, we can determine whether they belong to new or pre-existing taxa by studying their characters and using the resulting information to determine where they fall on our phylogenetic tree. Thus, our ability to discover taxa, determine relationships among them, and even determine whether a particular specimen is a member of a taxon ultimately is based upon characters. These ideas are not new (see e.g., de Queiroz & Gauthier 1990, 1992; Hennig 1966; Mishler & Theriot 2000a, 2000b, 2000c; Sober 1988; Wiley 1981 for pertinent reviews and discussions), but they are important to consider here. If characters are the basic data we need to build phylogenies and recognize the specimens we find in the fossil record as members of taxa, an obvious question is whether we can use characters to measure how well a particular phylogeny fits the known fossil record.

I argue that characters can measure the fit of a phylogeny to the fossil record. One way to implement such an approach is to consider the order in which character states appear. When characters are optimized on a phylogeny, a definite order of appearance is implied; relatively derived states must appear after more basal ones. Of course, the exact details of the order of appearance will vary with the optimization used. Often the reconstruction that requires the fewest instances of evolutionary change (i.e., the most parsimonious optimization) is used, although many other optimizations that may fit the fossil record better or have a higher likelihood under a given model are possible for a particular topology (Figure 1).

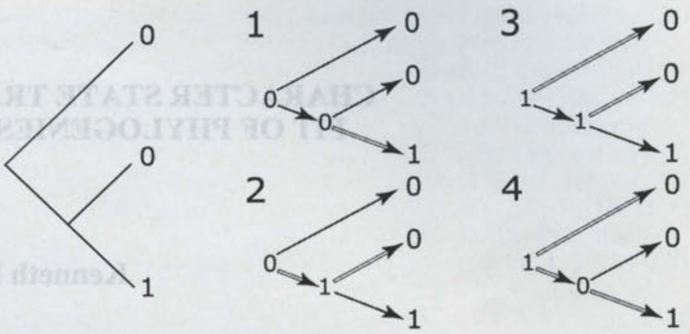

Figure 1: Alternative possible optimizations given the observed character states in the OTUs and the inferred cladogram, assuming that a maximum of one character state change can occur per branch. Open branches indicate where character state changes occur and numbers at nodes represent ancestral state reconstructions. Optimization 1 is the most parsimonious optimization. However, the other optimizations also may warrant consideration. For example, optimization 3 might have the highest likelihood under a model that assumes there is a driven trend that predisposes character state 1 to evolve into character state 0. Modified from Wagner (2000b).

The fossil record also provides a means to make predictions about the order of evolution of different character states. Indeed, the record itself does not document the appearance and disappearance of taxa, but rather the characters or suites of characters that diagnose them. If the fossil record was perfectly complete and resolved, and the true phylogeny of life was known, the stratigraphic and phylogenetic ordering of character state transformations would be the same. However, because neither the stratigraphic record nor our knowledge of the phylogenetic relationships of organisms is perfect or complete, discrepancies may become apparent when the predicted order of transformations of each are compared. I do not endorse the paleontological criterion for establishing polarity that some authors advocate (e.g., Fortey & Chatterton 1988; Gingerich & Schoeninger 1977; Harper 1976; Szalay 1977; although also see Nelson & Platnick 1981), because I am not attempting to establish the "correct" polarity of character states and I do not assume that "older" states must unequivocally be basal. My goal is only to compare the closeness of fit of the predicted order of character state changes in the two data sets under the assumption that ideally they should be the same.

CONSISTENCY

In this study I use specific definitions to recognize character state transformations that are congruent or incongruent with the fossil record. The definitions are based on the assumption that because sister taxa share a hypothetical common ancestor, they (as stem-groups) must be of the same age (Hennig 1965). Note that this is a simplifying assumption that assumes the actual phylogeny of a group is the same as the cladogram in question (which may not be true; several phylogenetic trees can be consistent with a cladogram; Eldredge 1979) and none of the OTUs is ancestral to any other OTUs. It is worth noting that elsewhere I described a method based on the ideas presented here as a way to measure the fit of a *cladogram* to the fossil record (Angielczyk 2002). I used this terminology because of the above assumption. However, as P. Wagner (personal communication, 2001) noted, because the method is based on reconstructions of character state evolution, it measures the fit of a *phylogeny* (that has the same topology as the original cladogram) to the fossil record. Although a minor point, I hope this clarification will help prevent confusion.

When a new character state appears in a particular taxon, the sister group of that taxon must be of the same age and often will have a more basal state of the character in question. I define a *Type I* consistent state change as a character state transformation that occurs such that a relatively derived state is found in a lineage whose sister taxon is of the same stratigraphic age (Figure 2). However, because the known fossil record is incomplete, sister taxa often have different stratigraphic ages. A *Type II* consistent state change is defined as a character state transformation that occurs such that a relatively derived state is found in a lineage whose sister taxon is of older stratigraphic age (Figure 2). Assuming the phylogeny is correct, this type of change could be the result of the non-preservation of some members of a lineage resulting in unequal stratigraphic ages of sister taxa, or it might indicate that sister taxa more closely related to the younger lineage have not been preserved, collected, or recognized. However, because the ordering of character state transformations predicted by phylogeny and stratigraphy is the same, the transformation is considered consistent. For both Type I and Type II changes, character state reversals can be considered consistent if they represent a locally derived state in a particular lineage or clade (Figure 2).

An inconsistent state change is defined as a character state transformation that occurs such that the relatively derived state is found in a lineage whose sister taxon is of a younger stratigraphic age (Figure 2). Several factors could cause such an inconsistency. For example, an incomplete fossil record or poor choice of outgroups could lead to an incorrect estimation of global polarity. Alternatively, inconsistencies could be caused by incorrect phylogenies or the erroneous optimization of a character or characters on a phylogeny. Regardless, the stratigraphy-based and phylogeny-based predictions of the order of character state transformations conflict, and the change is considered inconsistent.

An estimate of how closely a given phylogenetic tree fits the known fossil record can be obtained by examining all characters and finding the total number of inconsistent character state changes. This can be done as a two-step process in MacClade (Maddison & Maddison 1999). First, a stratigraphic character, coded so each state represents a first appearance in a particular stratigraphic interval, is optimized on the cladogram. This provides a simple graphic representation of the stratigraphic ages of all lineages in the cladogram, including interior branches. Character state transformations can then be optimized on the cladogram using one of the available resolving options

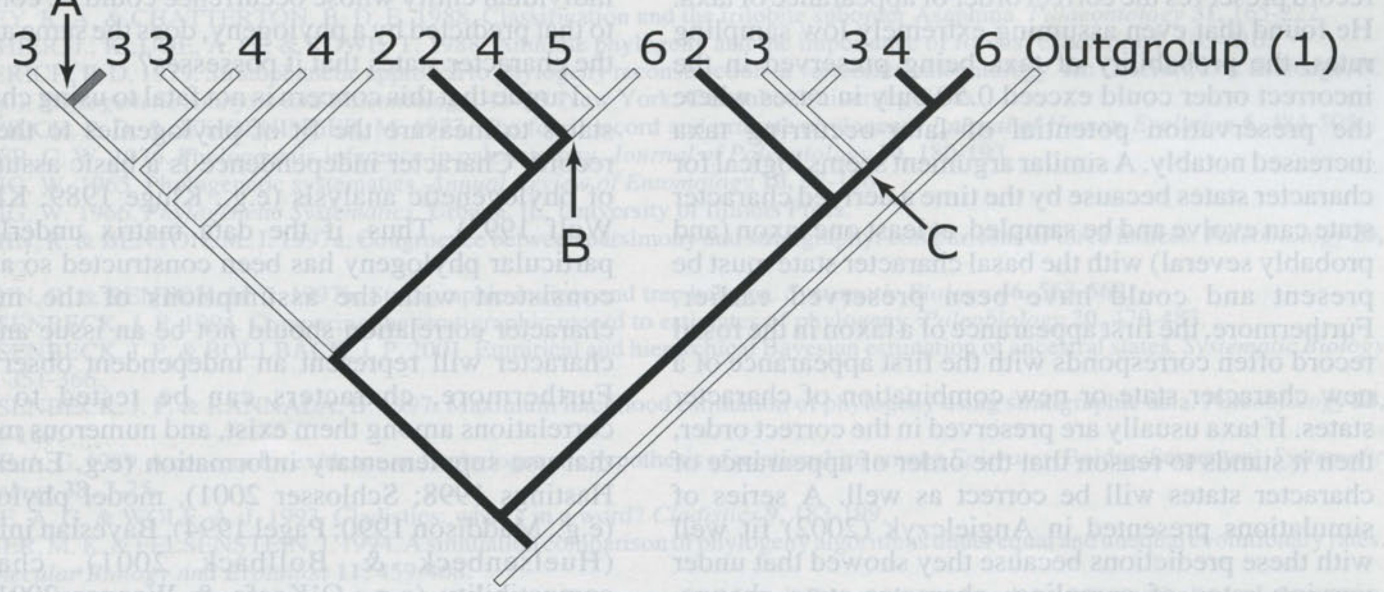

Figure 2: Hypothetical cladogram with three character state changes of interest. Numbers represent the stratigraphic interval in which each taxon appears, and branch shading reflects the character state reconstructed for a particular branch. 'A' represents a Type I consistent state change because both sister taxa are of equal stratigraphic age. 'B' represents a Type II consistent state change because the relatively derived state occurs in the stratigraphically younger lineage. Note also that although taxa 5 and 6 possess the globally basal character state, this state represents the derived state locally. 'C' represents an inconsistent state change because the relatively derived state occurs in the stratigraphically older sister taxon.

in MacClade (e.g., ACCTRAN, DELTRAN, Equivocal Cycling) and inconsistent changes noted (ambiguous changes are excluded). The optimizations available in MacClade are based on parsimony, but other optimizations (e.g., maximum likelihood given a particular model of evolution) could be examined as a separate step. Unfortunately, no method currently exists to automate the process completely, so it is time-consuming for large data sets. However, because the number of characters included in a phylogenetic analysis typically greatly exceeds the number of taxa, this type of estimate should provide a more sensitive measure of the fit of a cladogram to stratigraphy. The Character Consistency Ratio (CCR; Angielczyk 2002) represents a more formal implementation of this approach.

CONCERNS

Several issues are important to consider when comparing phylogeny-based and stratigraphy-based predictions of the order of character state transformations. Angielczyk (2002) noted a number of factors that can affect the frequency at which derived character states are sampled before basal states and that discussion will not be repeated here. However, two additional points related to using character states to measure the fit of phylogenies to the fossil record that were not discussed explicitly by Angielczyk (2002) are presented below.

Perhaps the most important of these considerations is how often we should expect to find conflicts between the predicted orders of appearance (i.e., how often are derived states sampled before basal states). If this is a rare occurrence, then a method like the CCR might be very useful for highlighting dubious phylogenies. Paul (1982) considered the question of how well the fossil record preserves the correct order of appearance of taxa. He found that even assuming extremely low sampling rates, the probability of taxa being preserved in the incorrect order could exceed 0.50 only in cases where the preservation potential of later-occurring taxa increased notably. A similar argument seems logical for character states because by the time a derived character state can evolve and be sampled, at least one taxon (and probably several) with the basal character state must be present and could have been preserved earlier. Furthermore, the first appearance of a taxon in the fossil record often corresponds with the first appearance of a new character state or new combination of character states. If taxa usually are preserved in the correct order, then it stands to reason that the order of appearance of character states will be correct as well. A series of simulations presented in Angielczyk (2002) fit well with these predictions because they showed that under varying rates of sampling, character state change, speciation, and extinction, the frequency at which derived states are sampled before basal states rarely exceeds 0.35 and usually is much lower.

Given these observations, it is interesting to note that there are scenarios when we might expect some phylogenetic reconstruction methods to make inaccurate predictions about the order of appearance of character states, whereas the correct order is preserved in the fossil record. For example, simulation and empirical studies (e.g., Kuhner & Felsenstein 1994; Lamboy 1994; Wagner 1999) have shown that parsimony can produce inaccurate results if character transition patterns are biased (e.g., driven trends *sensu* McShea 1994). The resulting phylogenies can imply large gaps in the fossil record and suggest incorrect polarities for characters that have biased transition patterns (Wagner 1999). Thus, if conflicts between the order of appearance of character states predicted by the two data sets arise, it may be unwarranted to assume *a priori* that the fossil record is wrong. In such cases it may be necessary to reexamine the phylogenetic data using methods such as the likelihood techniques described by Wagner (1999, 2000b) that can take into account more complex hypotheses of character evolution and fossil preservation.

A second consideration raised by D. Fisher (personal communication, 2001) concerns the question of whether character state transformations represent independent observations that can be used to compare phylogeny and stratigraphy. This is important for two reasons. First, if different characters are part of an adaptive complex or under common developmental control, they may be predisposed to change in a correlated fashion (see review in Emerson & Hastings 1998). Thus if one of the characters is preserved or reconstructed in the incorrect order, there is a high probability that other related characters will be as well, leading to a spuriously poor fit between a phylogeny and the fossil record. Second, character states always occur in groups in the form of individual specimens. Although a fossil found in a bed may represent an individual entity whose occurrence could be compared to that predicted by a phylogeny, does the same apply to the character states that it possesses?

I argue that this concern is not fatal to using character states to measure the fit of phylogenies to the fossil record. Character independence is a basic assumption of phylogenetic analysis (e.g., Kluge 1989; Kluge & Wolf 1993). Thus, if the data matrix underlying a particular phylogeny has been constructed so as to be consistent with the assumptions of the method, character correlation should not be an issue and each character will represent an independent observation. Furthermore, characters can be tested to see if correlations among them exist, and numerous methods that use supplementary information (e.g. Emerson & Hastings 1998; Schlosser 2001), model phylogenies (e.g., Maddison 1990; Pagel 1994), Bayesian inference (Huelsenbeck & Bollback 2001), character compatibility (e.g., O'Keefe & Wagner 2001), and likelihood (Wagner 2000b), have been proposed. A similar argument pertains to the question of whether character states found in a single specimen represent independent observations. If systematists consider each character state found in a specimen as an independent observation regarding the phylogenetic relationships of the specimen (or the taxon to which it belongs), then there seems to be no reason not to treat the character states as independent observations regarding stratigraphy as well.

CONCLUSIONS

The stratigraphic record is an important data set that can be used to test the results of phylogenetic analyses. To date, most methods proposed to measure the fit of a phylogeny to stratigraphy have used only one data set, the stratigraphic relationships of taxa. However, phylogenies also make predictions about the order of appearance of character states, and this order can be compared to that observed in the fossil record. Although a number of factors can affect the exact frequency at which derived character states are preserved before basal states in the fossil record, we should expect this frequency to be low. Phylogenies that imply a large number of character state changes that conflict with the fossil record may need to be reconsidered with different methodologies before the inadequacy of the fossil record is invoked.

ACKNOWLEDGEMENTS

A preliminary version of this paper was presented at the 2000 meeting of the Palaeontological Society of Southern Africa. I thank P. Wagner and an anonymous reviewer for many helpful comments. A. Aronowsky, B. Clemens, P. Holroyd, J. Hutchinson, L. Leighton, D. Lindberg, B. Mishler, K. Padian, J. Parham, and W. Alvarez also read various drafts and provided helpful discussions and comments throughout the course of this study. Any remaining mistakes or oversights are my own. This paper constitutes part of my doctoral dissertation research in the Department of Integrative Biology at the University of California. This is UCMP contribution 1763.

REFERENCES

ANGIELCZYK, K. D. 2002 A character-based method for measuring the fit of cladograms to the fossil record. *Systematic Biology* **51,** 176-191

BENTON, M. J. & STORRS, G. W. 1994. Testing the quality of the fossil record: paleontological knowledge is improving. *Geology* **22,** 111-114.

CLYDE, W. C. & FISHER, D. C. 1997. Comparing the fit of stratigraphic and morphologic data in phylogenetic analysis. *Paleobiology* **23,** 1-19.

DE QUEIROZ, K. & GAUTHIER, J. A. 1990. Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Systematic Zoology* **39**, 307-322.

DE QUEIROZ, K. & GAUTHIER, J. A. 1992. Phylogenetic taxonomy. *Annual Review of Ecology and Systematics* **23,** 449-480.

EMERSON, S. B. & HASTINGS, P.A. 1998. Morphological correlations in evolution: consequences for phylogenetic analysis. *Quarterly Review of Biology* **73,** 141-162.

FISHER, D. C. 1988. Stratocladistics: integrating stratigraphic and morphologic data in phylogenetic inference. *Geological Society of America Abstracts with Programs* **20,** A186.

FISHER, D. C. 1991. Phylogenetic analyses and its application in evolutionary paleobiology. **In:** Gilinsky, N. L. & Signor, P. W. (eds.), *Analytical Paleobiology*, 130-121. Knoxville, TN, Paleontological Society.

FISHER, D. C. 1992. Stratigraphic Parsimony. **In:** Maddison, W. P. & Maddison, D. R. (eds.), *MacClade: Analysis of Phylogeny and Character Evolution, Version 3,* 124-129. Sunderland, MA, Sinauer Associates.

FISHER, D. C. 1994. Stratocladistics: morphological and temporal patterns and their relation to phylogenetic process. **In:** Grande, L. & Reippel, O. (eds.), *Interpreting the Hierarchy of Nature – From Systematic Patterns to Evolutionary Theories*, 133-171. San Diego, Academic Press.

FISHER, D. C. 1997. Stratocladistics and hypothesis choice. *Journal of Vertebrate Paleontology* **17** (supplement to 3): 46A.

FORTEY, R. A. & CHATTERTON, B. D. E. 1988. Classification and the trilobite suborder Asaphina. *Palaeontology* **31,** 165-222.

GAUTHIER, J., KLUGE, A. G., & ROWE, T. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* **4,** 105-210.

GINGERICH, P. D. 1979. Stratophenetic approach to phylogeny reconstruction in vertebrate paleontology. **In:** Cracraft, J. & Eldredge, N. (eds.), *Phylogenetic Analysis and Paleontology*, 41-78. New York, Columbia University Press.

GINGERICH, P. D. & SCHOENINGER, M. 1977. The fossil record and primate phylogeny. *Journal of Human Evolution* **6,** 484-505.

HARPER, C. W. 1976. Phylogenetic inference in paleontology. *Journal of Paleontology* **50,** 180-193.

HENNIG, W. 1965. Phylogenetic systematics. *Annual Review of Entomology* **10,** 97-116.

HENNIG, W. 1966. *Phylogenetic Systematics*. Urbana, IL, University of Illinois Press.

HITCHIN, R. & BENTON, M. J. 1997a. Congruence between parsimony and stratigraphy: comparisons of three indices. *Paleobiology* **23,** 20-32.

HITCHIN, R. & BENTON, M. J. 1997b. Stratigraphic indices and tree balance. *Systematic Biology* **46,** 563-569.

HUELSENBECK, J. P. 1994. Comparing the stratigraphic record to estimates of phylogeny. *Paleobiology* **20,** 470-483.

HUELSENBECK, J. P. & BOLLBACK, J. P. 2001. Empirical and hierarchical Bayesian estimation of ancestral states. *Systematic Biology* **50,** 351-366.

HUELSENBECK, J. P. & RANNALA, B. 1997. Maximum likelihood estimation of phylogeny using stratigraphic data. *Paleobiology* **23,** 174-180.

KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* **38,** 7-25.

KLUGE, A. G. & WOLF, A. J. 1993. Cladistics: what's in a word? *Cladistics* **9,** 183-199.

KUHNER, M. K. & FELSENSTEIN, J. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Molecular Biology and Evolution* **11,** 459-468.

LAMBOY, W. F. 1994. The accuracy of the maximum parsimony method for phylogeny reconstruction with morphological characters. *Systematic Botany* **19,** 489-505.

MADDISON, W. P. 1990. A method for testing the correlated evolution of two binary characters: are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* **44,** 539-557.

MADDISON, W. P. & MADDISON, D. R. 1999. *MacClade: Analysis of Phylogeny and Character Evolution, Version 3.08.* Sunderland, MA, Sinauer Associates.

McSHEA, D. W. 1994. Mechanisms of large-scale evolutionary trends. *Evolution* **48,** 1747-1763.

MISHLER, B. D. & THERIOT, E. C. 2000a. The phylogenetic species concept (*sensu* Mishler and Theriot): monophyly, apomorphy, and phylogenetic systematics. **In:** Wheeler, Q. D. & Meier, R. (eds.), *Species Concepts and Phylogenetic Theory*, 44-54. New York, Columbia University Press.

MISHLER, B. D. & THERIOT, E. C. 2000b. A critique from the Mishler and Theriot phylogenetic species concept perspective: monophyly, apomorphy, and phylogenetic species concepts. **In:** Wheeler, Q. D. & Meier, R. (eds.), *Species Concepts and Phylogenetic Theory*, 119-132. New York, Columbia University Press.

MISHLER, B. D. & THERIOT, E. C. 2000c. A defense of the phylogenetic species concept (*sensu* Mishler and Theriot): monophyly, apomorphy, and phylogenetic species concepts. **In:** Wheeler, Q. D. & Meier, R. (eds.), *Species Concepts and Phylogenetic Theory*, 179-184. New York, Columbia University Press.

NELSON, G. & PLATNICK, N. 1981. *Systematics and Biogeography.* New York, Columbia University Press.

NORELL, M. A. & NOVACEK, M. J. 1992a. The fossil record and evolution. Comparing cladistic and paleontologic evidence for vertebrate history. *Science* **255,** 1690-1693.

NORELL, M. A. & NOVACEK, M. J. 1992b. Congruence between superpositional and phylogenetic patterns: comparing cladistic patterns with fossil records. *Cladistics* **8,** 319-337.

O'KEEFE, F. R. & WAGNER, P. J. 2001. Inferring and testing hypotheses of cladistic character dependence by using character compatibility. *Systematic Biology* **50,** 657-675.

PAGEL, M. D. 1994. Detecting correlated evolution on phylogenies: a general method for comparative analysis of discrete characters. *Proceedings of the Royal Society of London Series B* **255,** 37-45.

PAUL, C. R. C. 1982. The adequacy of the fossil record. **In:** Joysey, K. A. & Friday, A. E. (eds.), *Problems of Phylogenetic Reconstruction,* 75-117. New York, Academic Press.

POL, D. & NORELL, M. A. 2001. Comments on the Manhattan stratigraphic measure. *Cladistics* **17,** 285-289.

SCHLOSSER, G. 2001. Using heterochrony plots to detect the dissociated coevolution of characters. *Journal of Experimental Zoology (Mol Dev Evol)* **291,** 282-304.

SIDDALL, M. E. 1996. Stratigraphic consistency and the shape of things. *Systematic Biology* **45,** 111-115.

SIDDALL, M. E. 1997. Stratigraphic indices and tree balance: a reply to Hitchin and Benton. *Systematic Biology* **46,** 569-573.

SIDDALL, M. E. 1998. Stratigraphic fit to phylogenies: a proposed solution. *Cladistics* **14,** 201-208.

SOBER, E. 1988. *Reconstructing the Past.* Cambridge, MA, MIT Press.

SZALAY, F. S. 1977. Ancestors, descendants, sister groups and testing of phylogenetic hypotheses. *Systematic Zoology* **26,** 12-18.

WAGNER, P. J. 1995. Stratigraphic tests of cladistic hypotheses. *Paleobiology* **21,** 153-178.

WAGNER, P. J. 1999. The utility of fossil data in phylogenetic analyses: a likelihood example using Ordovician-Silurian species of the Lophospiridae (Gastropoda: Murchisoniina). *American Malacological Bulletin* **15,** 1-31.

WAGNER, P. J. 2000a. The quality of the fossil record and the accuracy of phylogenetic inferences about sampling and diversity. *Systematic Biology* **49,** 65-86.

WAGNER, P. J. 2000b. Phylogenetic analyses and the fossil record: tests and inferences, hypotheses and models. *Paleobiology* **26** (supplement to 4): 341-371.

WAGNER, P. J. & SIDOR C. A. 2000. Age rank/clade rank metrics – sampling, taxonomy, and the meaning of "stratigraphic consistency." *Systematic Biology* **49,** 463-480.

WILEY, E. O. 1981. *Phylogenetics: The Theory and Practice of Phylogenetic Systematics.* New York, John Wiley and Sons.

WILLS, M. A. 1999. Congruence between phylogeny and stratigraphy: randomization tests and the gap excess ratio. *Systematic Biology* **48,** 559-580.