

ON PHYLOGENETIC RELATIONSHIPS WITHIN *DENDROTRITON* (AMPHIBIA: CAUDATA: PLETHODONTIDAE): IS THERE SUFFICIENT EVIDENCE?

MARK WILKINSON

School of Biological Sciences, University of Bristol, Bristol, BS8 1UG, UK

Previous phylogenetic analyses of the relationships among five Central American salamanders of the genus *Dendrotriton* are reviewed. The available data was reanalysed using parsimony under a variety of analytical treatments. The results are highly sensitive to (1) the coding method used to convert quantitative characters into discrete character states; (2) different scalings (weighting) of multistate characters; and (3) the omission or inclusion of potentially problematic characters. Explorations of length differences between most parsimonious trees and selected less parsimonious alternatives reveal that under each treatment, most parsimonious trees are only marginally more parsimonious than alternatives and that Bremer support for the clades occurring in MPTs is always low. Tree length distributions are not highly left-skewed as would be expected of phylogenetically informative data. These analyses suggest that there is little phylogenetic signal in the available data and that these data provide little basis for well supported phylogenetic inferences. Both parsimony and compatibility-based randomization tests confirm this interpretation. The null hypotheses that the data are not significantly different from phylogenetically uninformative randomly permuted data cannot be rejected for any of the analytical treatments. Given failure to reject the null hypothesis, phylogenetic hypotheses for *Dendrotriton* based on the available data are unconvincing. Additional data are needed. Results of the randomization tests are consistent with the view that there has been extensive homoplasy in bolitoglossine salamanders.

INTRODUCTION

Dendrotriton was established by Wake & Elias (1983) for five nominate species of Central American salamanders that were previously considered part of the *bromeliacia* species group of *Chiropterotriton* beta. Phylogenetic relationships among these species were first investigated by Lynch & Wake (1975) who used numerical techniques to produce three different phylogenetic hypotheses. Relationships within *Dendrotriton* were also briefly addressed, as part of a broader study of the *Chiropterotriton* beta group, by Lynch & Wake (1978). More recently, Collins-Rainboth & Buth (1990) identified a number of problems with Lynch & Wake's (1975) original study, the most important of which they considered (p. 956) "relate to the limitations of the programs available" at that time. They sought an improved estimate of the phylogeny of *Dendrotriton* through the application of more recent parsimony analysis software to a revised interpretation of Lynch & Wake's (1975) data, and they presented a unique most parsimonious tree (MPT), as their single best estimate of the phylogeny of *Dendrotriton*. Here I review these studies, and show through reanalyses of both Lynch & Wake's (1975) and Collins-Rainboth & Buth's (1990) interpretations of the data, and by the application of randomization tests, that no hypothesis of phylogenetic relationships within *Dendrotriton* is well supported by the available data.

LITERATURE REVIEW

Lynch & Wake (1975) compiled data on variation in 11 features (Table 1), six morphometric (1-5, 8), two

meristic (6-7), and three qualitative osteological (9-11), that they used for inferring phylogenetic relationships within *Dendrotriton*. To convert quantitative data into discrete characters, they scored taxa with distinct character states if their sample means (standardized on body length if correlated with it) were statistically significantly different at a level of $P < 0.1$. Converted into discrete characters, their interpretation of the data yielded four binary and seven ordered multistate characters (Table 2). They made *a priori* assessments of character polarity, using the outgroup criterion, for five of their characters (7-11). In addition, they treated the median character state of the remaining morphometric and meristic characters as primitive, based on the assumption that the ancestral species was 'generalized'. Lynch & Wake (1975) experimented with a procedure suggested by Colless (1967) and termed condensation by Crovello (1968), scaling by Swofford (1985) and by Farris (1990) and ranging by Collins-Rainboth & Buth (1990), intended to give each character equal total weight in numerical analyses. They employed two methods of scaling characters, here termed *simple* and *bidirectional* approaches respectively. In the simple approach, the weight of each (adjacent) character state transition is the inverse of the one less than the number of character states of that character, so that all characters have a total weight of unity. In the bidirectional approach, deviations from the assumed primitive condition in different directions each have a total weight of unity.

For their phylogenetic analyses, Lynch & Wake (1975) constructed trees rooted on a hypothetical an-

TABLE 1. Variation in seven morphometric, two meristic and three osteological characters across the five species of *Dendrotriton*. The morphometric data provided are the means and (in parentheses) the 95% confidence intervals of the means. After Lynch & Wake (1975) and incorporating the corrections of Collins-Rainboth and Buth (1990). A = absent, P = present, p = minutely or occasionally present. ¹Reported as present in 2 of 15 specimens (Lynch & Wake, 1975). ²Reported as present in an additional specimen by Lynch & Wake (1978).

Character	<i>bromeliacia</i>	<i>cuchumatanus</i>	<i>megarhinus</i>	<i>rabbi</i>	<i>xolocalcae</i>	Av. confidence interval
sample size	42	9	18	17	38	-
1 Standard length	30.2	28.8	29.7	32.3	30.1	-
2 Tail length	33.2 (0.70)	29.7 (1.20)	32.4 (0.75)	33.2 (1.95)	31.1 (0.65)	1.05
3 Head width	4.60 (0.075)	4.72 (0.115)	4.86 (0.060)	4.73 (0.080)	4.83 (0.065)	0.079
4 Combined limb length	14.9 (0.20)	15.5 (0.25)	14.7 (0.25)	15.7 (0.30)	15.0 (0.20)	0.24
5 Foot width	2.76 (0.065)	2.80 (0.255)	3.07 (0.060)	2.91 (0.110)	2.96 (0.075)	0.113
6 Maxillary teeth	31.7 (1.30)	40.1 (5.10)	49.8 (1.65)	41.9 (2.70)	44.6 (2.40)	2.63
7 Vomerine teeth	10.1 (0.70)	12.7 (2.15)	10.9 (1.05)	17.1 (2.90)	19.9 (0.950)	1.15
8 Nostril diameter	1.89 (0.050)	1.47 (0.100)	2.27 (0.065)	0.72 (0.125)	0.21 (0.020)	0.072
sample size (osteology)	15	1	2	6	2	-
9 Septomaxillae	A ¹	A ²	P	p	A	-
10 Preorbital process of vomer	A	A	A	P	P	-
11 Columellar process	p	P	P	A	p	-

cestor using Wagner parsimony (Kluge & Farris, 1969), with either simple or bidirectional scaling, and the now little-used WISS (weighted invariant step strategy) method of Farris *et al.* (1970), which prohibits evolutionary reversals, using only bidirectional scaling. They also suggested that three of their characters might not be evolutionarily independent, noting that high vomerine tooth counts (character 7) are associated with small nostril sizes (character 8) and the presence of a (dentigerous) preorbital vomerine process (character 10). In view of these concerns, they performed parallel analyses either including or omitting characters (7 and 10). Their six analyses, yielded three distinct trees (Fig. 1, Trees A, B, and C), with the inferred results depending upon method of analysis, the scaling approach used, and the inclusion or exclusion of potentially non-independent characters. The strict component consensus of these three trees is completely unresolved, but there is a single strict reduced cladistic consensus (Wilkinson, 1994) that expresses the agreement among the three trees that *D. megarhinus* is always more closely related to *D. bromeliacia* than to *D. xolocalcae* (Fig. 2).

Lynch & Wake (1975) did not advocate any single, best or preferred phylogenetic hypothesis in the face of their varied results. They attributed the instability of their inferences in part to *D. megarhinus*, noting that it (p. 39) "is, on the average, the most distinctive in overall morphology, i.e. it has the greatest mean phenetic separation from the other species", and that consequently its position in their trees "tends to shift with practically every modification of the clustering procedure". This is not borne out by the strict reduced cladistic consensus (Fig. 2) which indicates that it is *D. cuchumatanus* and *D. rabbi* that have variable positions relative to a more stable phylogenetic framework provided by the other three taxa. They also correctly concluded that (p. 38) "a number of morphological convergences or reversals or both have occurred during the evolution of the group." Most importantly, Lynch & Wake (1975:41) emphasized the limitations of their study by suggesting that "Further refinement of our knowledge of the relationships within this group will depend upon the examination of new suites of characters".

TABLE 2. Lynch & Wake's (1975) data matrix for the five species of *Dendrotriton* and a hypothetical ancestor. The presentation is slightly modified from, but analytically equivalent to Lynch & Wake's (1975: Table 1) bidirectional scaling.

Taxa	Characters										
	1	2	3	4	5	6	7	8	9	10	11
Hypothetical Ancestor	4	0	4	0	4	4	0	0	0	0	0
<i>bromeliacia</i>	4	4	0	0	0	0	4	3	4	4	4
<i>cuchumatanus</i>	0	0	4	4	0	4	4	2	4	4	0
<i>megarhinus</i>	4	4	8	0	8	8	4	4	0	4	0
<i>rabbi</i>	8	4	4	4	4	4	0	1	2	0	4
<i>xolocalcae</i>	2	0	4	0	2	2	0	0	4	0	2

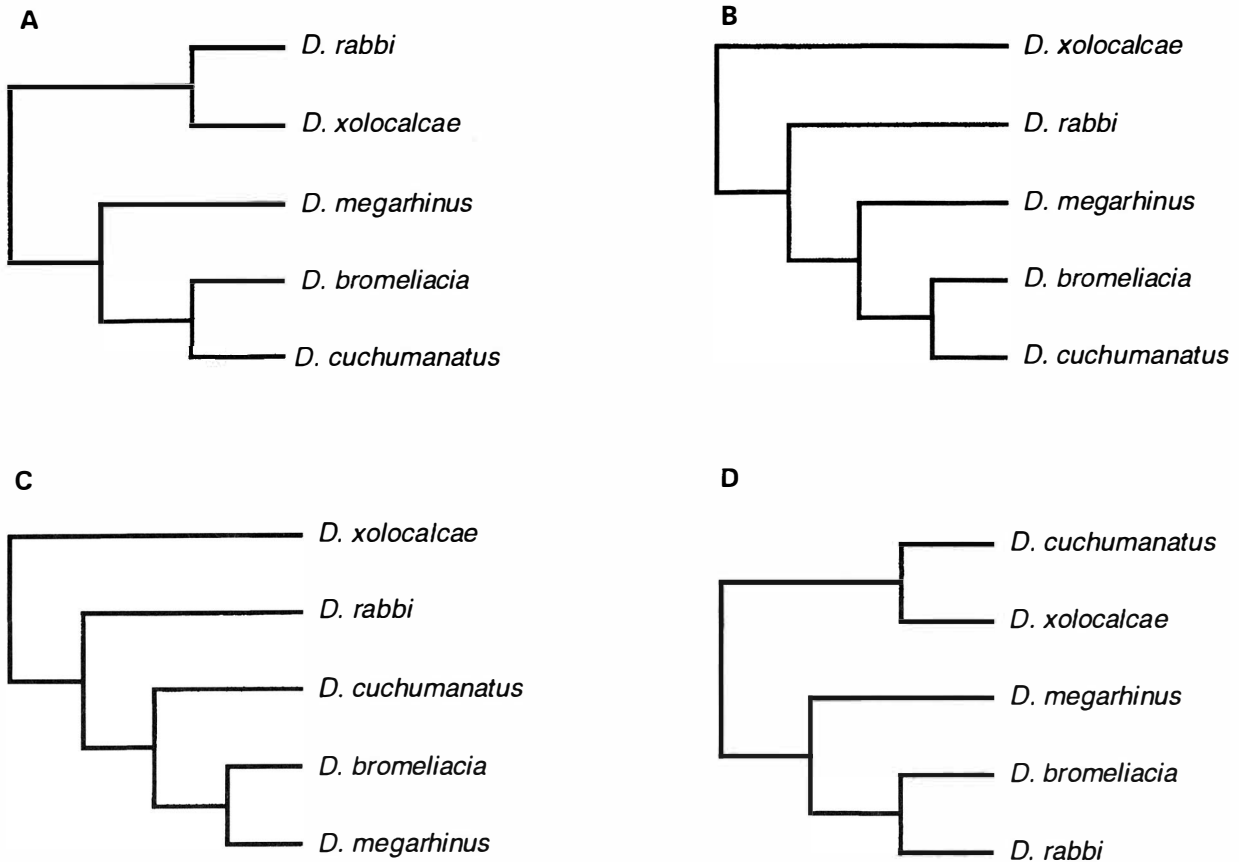


FIG. 1. Four phylogenetic hypotheses for *Dendrotriton* (Trees A - D) proposed in previous studies and the analytical conditions under which they were recovered. Numbers in parentheses indicate omitted characters: A, (Lynch & Wake, 1975) Wagner parsimony, simple bidirectional; B, (Lynch & Wake, 1975) Wagner parsimony, simple (7,10) WISS, bidirectional (7,10); C, (Lynch & Wake, 1975) WISS, bidirectional; D, (Collins-Rainboth and Buth, 1990) Wagner parsimony, no scaling.

Lynch & Wake (1978) presented a tree depicting relationships among eight species of *Chiropterotriton* Beta including the five species of *Dendrotriton* and three species that were subsequently transferred to *Nototriton* (Wake & Elias, 1983), and which, according to Wake & Elias (1983), are not closely related to *Dendrotriton* (but see Sessions & Kezer, 1991). Lynch & Wake's (1978) focus was primarily upon these latter three species, and their tree was not based on any numerical analysis. The relationships within *Dendrotriton* correspond to Tree C of Fig. 1, but no reason for this choice was presented. Lynch & Wake's (1978) study is of interest here primarily for their report that a second

specimen of *D. cuchumatanus* has a pair of distinct septomaxillaries, which has implications for the coding of character 9 of Lynch & Wake (1975).

Collins-Rainboth & Buth's (1990) interpretation and analysis of the data differs from those of Lynch & Wake (1975) in several important respects. Firstly, they did not include standard length (character 1). Secondly, for the remaining morphometric and meristic characters, they used a gap-coding method (Archie, 1985), in which a gap was the average of half the 95% confidence interval of the standardized means of the character for each species. As a result of this coding procedure, they were unable to detect any distinct character states in character 5 (foot width), but detected extra character states in characters 2, 6 and 7. Thirdly, they accepted Lynch & Wake's (1975) assessments of character polarity based on the outgroup criterion, but coded the hypothetical ancestor in their data matrix with missing entries for the remaining characters. The latter characters were therefore unpolarized in their analyses and have no effect upon the placement of the root in resulting trees. Fourthly, they did not employ any scaling of characters, so that each (adjacent) character state transition has equal weight independent of the numbers of character states within characters. Fifthly, they used only Wagner parsimony.

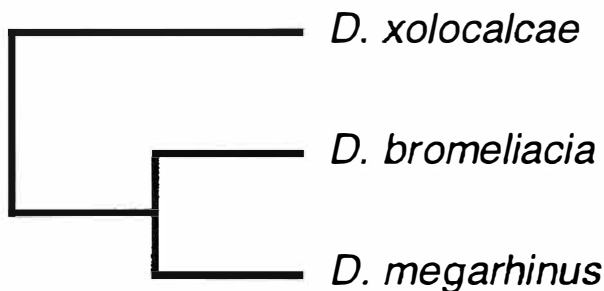


FIG 2. Unique primary reduced cladistic consensus tree for the Trees A-C of Fig. 1.

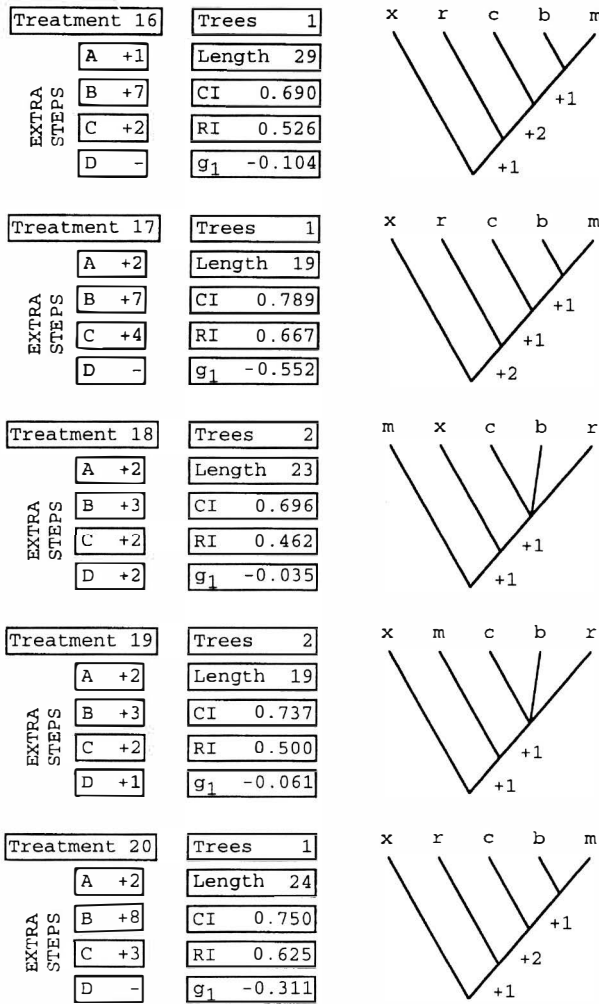


FIG 3. Summary of the parsimony analyses of the five treatments of Collins-Rainboth & Buth's 1990 data including the extra length of trees A-D. Where there are multiple MPTs for a treatment the tree shown is the strict component consensus. b, *D. bromeliacia*; c, *D. cuchumanatus*; m, *D. megarhinus*; r, *D. rabbi*; x, *D. xoloccalcae*; CI, consistency index; RI, retention index. Numbers next to branches are Bremer support values.

Collins-Rainboth & Buth (1990) also corrected three typological errors in Lynch & Wake's (1975) appendix. Of these changes, only one, relating to the condition of the columellar process of the opercular apparatus (character 11), leads to any modification of Lynch & Wake's (1975) data matrix. Lynch & Wake (1975) reported that 3 of 13 specimens of *D. bromeliacia* possessed a columellar process similar to the small but discrete processes of *D. xoloccalcae*, but coded the process as absent. Collins-Rainboth & Buth (1990) changed the coding of *D. bromeliacia* from absent to the intermediate state of occasionally or minutely present.

Despite parallel intraspecific variation, Collins-Rainboth & Buth (1990) proposed no modification to character 9, the condition of the septomaxillae. Their coding, like that of Lynch & Wake (1975), includes the three states: absent, minutely or occasionally present, and present. Lynch & Wake (1975) report that this bone is present in only two of six specimens of *D. rabbi* ex-

amined and it is coded as minutely or occasionally present in both Collins-Rainboth & Buth (1990) and Lynch & Wake (1975). In contrast, *D. bromeliacia*, for which Lynch & Wake (1975) reported that two of fifteen specimens had septomaxillae (large in one specimen and small in the other), was coded as absent in both studies. Thus there is some inconsistency, both in Lynch & Wake (1975) and Collins-Rainboth & Buth (1990) with regard to how intraspecific variation in the columellar process of the opercular apparatus and septomaxillae were handled. Collins-Rainboth & Buth's (1990) coding of the condition of character 9 also takes no account of Lynch & Wake's (1978) observation of septomaxillae in *D. cuchumanatus*, or their conclusion that (p. 294) "Septomaxillaries have not been observed in *C. xoloccalcae*, but experience has shown that these bones are found in at least a small percentage of other members of the group, and the absence here may reflect small sample size." Given that the conditions of both the columellar process of the opercular apparatus and the septomaxillae are known to vary within species, and that sample sizes are small for most of the species, it might be concluded that the boundaries between the character states of these characters and the somewhat arbitrary partitioning of variable species into these character states, are unlikely to reflect phylogeny.

Collins-Rainboth & Buth (1990) used PAUP version 2.4. (Swofford, 1985) to analyse their data matrix. Multistate characters were treated as linear ordered transformation series (i.e. they used Wagner parsimony) and two analyses utilizing either the full data matrix or omitting the osteological characters (9-11) were performed. The latter analysis was performed because, according to Collins-Rainboth & Buth (1990), the original study had explored these treatments (p. 958) "because of the limited number of individuals examined for the latter characters." This is inaccurate. As described above, Lynch & Wake (1975) performed parallel analyses with all the characters or without two of the characters (7 and 10). One of these excluded characters is osteological (10) the other is meristic (7), and the analysis without these characters was performed because Lynch & Wake (1975) considered that they might form part of a functional complex with another character (8). Of course, concern over small sample sizes, as expressed by Lynch & Wake (1975, 1978), may also provide good grounds for experimental omission of the osteological characters.

Both parsimony analyses performed by Collins-Rainboth & Buth (1990) yielded the same single MPT (Fig. 1: Tree D), different from the trees reported by Lynch and Wake (1975), though consistent with the reduced cladistic consensus of the latter trees. They reported that tree lengths among the 105 possible tree topologies (using all the characters) ranged from 29 to 38, that the frequency distribution of lengths of all topologies was not significantly skewed ($g_1 = 0.080$) or

TABLE 3. Collins-Rainboth & Buth's (1990) data matrix for the five species of *Dendrotriton* and a hypothetical ancestor. Character five is invariant and was not included in their analyses or the reanalyses.

Taxa	Characters									
	2	3	4	5	6	7	8	9	10	11
Hypothetical Ancestor	?	?	?	?	?	3	0	2	1	2
<i>bromeliacia</i>	2	0	0	0	0	0	3	0	0	1
<i>cuchumatanus</i>	0	1	1	0	1	1	2	0	0	2
<i>megarhinus</i>	2	2	0	0	3	0	4	2	0	2
<i>rabbi</i>	2	1	1	0	1	2	1	1	1	0
<i>xoloccalcae</i>	1	2	0	0	2	3	0	0	1	1

different from normal, and that the most parsimonious tree was significantly shorter than the mean tree length. In addition, the three tree topologies reported by Lynch and Wake (1975) (Fig. 1: Trees A-C) were shown to have lower consistency indices using their revised character coding. Given these results, Collins-Rainboth & Buth (1990) presented their MPT as the single best, and preferred, hypothesis of the phylogeny of *Dendrotriton*, and they attributed their resolution of this phylogenetic problem to their retroactive application of newer analytical techniques, without any accumulation of the new data that Lynch & Wake (1975) thought necessary.

It should be apparent from this review, that phylogenetic inferences for *Dendrotriton*, have been based on fairly limited data, and that these inferences appear to be sensitive to variation in (1) how qualitative characters are partitioned into character states, (2) scaling (weighting) of discrete characters, (3) method of analysis, and (4) inclusion or exclusion of characters that may not be independent. The instability of the results suggests that the available data may not provide a sufficient basis for robust inferences of phylogenetic relationships within *Dendrotriton*.

Following a suggestion from Fitch (1979), several workers have explored the frequency distributions of tree lengths supported by real, random and simulated data (e.g. Le Quesne, 1989; Hillis, 1991; Huelsenbeck, 1991) and argued that the skewness of such distributions can be used to assess whether the data contains any useful phylogenetic signal. Both theoretical and empirical studies indicate that data containing strong phylogenetic signal are expected to support a strongly left-skewed tree length distribution. The g_1 statistic of Sokal & Rohlf (1981) provides a measure of skewness, and Hillis (1991) has described critical values for a significantly more left-skewed distribution than that yielded by random, phylogenetically uninformative data. For six taxa, a g_1 of less than -0.51 indicates that the data is significantly more left-skewed than random data. Collins-Rainboth & Buth's (1990) reported g_1 for their analysis using all the characters is 0.08, indicating a slight right-skew and an insignificant departure from randomness. Källersjö *et al.* (1992) are highly critical

of using the skewness of tree length distributions to assess data, and have shown that it can give misleading results. However, g_1 values can be taken as suggestive of properties of the data, and are used in this way here. As we shall see, g_1 suggests conclusions that are fully supported by more direct randomization tests of the quality of the data.

MATERIALS AND METHODS

The potential for the available data to support robust inferences of phylogenetic relationships within *Dendrotriton* was assessed through multiple parsimony analyses using PAUP 3.1.1 (Swofford, 1993), and through the application of parsimony and compatibility-based randomisation tests.

Collins-Rainboth & Buth's (1990) data matrix was analysed, as in their study, without any scaling of characters. Also, as in their study, separate analyses were performed using all the characters and with the osteological characters (9-11) omitted. These two treatments replicate their original analysis. In addition, analyses were performed with the characters (7 and 10) considered potentially interdependent by Wake & Lynch (1975) omitted, with the problematic condition of the septomaxillae (character 9) omitted, or with all three of these characters omitted. Lynch & Wake's (1975) data matrix was reanalysed with the same five combinations of characters included or omitted. Parallel analyses were performed with simple and bidirectional scaling of characters, as in the original, and also without any scaling. The total of 20 analytical treatments used are summarized in Table 4.

All analyses used Wagner parsimony with all multistate characters linear ordered. Topological constraints were used to determine the extra length required to overturn clades common to the MPTs, i.e. the clades decay index (Donoghue *et al.*, 1992) or Bremer support (Bremer, 1988; Källersjö *et al.* 1992), and the comparative lengths of Trees A-D. PAUP also provided consistency and retention indices and g_1 statistics.

Each of the 20 different analytical treatments was used in corresponding parsimony and compatibility-

TABLE 4. Results of parsimony and compatibility-based randomization tests for 20 treatments of Lynch and Wake's (1975) and Collins-Rainboth and Buth's (1990) data.

Data	Characters Omitted	Scaling	Parsimony		PC	
			PTP	HER	PTP	IER
1 Lynch & Wake (1975)	None	None	0.62	0.023	0.511	0.000
2 Lynch & Wake (1975)	7 and 10	None	0.30	0.190	0.198	0.160
3 Lynch & Wake (1975)	9 - 11	None	1.00	-0.262	0.889	-0.130
4 Lynch & Wake (1975)	9	None	0.46	0.072	0.337	0.050
5 Lynch & Wake (1975)	7, 9 and 10	None	0.91	-0.099	0.955	-0.271
6 Lynch & Wake (1975)	None	Simple	0.40	0.041	0.715	0.055
7 Lynch & Wake (1975)	7 and 10	Simple	0.64	-0.039	0.527	-0.026
8 Lynch & Wake (1975)	9-11	Simple	0.79	-0.111	0.759	-0.124
9 Lynch & Wake (1975)	9	Simple	0.40	0.042	0.464	0.000
10 Lynch & Wake (1975)	7, 9 and 10	Simple	0.91	-0.148	0.768	-0.689
11 Lynch & Wake (1975)	None	Bidirectional	0.31	0.070	0.629	-0.078
12 Lynch & Wake (1975)	7 and 10	Bidirectional	0.49	0.013	0.476	0.006
13 Lynch & Wake (1975)	9-11	Bidirectional	1.00	-0.357	0.753	-0.010
14 Lynch & Wake (1975)	9	Bidirectional	0.35	0.091	0.424	0.008
15 Lynch & Wake (1975)	7, 9 and 10	Bidirectional	0.95	-0.184	0.702	-0.092
16 Collins-Rainboth & Buth (1990)	None	None	0.53	0.025	0.670	-0.041
17 Collins-Rainboth & Buth (1990)	7 and 10	None	0.15	0.025	0.178	0.190
18 Collins-Rainboth & Buth (1990)	9-11	None	0.99	-0.250	0.988	-0.260
19 Collins-Rainboth & Buth (1990)	9	None	0.20	0.220	0.273	0.084
20 Collins-Rainboth & Buth (1990)	7, 9 and 10	None	0.98	-0.220	0.963	-0.245

based randomization tests. Randomization tests allow the comparison of properties of real data (typically a measure of congruence) to be compared to those of similar but phylogenetically uninformative data. By randomly permuting the assignment of character states to taxa, congruence among the characters is reduced to that expected by chance alone but other features of the original data (numbers of taxa, characters, character states, and taxa in each character state) are unaltered. This allows the null hypothesis that the real data are no more congruent than random and phylogenetically uninformative data to be tested. If the null hypothesis cannot be rejected then the data would seem to provide no compelling basis for preferring any phylogenetic hypothesis.

The parsimony-based randomization test used is that developed independently by Archie (1989a) and by Faith & Cranston (1991) in which the length of most parsimonious trees (MPTs) supported by the original data is compared to the lengths of MPTs. Tree length is thus used as a measure of congruence. Faith & Cranston's (1991) parsimony permutation tail probability (PTP), defined as the proportion of data sets (original and randomly permuted) supporting MPTs as short or shorter than the original provides a test statistic for the null hypothesis and was determined using the Hennig86 (Farris, 1988) parsimony program and 99 randomly permuted data sets. Archie's (1989b) descriptive statistic, the homoplasy excess ratio (HER), was also determined. $HER = (E_L - O_L) / (E_L - M_L)$, where

M_L is minimum length if all characters are congruent, E_L is the expected or mean length of MPTs supported by randomly permuted data, and O_L is the observed length of MPTs supported by the real data.

HER is positive if the real data supports MPTs that are shorter than the expected length for randomly permuted data, has a maximum value of unity when the data includes no incongruence, and approaches zero as the level of incongruence approaches that expected for randomly permuted data. Thus values close to zero or negative values indicate that the data show little or no more congruence than expected by chance alone.

Faith & Cranston (1991) suggested that where the focus of the test is ingroup relationships, hypothetical ancestors or outgroups should be excluded from the random permutation, and thus maintained unaltered in all the randomly permuted data sets. I have followed this suggestion, but note that random permutation of the ingroup only will make any single hypothetical ancestor or outgroup essentially random with respect to the remaining taxa and has little or no impact upon test results (pers. obs). Faith & Cranston's (1991) suggestion is expected to be more important in cases where it would preserve non-random relations among multiple outgroups.

The compatibility-based randomization test used was developed independently by Wilkinson (1992) and Alroy (1994), and uses the number of pairwise (in)compatibilities among characters in the data as a measure of congruence. It yields a pairwise compat-

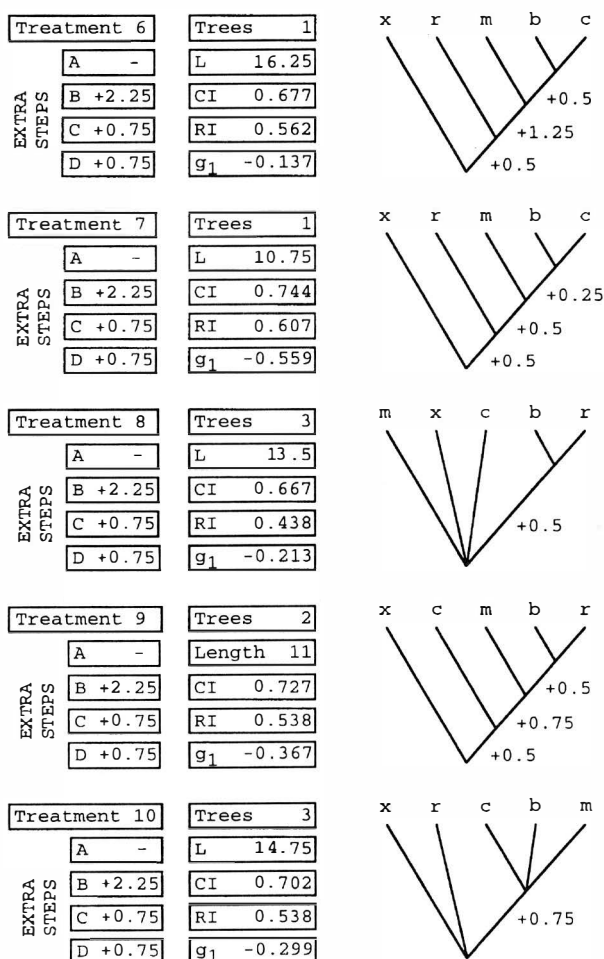


FIG. 4. Summary of the parsimony analyses of the five treatments of Lynch & Wake's (1975) data using simple scaling. Format and abbreviations as in Figure 3.

ibility (PC) PTP test statistic, defined as the proportion of data sets (original and randomly permuted) with as few or fewer pairwise incompatibilities as the original data. The PCPTP is analogous to Faith & Cranston's (1991) parsimony PTP, but differs most importantly in that the measure of congruence is not also a criterion of hypothesis choice. Thus the PCPTP test is tree or hypothesis independent (Alroy, 1994). An incompatibility excess ratio (IER) analogous to Archie's (1989b) HER was also determined. $IER = (E_r - O_r) / E_r$, where E_r is the expected or mean number of pairwise incompatibilities of randomly permuted data and O_r is the number of pairwise incompatibilities for the original data. IER is positive if the real data supports MPTs that are shorter than the expected length for randomly permuted data, has a maximum value of unity when the data includes no incongruence, and approaches zero as the level of incongruence approaches that expected for randomly permuted data. Thus values close to zero or negative values indicate that the data show little or no more congruence than expected by chance alone.

The use of g_1 to assess the phylogenetic signal in cladistic data is somewhat parallel to the parsimony and compatibility-based randomization tests used here. It

differs in that skewness is a much more indirect and unsatisfactory measure of congruence, and that Hillis's (1991) critical values were determined using randomly generated data, whereas random permutation produces critical values that are specific to the data at hand.

To facilitate the compatibility randomization tests and the parsimony analyses with scaling of multistate characters, the data were recoded into an analytically equivalent additive binary form. Interdependent binary factors (either of an originally multistate character or representing character weighting) cannot be incompatible or incongruent with each other and were 'tied' in the randomization tests so that they were randomly permuted with respect to other characters, but not with respect to each other.

RESULTS

Results of the parsimony analyses using Collins-Rainboth & Buth's (1990) interpretation of the data are summarized in Fig. 1. The treatments paralleling their original analyses (Treatments 16 and 17) replicate the original results (except for some discrepancies in reported consistency indices and g_1 statistics). In each of these treatments, Tree D is the single MPT and each of Lynch & Wake's (1975) three trees are less parsimonious. Note, however, that using all characters (Treatment 16) Tree A requires only a single extra step, and that Bremer support is low for all clades in the MPT. Excluding the osteological characters (Treatment 17) slightly increases the additional length required by Tree A, but Bremer support remains low for all clades. The same single MPT and parallel statistics are also produced by the omission of just character 9 (Treatment 20).

Treatment 18, in which the potentially interdependent characters identified by Lynch & Wake (1975) are omitted, yields rather different results. There are two MPTs neither of which corresponds to Trees A - D. These trees are unusual in placing *D. megarhinus* as the sister taxon of all other *Dendrotriton*. However, Bremer support is low for the clades common to both MPTs, and none of Trees A - D require more than three additional steps. Similarly, Treatment 19 yields two more distinct MPTs and comparable statistics.

Comparative results for the parsimony analyses using Lynch & Wake's (1975) interpretation of the data, using simple scaling, are summarized in Fig. 4. As in their study, analysis using all characters (treatment 6), yielded Tree A. Paralleling Collins-Rainboth & Buth's (1990) results, omission of the osteological characters (treatment 7) also yields the same unique MPT. Tree A is also one of three MPTs when just character 9 is omitted (treatment 10). However, in all treatments, Bremer support for all clades is low and one or more of Trees B - D are only marginally less parsimonious. In Lynch & Wake's (1975) study treatment 8 (omission of characters 7 and 10) yielded Tree B. In the reanalysis, Tree B is one of three MPTs, and the MPT recovered in treatment 9 represents an additional distinct MPT. Again,

neither of these treatments yield MPTs that are much shorter than one or more of their less parsimonious competitors among Trees A-D, and Bremer support for all clades is low. In all treatments g_1 is low.

Essentially similar results were found in the analyses using Lynch & Wake's (1975) interpretation of the data under bidirectional scaling or without scaling (data not shown), with MPTs depending upon which characters are included, always being only marginally more parsimonious than competitors from among Trees A-D, and with all clades having low Bremer support. In all treatments of both Lynch & Wake's (1975) and Collins-Rainboth & Buth's (1990) interpretations of the data, g_1 is either positive or slightly negative reflecting tree length distributions that are never strongly left-skewed as would be expected of phylogenetically informative data.

Results of the parsimony and compatibility-based randomization tests are summarized for all treatments in Table 4. In all cases, both parsimony PTPs and PCPTPs are not significant (> 0.05) and therefore do not allow the null hypothesis that the data are random with respect to phylogeny to be rejected under any of the diverse character coding and scaling (weighting) schemes used. For all analyses, the descriptive HER and IER statistics are close to zero or negative and also betray a lack of congruence within the data that might distinguish them from the expectations for phylogenetically uninformative, randomly permuted data.

DISCUSSION

Collins-Rainboth & Buth (1990:955) considered that Lynch & Wake's (1975) "failure to resolve a single most parsimonious cladogram for this group may not be due to limitations of their data, but rather to confusion with regard to the nature of the analyses and limitations of the algorithms available to them at that time." In particular, they noted that the programs available to Lynch & Wake (1975) required the input of ancestral character states, so that all characters had to be polarized, that they were limited to generating a single MPT rather than finding all such trees, and that the order of data input may have influenced clustering. They believed that their reanalyses solved these problems through the recoding of the data and use of more modern software that was free from these limitations. They aimed to support one of Lynch & Wake's (1975) three trees (Trees A-C), but their analysis yielded a different MPT (Tree D).

A major difference between Collins-Rainboth & Buth's (1990) and Lynch & Wake's (1975) analyses is in the coding of the hypothetical ancestor which was used to root the trees. Collins-Rainboth & Buth (1990) accepted Lynch & Wake's (1975) polarity assessments based on outgroup comparisons but rejected those based on the assumption that the ancestor *cx* was generalized, although (p. 957) they considered this to be "a reasonable supposition". Most phylogeneticists accept

that outgroup comparisons provide the best single guide to character polarity, thus in rejecting other inferences of polarity Collins-Rainboth & Buth (1990) adopt a sensible cautious approach that may have wrought some improvement over Lynch & Wake's (1975) character coding. However, this appears to have little practical effect. If Collins-Rainboth & Buth's (1990) interpretation of the data is amended by including all of Lynch & Wake's (1975) polarity assessments, Tree D remains the single MPT for the revised data, with no change in tree length.

A second major difference results from Lynch and Wake's (1975) experimentation with scaling their multistate characters. Collins-Rainboth & Buth (1990) weighted all characters equally, which is the commonest approach in numerical phylogenetics, but they did not discuss or attempt to justify their rejection of scaling. Farris (1990) has recently argued against scaling of phylogenetic data but cogent arguments for scaling have been presented by Thiele (1993). Use of scaling relates to the thorny issue of differential weighting of evidence, attitudes to which tend to be highly polarized. However, we would expect good phylogenetic data to support inferences that are insensitive to differences in potentially reasonable weighting strategies. Conversely, instability is indicative of limitations of the data, and the tentative nature of inferences based upon any preferred weighting scheme (Wilkinson and Benton, 1996). In practice, both Lynch & Wake's (1975) and Collins-Rainboth & Buth's (1990) interpretations of the data yield Tree A and Tree D respectively, irrespective of whether simple scaling or no scaling is employed. Thus, differences in scaling do not account for their different results. In contrast, bidirectional scaling, which is probably the least satisfactory of the methods used by Lynch & Wake (1975), particularly as it depends upon uncertain polarity assessments, does yield different trees.

Differences in the software used also do not fully explain the differences in Lynch & Wake's (1975) and Collins-Rainboth & Buth's (1990) results. Reanalysis using Lynch & Wake's (1975) interpretation of the data shows that the software they employed did not prevent them from correctly identifying the MPTs supported by their data. A single exception is their analysis using simple scaling with characters 7 and 10 omitted (treatment 8). Here the original analysis yielded only one of three MPTs.

A third major difference is in the method used to convert qualitative data into discrete characters. Many such methods have been proposed, though none are immune from criticism (Archie, 1985; Farris, 1990; Thiele, 1993). Collins-Rainboth & Buth (1990) presented no reasons for preferring their method over that of Lynch & Wake (1975), although their different results probably depend, at least in part, upon the different approaches. In practice, Collins-Rainboth & Buth's (1990) method fails to find any discrete character states for one of Lynch & Wake's (1975) characters,

but finds additional character states for three others. Thus, neither approach can be categorized as the generally more conservative and, judged on their randomization test results, neither approach can claim to have yielded a more phylogenetically informative data set.

Other differences include Collins-Rainboth & Buth's (1990) minor change of coding of the septomaxillae and the omission of character 1 (standard length) from their reanalysis. This latter difference was also not discussed, but presumably reflects a desire to avoid non-independence (and thus overweighting) of this and other size related characters. In practice, this character is phylogenetically uninformative under parsimony and its omission can have no effect. In contrast, Collins-Rainboth & Buth (1990) showed no such concern for the possible independence of character 7, 8 and 10 discussed by Lynch & Wake (1975) and their recoding of character 9 ignored Lynch & Wake's (1978) observations and concerns.

In summary, the differences between Lynch & Wake's (1975) and Collins-Rainboth & Buth's (1990) results appear to depend primarily upon a few differences in their character coding, rather than, as Collins-Rainboth & Buth's (1990) believed, their use of more advanced software. Furthermore, neither of Lynch & Wake's (1975) or Collins-Rainboth & Buth's (1990) interpretations of the data are free from problems, and neither is demonstrably better than the other. The sensitivity of the results to character coding indicates which, if any, phylogenetic hypothesis is best supported by the observed variation is not clear cut.

Having found a single MPT supported by their revised data Collins-Rainboth & Buth (1990) addressed which of the four trees (Trees A-D) represent the best estimate of the phylogeny of *Dendrotriton*. Although they noted that the trees resulted from analyses that differed in coding of the data, scaling (weighting) and, in the case of Tree C, in method of analysis, they considered that the consistency indices for trees, based on their unscaled and revised character data, provided an appropriate comparative measure. Not surprisingly, consistency indices for Trees A-C were lower than those for Tree D (because they are not MPTs for these data), and on this basis, they proposed Tree D as the single best estimate of the phylogeny of *Dendrotriton*. However, and conversely, under Lynch & Wake's (1975) character coding and scaling Tree D is not most parsimonious. Thus, their preference for Tree D rests upon their preference for their character coding. Collins-Rainboth & Buth's (1990) enthusiasm for Tree D was not tempered by the observation that quite different trees result from minor differences in the treatment of the same underlying data or the observation that the differences in the consistency indices they considered are only marginal for Trees A and C.

Their confidence in Tree D was also strengthened by its insensitivity to the omission of the osteological characters. My reanalyses confirm this and also

demonstrate insensitivity to the omission of just the problematic character 9 (septomaxillae). Collins-Rainboth & Buth (1990) claimed to be following Lynch & Wake (1975) in performing their parallel analysis. However, as we have seen, Lynch & Wake (1975) omitted a different set of characters (7 and 10) because of their possible interdependence with character 8. If these characters are omitted from Collins-Rainboth & Buth's (1990) revised interpretation of the data (Fig. 3, treatment 18), there are dramatic changes in the MPTs supported by their data. Non-independence is a potentially serious problem because it can lead to overweighting of misleading phylogenetic signal and a false sense of confidence (Wilkinson, 1995). In as much as Lynch & Wake's (1975) concerns are reasonable, sensitivity of Collins-Rainboth & Buth's (1990) preferred tree to the omission of characters 7 and 10 must diminish confidence in that tree.

Collins-Rainboth & Buth's (1990) demonstration that Tree D is significantly shorter than the mean length of all trees appears to be intended to support their confidence in that tree, but this result is not surprising and has little phylogenetic significance. For example, the same significant difference obtains in the analyses of Lynch & Wake's (1975) data. What is more suggestive is the shape of the tree length distributions and g_1 , which, in their analyses and in all treatments in my reanalyses, suggests that the data do not contain strong phylogenetic signal.

Results from the reanalyses using Lynch & Wake's (1975) interpretation of the data show a strong parallel to those for Collins-Rainboth & Buth's (1990) revised interpretation. In both analyses there is a single MPT supported by the full data, and insensitivity to the omission of some characters, but sensitivity to the omission of others, particularly those that may lack independence. In each case Bremer support for clades is low, differences in tree length between MPTs and one or more alternatives from among Trees A - D are marginal and g_1 statistics are unimpressive. The parallel suggests that Tree D is no better supported by Collins-Rainboth & Buth's (1990) interpretation of the data than are Trees A, B and C by some treatments using Lynch & Wake's (1975) interpretation.

A cautious view of the results of the reanalyses is that the sensitivity of phylogenetic inferences to variations in coding of data and analytical treatment, low Bremer support values and insignificant g_1 statistics imply that the data contains only weak (if any) phylogenetic signal, and that phylogenetic hypotheses based on these data should be viewed sceptically. This conclusion is fully supported by the results of the randomization tests. In none of the twenty treatments of the data can it be distinguished, based on tree length or compatibility, from randomly permuted data. That the null hypothesis that the data is phylogenetically uninformative cannot be rejected entails that hypotheses based on the data should be invested with no more confidence than a hypothesis based on random data or a

randomly selected hypothesis, and that the available data are simply insufficient to resolve the phylogeny of *Dendrotriton*.

Collins-Rainboth & Buth (1990:960) commented that "Lynch and Wake (1975) predicted that further refinement of the knowledge of relationships within this group might come from the acquisition of new suites of characters, especially those generated using molecular technology. More data may strengthen our basis for estimation of relationships. However, advances in analytical methods may be as, or more important as technological advances in the generation of data." The present study reveals that Collins-Rainboth & Buth's (1990) results, and differences from those of Lynch & Wake's (1975), reflect more their modified interpretation of the data than they do advances in analytical methods and support Lynch & Wake's (1975) view that more data are needed.

Randomization tests are themselves a relatively recent addition to the analytical methods available in numerical phylogenetics, reflecting the increasing concern for the strengths and weaknesses of phylogenetic hypotheses and a shift away from more blinkered search for and acceptance of MPTs. As yet, they have not been widely applied, and some attitudes toward them are dismissive (e.g. Carpenter, 1992). However, this and other studies (e.g. Archie, 1989c; Faith, 1990) have identified real data that cannot be distinguished from randomly permuted data, with serious implications for the assessment of phylogenetic inferences based on the data.

The major obstacle to phylogenetic inference is the misleading evidence provided by homoplasy (Wilkinson, 1991). Interestingly, Wake (1991) has argued that homoplasy is rife among bolitoglossine salamanders. Three possible causes of this are (1) high rates of evolution in the variable characters of the group; (2) relatively rapid cladogenesis and thus short interior branch lengths such that most change occurs independently in terminal branches; and (3) a combination of 1 and 2. In as much as randomly permuted data are consistent with all three models of evolution, failure to discriminate between the data for *Dendrotriton* and random permutations of that data does not allow these models of evolution to be rejected, and supports Wake's (1991) view that there are high levels of homoplasy in bolitoglossine salamanders.

ACKNOWLEDGEMENTS

I am grateful to Donald Buth for his constructive comments. This work was supported in part by SERC GR/F 87912. The software used for randomly permuting data and computing parsimony and PCPTs are available from the author upon receipt of a (PC formatted) disc.

REFERENCES

- Alroy, J. (1994). Four permutation tests for the presence of phylogenetic structure. *Syst. Biol.* **43**, 430-437.
- Archie, J. W. (1985). Methods for coding variable morphological features for numerical taxonomic analysis. *Syst. Zool.* **34**, 326-345.
- Archie, J. W. (1989a). A Randomiation test for phylogenetic information in systematic data. *Syst. Zool.* **38**, 239-252.
- Archie, J. W. (1989b). Homoplasy excess ratios: New indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. *Syst. Zool.* **38**, 253-269.
- Archie, J. W. (1989c). Phylogenies of plant families: a demonstration of phylogenetic randomness in DNA sequence data derived from proteins. *Evolution* **43**, 1796-1800.
- Bremer, K. (1988). The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**, 795-803.
- Carpenter, J. M. (1992). Random cladistics. *Cladistics* **8**, 147-153.
- Colless, D. H. (1967). An examination of certain concepts in phenetic taxonomy. *Syst. Zool.* **16**, 6-27.
- Collins-Rainboth, A. & Buth, D. G. (1990). A reevaluation of the systematic relationships among species of the genus *Dendrotriton* (Caudata: Plethodontidae). *Copeia* **1990**, 955-960.
- Crovello, T. J. (1968). The effect of alteration of technique at two stages in a numerical taxonomic study. *Univ. Kansas Sci. Bull.* **47**, 761-768.
- Donoghue, M. J., Olmstead, R. G., Smith, J. F., & Palmer, J. D. (1992). Phylogenetic relationships of Dipsacales based on *rbcL* sequence data. *Ann. Missouri Bot. Gard.* **80**, 672-685.
- Faith, D. P. (1990). Chance marsupial relationships. *Nature* **345**, 393-394.
- Faith, D. P. & Cranston, P. S. (1991). Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. *Cladistics* **7**, 1-28.
- Farris, J. S. (1988). *Hennig86. Version 1.5. Software and Documentation*. Distributed by the author. Port Jefferson Station, New York.
- Farris, J. S. (1990). Phenetics in camouflage. *Cladistics* **6**, 91-100.
- Farris, J. S., Kluge, A. G. & Eckhardt, M. J. (1970). A numerical approach to phylogenetic systematics. *Syst. Zool.* **19**, 172-189.
- Fitch, W. M. (1979). Cautionary remarks on using gene expression events in parsimony procedures. *Syst. Zool.* **28**, 375-379.
- Hillis, D. M. (1991). Discriminating between phylogenetic signal and random noise in DNA sequences. In *Phylogenetic analysis of DNA*

- sequences, 278-294. Miyamoto, M. M. and Cracraft, J. (Eds.). Oxford: OUP.
- Huelsenbeck, J. P. (1991). Tree-length distribution skewness: an indicator of phylogenetic information. *Syst. Zool.* **40**, 257-270.
- Källersjö, M., Farris, J. S. Kluge, A. G. & Bult, C. (1992). Skewness and permutation. *Cladistics* **8**, 275-287.
- Kluge, A. G. & Farris, J. S. (1969). Quantitative phyletics and the evolution of anurans. *Syst. Zool.* **18**, 1-32.
- Le Quesne, W. J. (1989). Frequency distributions of lengths of possible networks from a data matrix. *Cladistics* **5**, 395-407.
- Lynch, J. F. & Wake, D. B. (1975). Systematics of the *Chiropterotriton bromeliacia* group (Amphibia: Caudata) with descriptions of two new species from Guatemala. *Contrib. Sci. Nat. Hist. Mus. Los Angeles Co.* **265**, 1-45.
- Lynch, J. F. & Wake, D. B. (1978). A new species *Chiropterotriton* (Amphibia: Caudata) from Baja Verapaz, Guatemala, with comments on relationships among Central American species of the genus. *Contrib. Sci. Nat. Hist. Mus. Los Angeles Co.* **294**, 1-22.
- Sessions, S. K. & Kezer, J. (1991). Evolutionary cytogenetics of bolitoglossine salamanders (Family Plethodontidae). In *Amphibian Cytogenetics and Evolution* 89-129. S. K. Sessions and Green, D. M. (Eds.). San Diego: Academic Press.
- Sokal, R. R. & Rohlf, F. J. (1981). *Biometry*. San Francisco: Freeman.
- Swofford, D. L. (1985). *PAUP: Phylogenetic analysis using parsimony*. Version 2.4. Champaign, Illinois: Illinois Natural History Survey.
- Swofford, D. L. (1993). *PAUP: Phylogenetic analysis using parsimony*. Version 3.1.1. Champaign, Illinois: Illinois Natural History Survey.
- Thiele, K. (1993). The holy grail of the perfect character: the cladistic treatment of morphometric data. *Cladistics* **9**, 275-304.
- Wake, D. B. (1991). Homoplasy: the result of natural selection, or evidence of design limitations. *Am. Nat.* **138**, 543-567.
- Wake, D. B. & Elias, P. 1983. New genera and a new species of Central American salamanders, with a review of the tropical genera (Amphibia, Caudata, Plethodontidae). *Contrib. Sci. Nat. Hist. Mus. Los Angeles Co.* **345**, 1-19.
- Wilkinson, M. (1991). Homoplasy and parsimony analysis. *Syst. Zool.* **40**, 107-109.
- Wilkinson, M. (1992). *Compatibility, consensus and missing data in phylogenetic inference*. Unpubl. Ph.D. Thesis. University of Bristol, Bristol.
- Wilkinson, M. (1994). Common cladistic information and its consensus representations: reduced Adams and reduced cladistic consensus trees and profiles. *Syst. Biol.* **43**, 343-368.
- Wilkinson, M. (1995). A comparison of two methods of character construction. *Cladistics* **11**, 297-308.
- Wilkinson, M. & Benton, M. J. (1996). Sphenodontid phylogeny and the problems of multiple trees. *Phil. Trans. R. Soc. B* **351**, 1-16.

Accepted: 8.11.96