Reply: Xantusiid Lizards and Phylogenetic Inference

To the Editor:

Crother *et al.* (1986) performed a parsimony analysis on some published morphological data of xantusiid lizards and concluded that the Cuban endemic genus *Cricosaura* is the sister group to the Middle American genus *Lepidophyma* and that *Klauberina* (California Islands) is a sister group to *Xantusia* (United States and Mexico). Hedges *et al.* (1991) presented DNA sequence data from two mitochondrial genes supporting a different phylogeny: *Lepidophyma* and *Xantusia* as sister genera (bootstrap *P* value = 99%), with the species *Xantusia riversiana* ("*Klauberina*") nestled within *Xantusia* and as a sister species (*P* = 100%) to *Xantusia vigilis*. Crother and Presch (1992) reanalyzed our DNA sequence data, along with the morphological data, and concluded that the phylogeny of xantusiid lizards is "not yet a robust hypothesis." In our reply (Hedges and Bezy, 1993), we (1) pointed out the strong statistical support for the results of our earlier study, (2) investigated the pivotal morphological characters used by Crother *et al.* (1986) and Crother and Presch (1992) and found them to be incorrectly scored, thus putting into question their morphological analysis, and (3) presented new DNA sequence data from a third

mitochondrial gene (16S rRNA) that support the same phylogenetic relationships that we obtained earlier.

Crother and Presch (1994) now have accepted our phylogeny for xantusiid genera and our criticisms of their morphological analysis, but wish to debate further some methodological aspects of phylogenetic analysis. Our responses to each of their points are listed below.

Morphology and Linear Transformation Series

The reason as given by Hedges and Bezy (1993, p. 77) for questioning the informativeness of four of the five multistate characters used by Crother and Presch (1992) is not that the characters were ordered into transformation series, but rather that they are not informative (for parsimony analysis) by the criterion of having two or more states that are present in two or more OTUs (each state), and thus their only information content is derived from their *a priori* ordering. Slowinski (1993, p. 163) recommended using ordered characters when one character state tree is "clearly superior to others." Reanalyses of the xantusiid morphological data indicate that ordering these particular multistate characters does not increase phylogenetic resolution or "congruence" (Houser and Presch, 1991, Table 2; Slowinski, 1993, Table 1). We continue to assert that counting these four characters as informative for parsimony analyses is problematic.

Distant Outgroups

Crother and Presch (1994) have criticized our use of a bird (*Gallus*) as the root for some of our phylogenetic trees, noting that such a distant outgroup "may be equivalent to a random collection of character states and not a reflection of history." However, they apparently have overlooked the distance scales in our figures (Hedges and Bezy, 1993, Figs. 3 and 4) which indicate much less divergence than would be expected based on random variation. For example, the average uncorrected distance (p) between *Gallus* and the ingroup taxa for our entire data set is 0.276, compared with 0.750 which would be expected in a random situation. As can be seen by examining our sequence alignment (Hedges and Bezy, 1993, Fig. 2), these are relatively conserved genes and the variation is clearly not random. These genes useful for investigating relationships not only among amniotes, but also among vertebrates in general (e.g., Hedges *et al.*, 1993). Furthermore, our analyses also included a closer outgroup (teiid lizard), and the results using the two different outgroups in the distance analyses of the complete data set were identical (Hedges and Bezy, 1993, Fig. 4).

Neighbor-Joining Phylogeny Estimates

Crother and Presch (1994) state that "given that data sets used to reconstruct phylogeny are small, the probability of accurately recovering the singular [phy-

logenetic] pattern is low." This is not true. The probability of recovering the true phylogeny depends on many factors besides the size of the data set, such as number of taxa, topology of the true tree, rate of change of the characters, method of analysis, etc., and these have been investigated in computer simulations (see review in Nei, 1991). In many cases, this probability can be quite high. Crother and Presch (1994) continue to confuse (see also Crother and Presch, 1992) the method of maximum parsimony with other methods of analysis such as neighbor-joining. The comparison of alternative topologies is built into the neighbor-joining algorithm (Saitou and Nei, 1987). The minimum evolution method (e.g., Rzhetsky and Nei, 1992) also can be used to examine the sum (S) of branch lengths for different topologies, and as we stated, this was done in our analysis (Hedges and Bezy, 1993).

The Bootstrap

Crother and Presch (1994) state (incorrectly) that we "used the bootstrap as a minimum evolution method, to choose among most or near most-parsimonious trees. . . ." Here, they have confused a method for applying statistical confidence (bootstrap) with two very different methods of phylogenetic analysis (minimum evolution and maximum parsimony). They state that "most" of our "uses of the bootstrap are inappropriate," but do not detail which are appropriate and which are not (and specifically why they are or are not appropriate). Therefore it is difficult to reply to such vague criticisms.

The quote taken from Hillis and Bull (1993) is not a criticism of the bootstrap, but a statement of the obvious: that the value of the test depends on the data and method of analysis used. Clearly it is not a magical test that can make probability statements about a true phylogeny that are independent of the data at hand. That the bootstrap is a conservative test (Zharkikh and Li, 1992) means that actual P values in our analysis of xantusiid phylogeny are likely to be higher than those reported. As the P values for most of the nodes in the tree (Hedges and Bezy, 1993, Fig. 4) already exceed 95%, this indicates that the tree is very strongly supported by the data.

By contrasting the bootstrap test with consistency index, tree length, and "alternative phylogenetic hypotheses," Crother and Presch (1994) imply that the latter are statistical tests of phylogenetic inference. This is incorrect. The consistency index is a measure of the amount of character conflict (homoplasy) in a parsimony tree and the length of a tree is simply the quantity being optimized in a parsimony analysis. Examining the lengths of trees corresponding to alternative phylogenetic hypotheses is a useful observation, but if they differ, is it a significant difference? Crother

and Presch (1994) appear to misunderstand what constitutes a statistical test.

Tree Length

Crother and Presch (1994) state that "finding the simplest explanation (i.e., the shortest tree) and holding it as the best estimate is the very essence of the parsimony paradigm." This is simply restating the obvious, but raises the question as to whether the shortest tree is significantly shorter than other topologies. Where is the line drawn: one step longer, two steps longer, three steps longer? Statistical methods that allow one to place error estimates on phylogenies, such as the bootstrap method (Felsenstein, 1985) and the standard error test for branch lengths (Rzhetsky and Nei, 1992), provide much more information for drawing conclusions from a phylogenetic analysis than simply finding the shortest tree or trees.

Insertions and Deletions

Crother and Presch (1994) suggest that we have overlooked the phylogenetic information in gaps, but this is not true. As we pointed out (Hedges and Bezy, 1993, p. 85), the insertions and deletions (examined separately) support the same topology that we obtained with only nucleotide data.

Transitions and Transversions

Crother and Presch (1992) weighted transversions more heavily than transitions in their reanalysis of our DNA sequence data. We detailed our criticisms of this approach elsewhere (Hedges and Bezy, 1993; see also Hedges and Maxson, 1992) and refer the reader to those discussions, but will respond briefly to some points raised by Crother and Presch (1994).

Crother and Presch (1994) cite Marshall (1992) and Fitch (1992) in support of weighting all sequence data for parsimony analysis. However, Hedges and Maxson (1992) pointed out the theoretical problem with that approach. It is undesirable to apply the same weighting scheme to sequence data that have not diverged greatly because they have not incurred multiple hits. Weighting data in those cases will reduce the information content of the sequence data and bias the analysis in favor of particular substitution types. At the other end of the spectrum, rare substitution types also can undergo multiple hits, and giving high weight to the those "rare" substitution types does not help estimate the number of such multiple hits. Because any given sequence data set can span this spectrum of variation, these theoretical problems with weighting in parsimony analysis apply to virtually all data sets. Multiple hit corrections for distance data (e.g., Jukes

and Cantor, 1969; Kimura, 1980) are scaled and therefore avoid that problem.

Consistency Index

Crother and Presch (1994) state that "the use of the consistency index for comparison of trees is not problematical" and refer us to their previous paper for discussion. However, in their previous paper, they state that "because of these problems [with the consistency index] the Retention Index (Farris, 1989a, b) is a superior method of character data fit to a tree" (Crother and Presch, 1992, p. 291). The reason that the consistency index is recognized to be problematical (Forey et al., 1992) is because it is correlated with both the number of taxa and the number of characters, and because the value is commonly inflated by including uninformative characters (e.g., Crother et al., 1986).

Total Evidence and Weighting

In order to prevent the large molecular data set from "swamping out" the much smaller morphological data set, Crother and Presch (1992) weighted the morphological characters in inverse proportion to their contribution in the combined data set. We criticized that approach because of its arbitrary nature and the greatly inflated weight assigned to each morphological character. Crother and Presch (1994) state that they followed Kluge (1983). However, Kluge (1983) did not suggest such an approach. He suggested comparing the results of different classes of data in search for "consilience" rather than combining those data in one analysis with differential weights applied to the different data sets. Later, Kluge (1989) suggested combining molecular and morphological data, but not in the manner (inverse weighting) that Crother and Presch (1992) have done. In citing other studies critical of inverse weighting (e.g., Barrett et al., 1991; Jones et al., 1993), Crother and Presch (1994) now appear to have serious reservations about this approach.

Conclusion

Crother and Presch (1994) appear to have accepted our phylogeny of xantusiid genera (Hedges et al., 1991; Hedges and Bezy, 1993) and our criticisms of their morphological analysis. They also now appear to have serious reservations about their previous approach of assigning differential weights to morphological and molecular characters in a combined data set. However, the analytical points raised by them still indicate a confusion of different methodologies and a lack of appreciation for the crucial role that statistical methods play in phylogenetic inference.

REFERENCES

Barrett, M., Donoghue, M. J., and Sober, E. (1991). Against consensus. Syst. Zool. 40: 486–493.

Crother, B. I., Miyamoto M. M., and Presch, W. F. (1986). Phylogeny and biogeography of the lizard family Xantusiidae. Syst. Zool. 35: 37–45.

Crother, B. I., and Presch, W. F. (1992). The phylogeny of xantusiid lizards: The concern for analysis in the search for a best estimate of phylogeny. Mol. Phylogenet. Evol. 1:289–294.

Crother, B. I., and Presch, W. F. (1994). Xantusiid lizards, concern for analysis, and the search for the best estimate of phylogeny: Further comments. Mol. Phylogenet. Evol. 3.

Farris, J. S. (1989a). The retention index and homoplasy excess. Syst. Zool. 38: 406–407.

Farris, J. S. (1989b). The retention index and the rescaled consistency index. Cladistics 5: 417–419.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791.

Fitch, W. M. (1992). Commentary on letters of Marshall and of Hedges and Maxson. Mol. Biol. Evol. 9: 378–379.

Forey, P. L., Humphries, C. J., Kitching, I. J., Scotland, R. W., Siebert, D. J., and Williams D. M. (1992). "Cladistics: A Practical course in Systematics," Clarendon Press, Oxford.

Hedges, S. B., and Bezy, R. L. (1993). Phylogeny of xantusiid lizards: Concern for data and analysis. Mol. Phylogenet. Evol. 2: 76–87.

Hedges, S. B., Bezy, R. L., and Maxson, L. R. (1991). Phylogenetic relationships and biogeography of xantusiid lizards, inferred from mitochondrial DNA sequences. Mol. Biol. Evol. 8: 767–780.

Hedges, S. B., Hass, C. A., and Maxson, L. R. (1993). Relations of fish and tetrapods. Nature 363: 501–502.

Hedges, S. B., and Maxson, L. R. (1992). 18S rRNA sequences and amniote phylogeny: Reply to Marshall. Mol. Biol. Evol. 9: 374–377.

Hillis, D. M., and Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42: 182–192.

Houser, D. L., and Presch, W. F. (1991). The effect of ordered characters on phylogenetic reconstruction. Cladistics 7: 243–265.

Jones, T. R., Kluge, A. G., and Wolf, A. J. (1993). When theories and methodologies clash: A phylogenetic reanalysis of the North American Ambystomatid salamanders (Caudata: Ambystomatidae). Syst. Biol. 42: 92–102.

Jukes, T. H., and Cantor, C. R. (1969). Evolution of protein molecules. In "Mammalian Protein Metabolism" (H. N. Munroe, Ed.), pp. 21–132, Academic Press, New York.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111–120.

Kluge, A. G. (1983). Cladistics and the classification of the great apes. In "New Interpretations of Ape and Human Ancestry" (R. L. Ciochon and R. S. Corruccini, Eds.), pp. 151–177, Plenum, New York.

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae, Serpentes). Syst. Zool. 38: 7–25.

Marshall, C. R. (1992). Substitution bias, weighted parsimony, and amniote phylogeny as inferred from 18S rRNA sequences. Mol. Biol. Evol. 9: 370–373.

Nei, M. (1991). Relative efficiencies of different tree-making methods for molecular data. In "Phylogenetic Analysis of DNA Sequences" (M. M. Miyamoto and J. Cracraft, Eds.), pp. 90–128, Oxford Univ. Press, Oxford.

Rzhetsky, A., and Nei, M. (1992). A simple method for estimating and testing minimum-evolution trees. *Mol. Biol. Evol.* **9:** 945–967.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.

Slowinski, J. B. (1993). "Unordered" versus "ordered" characters. *Syst. Biol.* **42:** 155–165.

Zharkikh, A., and Li, W.-H. (1992). Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock. *Mol. Biol. Evol.* **9:** 1119–1147.

S. BLAIR HEDGES

Department of Biology and Institute of Molecular Evolutionary Genetics Pennsylvania State University University Park, Pennsylvania 16802

ROBERT L. BEZY

Section of Herpetology Natural History Museum of Los Angeles County Los Angeles, California 90007