Syst. Biol. 49(1):143–159, 2000

War of the Iguanas: Conflicting Molecular and Morphological Phylogenies and Long-Branch Attraction in Iguanid Lizards

JOHN J. WIENS[1,4] AND BRADFORD D. HOLLINGSWORTH[2,3]

[1]*Section of Amphibians and Reptiles, Carnegie Museum of Natural History, Pittsburgh, Pennsylvania 15213–4080, USA; E-mail: wiensj@clpgh.org*
[2]*Department of Natural Sciences, Loma Linda University, Loma Linda, California 92350, USA*
[3]*Present address: Department of Herpetology, San Diego Natural History Museum, San Diego, California 92112–1390, USA; E-mail: bhollingsworth@sdnhm.org.*

Abstract.—Recent studies based on different types of data (i.e., morphology, molecules) have found strongly conflicting phylogenies for the genera of iguanid lizards but have been unable to explain the basis for this incongruence. We reanalyze published data from morphology and from the mitochondrial ND4, cytochrome *b*, 12S, and 16S genes to explore the sources of incongruence and resolve these conflicts. Much of the incongruence centers on the genus *Cyclura*, which is the sister taxon of *Iguana*, according to parsimony analyses of the morphology and the ribosomal genes, but is the sister taxon of all other Iguanini, according to the protein-coding genes. Maximum likelihood analyses show that there has been an increase in the rate of nucleotide substitution in *Cyclura* in the two protein-coding genes (ND4 and cytochrome *b*), although this increase is not as clear when parsimony is used to estimate branch lengths. Parametric simulations suggest that *Cyclura* may be misplaced by the protein-coding genes as a result of long-branch attraction; even when *Cyclura* and *Iguana* are sister taxa in a simulated phylogeny, *Cyclura* is still placed as the basal member of the Iguanini by parsimony analysis in 55% of the replicates. A similar long-branch attraction problem may also exist in the morphological data with regard to the placement of *Sauromalus* with the Galápagos iguanas (*Amblyrhynchus* and *Conolophus*). The results have many implications for the analysis of diverse data sets, the impact of long branches on parsimony and likelihood methods, and the use of certain protein-coding genes in phylogeny reconstruction. [Data set incongruence; Iguanidae; likelihood; long-branch attraction, parsimony.]

As the use of molecular data in phylogenetic analyses continues to expand, issues of agreement, conflict, and reconciliation between trees from molecular and morphological data have become increasingly prominent (e.g., Hillis, 1987; Swofford, 1991; Donoghue and Sanderson, 1992; Patterson et al., 1993; Larson, 1994; Hedges and Maxson, 1996; Baker et al., 1998). A major question is whether molecules and morphology yield strongly supported, conflicting hypotheses of relationships, or whether the incongruence is instead due to undersampling of characters in one or both types of data (e.g., Bull et al., 1993; de Queiroz, 1993; Rodrigo et al., 1993). Cases of significant incongruence are important because they suggest the disturbing possibility that phylogenetic analyses based on only one type of data might produce well-supported but incorrect estimates of organismal phylogeny (e.g., Huelsenbeck and Bull, 1996). On the positive side, these well-supported conflicts can provide valuable insights into the causes of failure of certain methods or types of data, and they can allow us to explore ways of detecting and avoiding these errors. Several studies have now reported statistically significant conflict between molecular and morphological data sets (e.g., Miyamoto, 1996; Poe, 1996; Baker et al., 1998), but few studies have been able to identify specific causes for these conflicts (e.g., Normark and Lanteri, 1998).

Recent studies have estimated highly incongruent molecular and morphological phylogenies for iguanid lizards. The family Iguanidae (formerly the iguanines; de Queiroz, 1987; Etheridge and de Queiroz, 1988; Frost and Etheridge, 1989), known informally as iguanas, consists of eight extant genera and ~35 living species (Table 1). Iguanids are found in North, Central, and South America, the West Indies, and three

[4]Corresponding author.

TABLE 1. Genera, species, and distribution of extant iguanid lizards (after de Queiroz, 1987; Hollingsworth, 1998).

Genus	No. species	Distribution
Amblyrhynchus	1	Galápagos Islands
Brachylophus	2	Fiji and Tonga island groups
Conolophus	2	Galápagos Islands
Ctenosaura	13	Mexico and Central America
Cyclura	8	West Indies (Greater Antilles)
Dipsosaurus	2	Southwestern U.S. and Mexico
Iguana	2	Mexico to South America and West Indies (Lesser Antilles)
Sauromalus	5	Southwestern U.S. and Mexico

groups of Pacific ISlands (Figi, Tongo, and Galapagas) and are the only major clade of primarily herbivorous squamates (Etheridge and de Queiroz, 1988). Seemingly because of their relatively large body sizes and high degree of endemism on islands, the family also contains a large number of critically endangered species (Burghardt and Rand, 1982).

Iguanid lizard phylogeny has been the subject of several recent phylogenetic analyses. The first modern analysis of the family was the morphological study by de Queiroz (1987), which was expanded slightly by Norell and de Queiroz (1991). Sites et al. (1996) investigated iguanid phylogeny by using sequences from the mitochondrial ND4 gene (and adjacent transfer RNAs), comparing and combining these data with the morphological data published by de Queiroz (1987) and Norell and de Queiroz (1991). Petren and Case (1997) used mitochondrial cytochrome *b* sequences to examine relationships within the genus *Sauromalus* and its position relative to other iguanid genera. Rassmann (1997) sequenced the 12S and 16S ribosomal mitochondrial genes to test the monophyly of the Galápagos iguanas (*Amblyrhynchus* and *Conolophus*) relative to other iguanid genera. Hollingsworth (1998) examined the species and genus-level phylogeny of the Iguanini (iguanids exclusive of the basal *Brachylophus* and *Dipsosaurus*) using a morphological data set that included the characters used by de Queiroz (1987).

Sites et al. (1996) found that the trees they estimated from ND4 sequences and morphological data were significantly in conflict; however, the morphological trees they generated were very poorly resolved. Furthermore, they were unable to find a definitive explanation for the conflict, although they discussed several possible causes, including long-branch attraction (LBA). Nevertheless, they combined the conflicting data sets and used this combined-data tree, which was largely the same as the molecular tree, as their preferred estimate of phylogeny for the family.

In this study we reanalyze all of the published morphological and molecular data sets for iguanid lizards, to further explore this incongruence and its possible sources. Our results suggest that at least part of the conflict may result from LBA in the protein-coding genes associated with an accelerated rate of change in the genus *Cyclura.* Other conflicting relationships may be associated with long branches in the morphological data, in a different set of genera. The results have many implications for the analysis of diverse data sets, LBA, the use of parsimony versus likelihood methods, rates of molecular and morphological evolution, and the use of certain protein-coding genes in phylogeny reconstruction.

MATERIALS AND METHODS

DNA sequence data were provided by the respective authors or obtained from GenBank. Alignment of the protein-coding sequences (cytochrome *b* and ND4) was relatively straightforward and was performed with Sequencher 3.0 (Gene Codes Corporation). The 12S and 16S ribosomal DNA sequences were realigned using information on secondary structure (following Kjer, 1995; Titus and Frost, 1996): Homologous stem regions were constrained to align to one another while allowing for optimal placement of gaps in other regions. Alignments were performed using CLUSTAL X (a modification of CLUSTAL W; Thompson

et al., 1994), following the procedure outlined in Wiens and Reeder (1997), which was based on that of Titus and Frost (1996). Homologous stems were constrained to align to one another by placing an identical 10-mer sequence ("GATCATCTAG") before and after each of the stems hypothesized by Rassmann (1997) in all species prior to alignment (these were then removed after alignment). Given these structural constraints, different gap costs were explored for both the 12S and 16S genes (by considering gap-opening penalties of 5, 10, and 15). Regions that differed in their alignment under the different gap costs were considered to be "alignment ambiguous" and were removed from the analysis. Other alignment parameters were gap extension penalty = 0.10, delay divergent sequences = 40%, and transitions = transversions. Testing different values for these parameters caused changes only in areas that were considered to be ambiguously aligned under different gap-opening penalties (Wiens, unpubl. data).

All phylogenetic analyses were performed using PAUP* (versions 4.0.0d63 and 4.0b1; Swofford, 1998); the data matrices used are available from the authors on request. The morphological data were those of Hollingsworth (1998), but with two minor corrections to his published data matrix regarding *Iguana delicatissima* (character 20 = state 2, character 21 = state 0). DNA sequence data sets were initially analyzed using unweighted parsimony. The shortest trees were sought by using heuristic searches, with 50 random-addition sequence replicates per search and TBR branch swapping. Support for individual clades was evaluated by nonparametric bootstrapping (Felsenstein, 1985a), using 1,000 pseudoreplicates per analysis with five random-addition sequences per pseudoreplicate. Our cutoff value for "strongly supported" was a bootstrap value of ~70% or higher (based on Hillis and Bull, 1993; but see their caveats). Gaps in DNA sequences were treated as an alternative character state in parsimony analyses rather than as missing data, assuming that insertions and deletions also represent evolutionary changes. Although the nonindependence of nucleotide positions associated with contiguous gaps may be problematic, contiguous gaps in this study were all at phylogenetically uninformative sites.

Maximum likelihood was used to compare the relative likelihoods of trees from the parsimony analyses, to compare the goodness-of-fit of different models of sequence evolution to the observed data, and to search for optimal likelihood trees. Trees from the parsimony analysis were compared using six nested models of increasing complexity (loosely following Huelsenbeck and Crandall, 1997; Sullivan et al., 1997): (1) Jukes–Cantor (JC; Jukes and Cantor, 1969: assuming equal rates of change for transitions and transversions and equal base frequencies), with no invariable sites and no among-site rate variation; (2) Kimura two-parameter (K2P; Kimura, 1980: assuming different rates of change for transitions and transversions and equal base frequencies), with no invariable sites or among-site rate variation; (3) Hasegawa– Kishino–Yano (HKY85; Hasegawa et al., 1985: assuming different rates for transitions and transversions and unequal base frequencies), with no invariable sites or among-site rate variation; (4) HKY85 with some sites assumed to be invariable but equal rates of change assumed at variable sites (HKY85 + I; Hasegawa et al., 1985); (5) HKY85 with some sites assumed to be invariable, and variable sites assumed to follow a gamma distribution (HKY85 + I + Γ; Gu et al., 1995); and (6) general time-reversible (GTR; Yang, 1994: assuming a different rate for all six classes of substitutions), with some sites assumed to be invariable, and variable sites assumed to follow a gamma distribution (GTR + I + Γ). Specific model parameters for likelihood analyses (e.g., base frequencies, transition– transversion ratios, proportion of invariable sites, gamma distribution shape parameter) were estimated from the data using PAUP*. Using maximum likelihood, the goodness-of-fit of different models to the observed data can be evaluated by comparing likelihoods for different models for the same tree. Although the tree with the overall highest likelihood for a given model might be best found by extensively searching the tree space, the comparison of different models on the same tree allows the use of the χ^2 test to compare how well different models fit the data. The statistical

significance of differences in likelihoods of the models was evaluated by using the likelihood-ratio test statistic $-2\log\Lambda$ (the difference between the negative log likelihoods for the two models, multiplied by two), which should approximate a χ^2 distribution with the degrees of freedom being equal to the difference in the number of parameters between the two models (Yang et al., 1995). The best-fitting model was then used in a heuristic search to find the overall best likelihood topology, using TBR branch swapping and 10 random-addition sequence replicates. Support for likelihood trees was evaluated using nonparametric bootstrapping, with 100 pseudoreplicates and one random-addition sequence replicate per bootstrap pseudoreplicate.

Hollingsworth (1998) analyzed his morphological data using both the polymorphic (or baseline) and frequency-bin methods to code within-species variation (Wiens, 1995). Those methods give very similar results for these data (probably because there is relatively little intraspecific variation), although frequency coding generally performs better than baseline/polymorphic coding in simulation and congruence analyses (Wiens and Servedio, 1997, 1998; Wiens, 1998a) and was the method used in this study unless noted otherwise.

Incongruence between the molecular and morphological data sets was examined in three ways. First, data sets were analyzed separately, and the support for conflicting clades was evaluated using nonparametric bootstrapping. Although this is not a statistical test of incongruence in the strict sense, it is a test of whether or not any clades are in strongly supported disagreement between data sets, and unlike many other incongruence tests, can identify which clades are in conflict. Second, we used the incongruence length difference test (ILD) proposed by Farris et al. (1994) to test for overall significant conflict between the data sets. This test involves finding the incongruence index (Mickevich and Farris, 1981) for the data sets when analyzed separately and then randomly repartitioning the combined data to generate a null distribution of this statistic, assuming no significant conflict. All parsimony-uninformative characters were removed from each data set before analysis (not merely the invariant characters, as was recommended by Cunningham, 1997), and 1,000 replicates were analyzed for each pair of data matrices using the "partition homogeneity" test in PAUP* (Swofford, 1998). For each repartitioned data set, heuristic searches with five random-addition sequence replicates each and TBR branch swapping were used. Finally, we used the Wilcoxon signed ranks test (WSR; Templeton, 1983) to test whether or not a given data set significantly rejected the best tree from another data set (as recommended by Larson, 1994), based on the number of changes in each character on each topology. A list of changes was obtained using the "compare two trees" option in MacClade (Maddison and Maddison, 1992). The critical values for a two-tailed test (Rohlf and Sokal, 1981; their Table 30) were used to determine statistical significance (Felsenstein, 1985b; Larson, 1994). Other tests of incongruence have been proposed but are either difficult to apply to morphological data (e.g., likelihood ratio test; Huelsenbeck and Bull, 1996) or are similar to the ones utilized here (e.g., Rodrigo et al., 1993). To test for the presence of substantial conflict in different parts of the phylogeny, the ILD test was rerun, but with the clades previously identified as conflicting constrained to be monophyletic. This tested for the presence of significant conflict apart from these constrained clades (A. de Queiroz, pers. comm.).

The data sets used in this study contained very different numbers of terminal taxa, ranging from 10 (Rassmann, 1997) to 45 (Hollingsworth, 1998). For many of the analyses, we considered it important to have comparable numbers of taxa in each data set, and so some taxa were excluded. For all analyses, we excluded a large number of taxa in the data sets of Petren and Case (1997) and Hollingsworth (1998), because many of these were merely different populations of *Sauromalus.* Although incomplete taxon sampling can have a negative impact on the accuracy of higher-level analyses (e.g., Graybeal, 1998; Hillis, 1998; Wiens, 1998b), our analyses of the complete data sets from these two studies suggest that our exclusion of taxa has little impact on the topology or support for the generic-

level relationships (Petren and Case, 1997; Hollingsworth, 1998; Wiens, unpubl.). The data set of Sites et al. (1996), which has the most generally thorough taxon sampling among the molecular studies, was used as a template to guide the inclusion and exclusion of taxa in other data sets. In a few cases, data from congeners were used interchangeably, to allow generic- level relationships to be compared when a different species was sampled from each genus. The Arizona/Mainland population of *Sauromalus ater* (= *S. obesus*) sampled by Petren and Case (1997) and Hollingsworth (1998) was used to represent *S. ater,* because this was the population sampled in the other molecular analyses. Hollingsworth (1998) synonymized *Sauromalus obesus* with *S. ater,* and we use *S. ater* for the remainder of the paper.

Because all data sets agree that extant iguanid genera exclusive of *Dipsosaurus* and *Brachylophus* form a monophyletic group (the Iguanini of de Queiroz, 1987), these two genera were used to root the tree, rather than referring to some outgroup external to the Iguanidae. This means that our analysis did not address the position of *Brachylophus* and *Dipsosaurus* relative to the iguanid root; but this eliminated potential problems associated with uncertain outgroup relationships (Frost and Etheridge, 1989) and inclusion of long-branch outgroup taxa. Hollingsworth (1998) did not use species of *Brachylophus* and *Dipsosaurus* as units in his analysis, but we used his unpublished data for these species to include them as terminal taxa in our study.

A major question in the analysis of diverse data is what constitutes a "data set" (e.g., Kluge and Wolf, 1993; Chippindale and Wiens, 1994)? In this study, we used the data from each of the major studies of iguanid phylogeny as a data set; for the molecular studies each of these roughly corresponds to a different gene. However, the molecular data of Sites et al. (1996), although consisting largely of ND4 sequences, also contained a small region with the complete sequences of three tRNA genes (217 bp total). The data from the two ribosomal genes also were combined. This pooling avoided the use of any data sets that were extremely small (in terms of numbers of informative characters); moreover, our analyses and those of Sites et al. (1996) suggest that neither the two ribosomal genes nor the ND4 and tRNA sequences give trees that conflict strongly with each other.

The various molecular data sets and the morphological characters were also combined and analyzed together. The combination of strongly conflicting data sets is controversial (see reviews by de Queiroz et al. [1995] and Huelsenbeck et al. [1996a]). We combined these data to examine the effects of the practice in this case, rather than simply use the combined-data tree as the best estimate of iguanid phylogeny. All characters were weighted equally in the combined analyses. Some of the taxa in the combined analyses were missing data from one or more data sets, but resampling and simulation studies suggest that inclusion of these incomplete taxa should not greatly disrupt analyses (Wiens and Reeder, 1995; Wiens, 1998c).

Initial results suggested that LBA (Felsenstein, 1978) might be occurring in the two protein-coding genes (ND4 and cytochrome *b*) and in the morphological data. To test this hypothesis, we used parametric bootstrapping as described by Huelsenbeck et al. (1996b) and Huelsenbeck (1997). For the DNA sequence data, we used the program Siminator (by J. P. Huelsenbeck) to simulate a number of data sets with parameters identical to those estimated for the combined ND4 and cytochrome *b* data sets, but with the hypothesized long branches being separate in the true (model) phylogeny. We then analyzed these simulated data sets to determine how often parsimony and likelihood analyses correctly estimate trees in which these branches are separate versus estimating trees in which the long branches are incorrectly placed together. Parameters for the simulations (number of nucleotides, branch lengths, base frequencies, transition:transversion ratio shape of the Γ distribution of rate variation among sites) were estimated using maximum likelihood (HKY + Γ model; more complex models cannot be simulated with current versions of Siminator). One hundred replicates of each of the simulated trees were analyzed using unweighted par-

simony and maximum likelihood (HKY + Γ; using the simulated base frequencies, transition : transversion ratio, and shape of the Γdistribution of rate variation among sites). As a control, we also simulated a phylogeny in which the long branches were sister taxa, as in the observed trees from parsimony analyses.

The possibility of LBA in the morphological data was also investigated by using parametric bootstrapping. Unfortunately, available methods for estimating branch lengths for trees from discrete morphological data are limited. In this study, branch lengths were estimated based on the proportion of morphological characters changing on a branch over the total number of morphological characters in the analysis; rates of change were based on polymorphic coding (changes would be considerably more difficult to estimate and simulate if treated as frequencies). Trees in which the putative long branch taxa were separated and those where they were sister taxa were both simulated. For each simulation, each data set contained 142 characters (the number of characters in the data set of Hollingsworth, 1998). For the sake of simplicity, all characters were assumed to be binary and fixed (intraspecifically invariant). One hundred replicates were analyzed for each model tree using parsimony. Simulations were performed using a program written in C by Wiens.

RESULTS

For all three molecular data sets, the GTR + I + Γ model (the most parameter-rich model) has the best goodness-of-fit (using trees from the unweighted parsimony analysis); and likelihood ratio tests show that this model substantially improves the likelihood score relative to the next most complex model in all three cases (results not shown). This model was used to estimate phylogenies for all three data sets using maximum likelihood.

Trees from the separate and combined analyses are summarized in Figure 1, and basic statistics for each data set are given in Table 2. Parsimony analysis of the combined ND4 and tRNA sequences (referred to as the ND4 data set hereafter) strongly supports the following intergeneric clades: (1) *Cyclura* as the sister taxon of all other iguanid genera above *Brachylophus* and *Dipsosaurus,* (2) *Ctenosaura* as the sister taxon of the Galápagos iguanas (*Amblyrhynchus* and *Conolophus*), and (3) *Sauromalus* as the sister taxon of *Iguana.* The likelihood tree for these data is nearly identical to the parsimony tree, except that the Galápagos iguanas are paraphyletic in the likelihood tree. However, the three intergeneric clades that are strongly supported by parsimony are only weakly supported by likelihood. The cytochrome *b* data support a parsimony phylogeny that is nearly identical to that based on ND4 (Fig. 1), the only difference being that *Ctenosaura* is paraphyletic with respect to *Conolophus* in one of the two shortest trees. However, this tree is not generally well supported. The likelihood tree for the cytochrome *b* data is the same as one of the parsimony trees except that *Cyclura* is placed with *Iguana* and *Sauromalus* rather than as the sister taxon of all other Iguanini. The parsimony tree based on the ribosomal 12S and 16S genes is weakly supported except for the monophyly of the Iguanini. *Ctenosaura similis* is placed with the Galápagos iguanas and *Ctenosaura quinquecarinata* is placed in a clade with *Sauromalus, Cyclura,* and *Iguana; Cyclura* and *Iguana* are sister taxa. The likelihood tree is similar (and also weakly supported), except that both species of *Ctenosaura* sampled are in a clade with the Galápagos iguanas (which are monophyletic) and *Cyclura* is the sister taxon of *Sauromalus* instead of *Iguana.* The morphological data strongly support the following intergeneric clades: (1) the Galapagos iguanas (*Amblyrhynchus* + *Conolophus*), (2) the Galápagos iguanas plus *Sauromalus,* and (3) *Iguana* and *Cyclura.* Combined analysis of the three molecular data sets and the morphological data supports the generic-level relationships suggested by the parsimony analyses of the ND4 and cytochrome *b* data sets. The latter two data sets contain more than two-thirds of the parsimony-informative characters in the combined data.

The results of the ILD tests (Table 3) suggest substantial conflict between the morphological data set and the ND4 and cytochrome *b* data sets (but not the 12S–16S

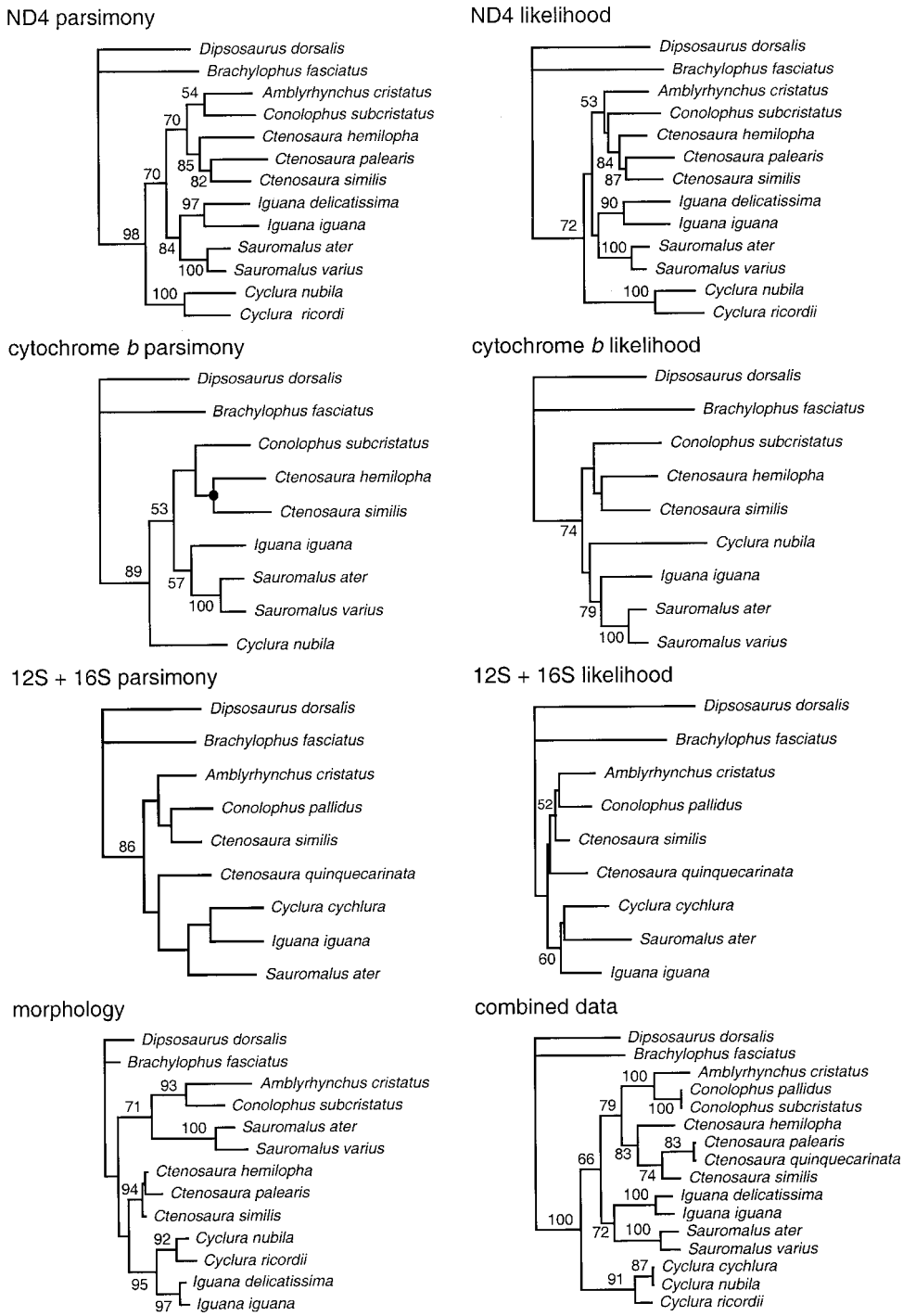

FIGURE 1. Estimated phylogenies of iguanid lizards based on parsimony and likelihood analyses of each of the separate data sets (see Table 2 for description of data sets and trees). Trees are drawn as phylograms, with branch lengths reflecting the estimated amount of evolutionary change. Numbers adjacent to branches are bootstrap values (values <50% not shown). Trees are unrooted, but the root is most likely on the branch leading to *Brachylophus* or *Dipsosaurus*. The cytochrome *b* tree is one of two shortest trees; the black bullet indicates a clade that is collapsed in the strict consensus of the two trees.

TABLE 2. Basic description of the morphological, molecular, and combined data sets for iguanid lizards (trees shown in Fig. 1). Taxa refer only to the number used in these analyses, characters refer only to parsimony-informative characters (total number of base pairs for sequence data are given in parentheses), and trees refers to the number of equally parsimonious topologies from an unweighted analysis.

Data type	Taxa	Characters	Trees	Length	−Ln likelihood
Morphology	13	84	1	195.7	—
ND4 (+ tRNAs)	13	229 (901)	1	865	4772.670
Cytochrome *b*	10	218 (903)	2	735	4196.408
12S + 16S	9	92 (876)	1	354	2672.190
Molecules + morphology	16	618	1	2,178.4	—

data set), and the combined molecular data. There is no significant conflict between any of the three molecular data sets. The results of the WSR tests (Table 4) show that the morphological data strongly reject the shortest trees from all three molecular data sets and the shortest morphological tree is rejected by all three molecular data sets. The ND4 data do not reject the shortest 12S–16S topology but do reject one of the shortest cytochrome *b* trees without rejecting the other. The cytochrome *b* data do not reject the ND4 tree but do reject the 12S–16S trees. The 12S–16S data do not reject either the ND4 or cytochrome *b* trees.

Examination of the separately analyzed data sets suggests two major points of strongly supported incongruence between the morphology and two or more of the molecular data sets: (1) the placement of *Cyclura* as basal within Iguanini by the molecular data (but not by the ribosomal genes or by cytochrome *b* using likelihood) and as the sister taxon of *Iguana* by the morphological data, and (2) the placement of *Sauromalus* in a clade with *Iguana* (along with *Cyclura* in some analyses and data sets) by the molecular data sets and in a clade with the Galápagos iguanas by the morphological data. In fact, these conflicts involve most of the genera within the Iguanini. Constrained ILD tests confirm the presence of multiple points of significant conflict between the molecular and morphological data (Table 5). When the clade *Cyclura* + *Iguana* is constrained to be monophyletic, however, the conflict is no longer significant, suggesting that this clade accounts for much of the incongruence between the molecular and morphological data for iguanid lizards. Nevertheless, when the ILD test is run with *Cyclura* and *Iguana* deleted, highly significant conflict ($P = 0.006$) remains.

TABLE 3. Results of ILD tests between data sets for iguanid lizards.

Data sets compared	*P*-value
Morphology vs. ND4	0.002
Morphology vs. cytochrome *b*	0.006
Morphology vs. 12S–16S	0.257
Morphology vs. combined molecular data	0.002
ND4 vs. cytochrome *b*	1.000
ND4 vs. 12S–16S	0.388
Cytochrome *b* vs. 12S–16S	0.364

The statistically significant, strongly supported conflict between the molecular data (especially the protein-coding genes) and the morphological data suggests the presence of some systematic error in one or both types of data. Examination of the branch lengths estimated by maximum likelihood for the ND4 and cytochrome *b* trees suggests that the branches associated with the genus *Cyclura* are relatively long, which raises the possibility that LBA is the source of error (Fig. 1). A parametric bootstrapping simulation of the combined ND4 and cytochrome *b* data (Fig. 2) shows that when *Cyclura* and *Iguana* are sister taxa in the simulated model tree, parsimony recovers this clade correctly in only 10% of the replicates (Fig. 3). In the majority of replicates (55%), *Cyclura* is placed as the basal lineage of the Iguanini, just as it is in the parsimony estimates from the empirical data (Fig. 2); the ancestor of the Iguanini also represents a relatively long branch. This result strongly suggests that the *Cyclura* branch is long enough to mislead parsimony analysis. Maximum likelihood analysis of the simu-

TABLE 4. Results of WSR tests between trees from different data sets for iguanid lizards. Intraspecific variation in morphology was coded by using the polymorphic method (to avoid the complex weighting of the frequency method), but the single shortest tree from the frequency method was used as the morphology tree (this is one of the shortest trees obtained using polymorphic coding).

Data set	Tree 1	Tree 2	*P*-value
Morphology	Morphology	ND4	<0.01
	Morphology	Cytochrome *b*–tree 1	<0.01
	Morphology	Cytochrome *b*–tree 2	<0.01
	Morphology	12S–16S	<0.01
ND4	ND4	Morphology	<0.01
	ND4	Cytochrome *b*–tree 1	0.02
	ND4	Cytochrome *b*–tree 2	>0.05
	ND4	12S–16S	>0.05
Cytochrome *b*	Cytochrome *b*–tree 1	Morphology	0.02
	Cytochrome *b*–tree 2	Morphology	<0.01
	Cytochrome *b*–tree 1	ND4	>0.05
	Cytochrome *b*–tree 2	ND4	>0.05
	Cytochrome *b*–tree 1	12S–16S	0.02
	Cytochrome *b*–tree 2	12S–16S	0.02
12S–16S	12S–16S	Morphology	0.01
	12S–16S	ND4	>0.05
	12S–16S	Cytochrome *b*–tree 1	>0.05
	12S–16S	Cytochrome *b*–tree 2	>0.05

lated data sets using the HKY + Γ model recovers the correct *Cyclura* + *Iguana* clade in 25% of the replicates (Fig. 3). In 31% of the replicates, *Cyclura* is placed as the sister taxon of *Iguana* + *Sauromalus* (as it is in the likelihood trees in the real data; Fig. 2), whereas in 25% of the replicates *Cyclura* is the sister taxon of *Sauromalus.* In contrast to parsimony, maximum likelihood never places *Cyclura* as the basal member of the Iguanini. When *Cyclura* is the basal member of the Iguanini in the model trees, parsimony always recovers this aspect of the tree correctly, whereas maximum likelihood is successful in 88% of the replicates but places *Cyclura* as the sister taxon of *Iguana* + *Sauromalus* in 7%.

TABLE 5. Results of ILD tests (morphology vs. combined molecular data) with certain conflicting clades identified in the separate analyses constrained to be monophyletic.

Constrained clade	*P*-value
No constraints	0.002
(*Sauromalus, Iguana, Ctenosaura, Amblyrhynchus, Conolophus*)	0.008
(*Ctenosaura, Amblyrhynchus, Conolophus*)	0.001
(*Iguana, Sauromalus*)	0.019
(*Cyclura, Iguana*)	0.126
(*Sauromalus, Amblyrhynchus, Conolophus*)	0.004

There is also a possibility of LBA in the morphological data set. In the trees based on the morphological data, the branches associated with *Sauromalus* and the Galápagos iguanas are relatively long, and the grouping of these three genera is strongly contradicted by all three molecular data sets. Parametric bootstrapping of the morphological data set, assuming a tree in which *Sauromalus* is related to *Cyclura* and *Iguana* (as suggested by the molecular data sets), provides some evidence, albeit weak, for LBA (Fig. 4). The correct (simulated) placement of *Sauromalus* is recovered in less than half of the replicates (43%), but the expected result (given LBA)—*Sauromalus* grouping with the Galápagos iguanas regardless of its true relationships—appears in only 24% of the replicates. However, it is possible that the lengths of these branches used in the simulations were underestimated, and that longer branches would provide stronger ev-

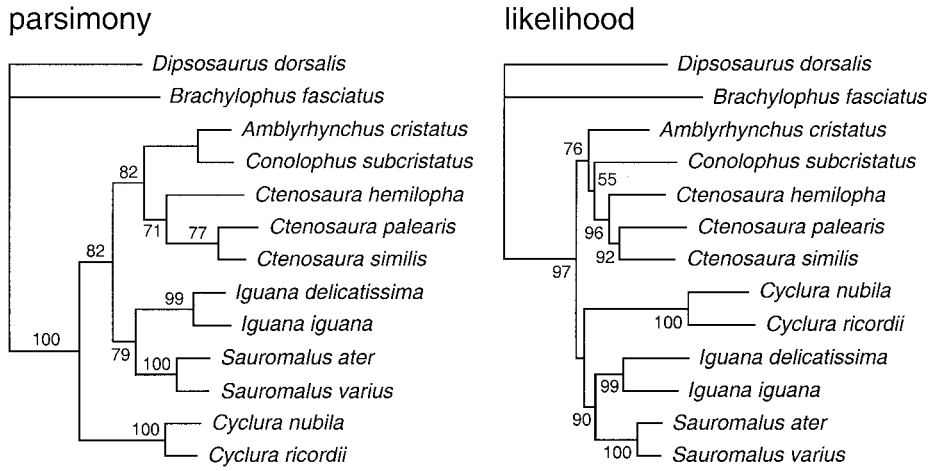

FIGURE 2. Parsimony and likelihood trees for the combined ND4 and cytochrome *b* data sets. Numbers adjacent to branches are bootstrap values (values <50% not shown). Trees are drawn as phylograms, with branch lengths reflecting the estimated amount of evolutionary change. Parsimony length = 1598. −Ln likelihood = 8992.595.

idence for LBA. When the estimated morphological phylogeny is assumed to be true (Fig. 4), the correct placement of *Sauromalus* is recovered in 52% of the replicates, the next most common result (31%) being an unresolved placement.

DISCUSSION

Long-Branch Attraction

The results of our analyses suggest that LBA, the tendency for taxa with long branches to be placed together in an estimated phylogeny regardless of their actual relationships, may explain much of the data set incongruence in iguanid lizards. LBA is an important issue in phylogenetics because it can cause methods to estimate phylogenies that are both incorrect and statistically well-supported. Yet, despite its potential importance, few well-documented empirical examples of this phenomenon have been presented (amniotes [Huelsenbeck et al., 1996b]; insects [Huelsenbeck, 1997]; rodents [Sullivan and Swofford, 1997]).

Huelsenbeck (1997) proposed two criteria for deciding whether or not there is sufficient evidence in a given empirical study to invoke LBA: (1) the long branches must be shown to be long enough to attract each other (i.e., in simulations in which the simulated branch lengths are equivalent to the observed empirical branch lengths, the long branch taxa will be placed together even when they are not sister taxa); and (2) a method that is relatively insensitive to LBA (i.e., maximum likelihood) must be shown to separate the long branches in the estimated trees. Both of these criteria are met in the iguanid example, at least for the ND4 + cytochrome *b* data.

A third criterion that might be added to this list is to provide evidence (i.e., from other data sets) that the long branches are not actually sister taxa. In the case of iguanid lizards, the morphological data and ribosomal DNA sequences suggest that *Iguana* and *Cyclura* are sister taxa, and that *Cyclura* is not at the base of the Iguanini. Other studies also have discussed evidence from independent data sets (i.e., morphology) that the putative long-branch taxa are not closest relatives (e.g., Sullivan and Swofford, 1997). However, in the case of the strepsipteran insects, evidence from morphological data suggests that the long branches may be sister taxa (Whiting et al., 1997), and that this example may actually represent long-branch repulsion—the tendency of maximum likelihood and similar methods to separate long branches that are actually closest relatives (Siddall, 1998).

Our results from iguanid lizards have interesting implications for the ongoing debate over the relative superiority of parsi-

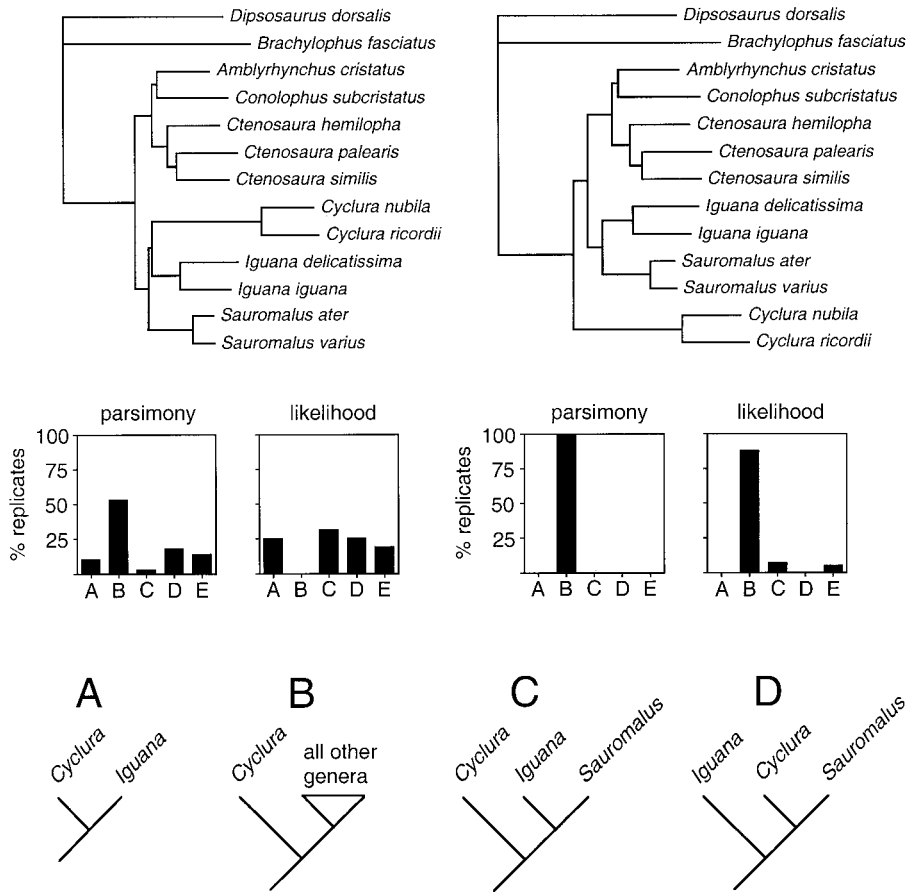

FIGURE 3. Results of a parametric simulation of a potential long-branch problem in the combined protein-coding genes (ND4 + cytochrome *b*) for iguanid lizards. Two trees were simulated, one in which *Cyclura* and *Iguana* are sister taxa, the other in which *Cyclura* is the basal member of the Iguanini. The trees are the true phylogenies in each set of simulations, and the branch lengths were estimated using likelihood (the length of the *Iguana* + *Cyclura* clade has been lengthened slightly for illustrative purposes). The graphs show how often each placement of *Cyclura* (A–D) is supported by a given method in the 100 replicated data sets. E indicates alternative placements and unresolved relationships.

mony and likelihood methods and the effects of long branches on these methods (Huelsenbeck, 1995, 1997, 1998; Yang, 1996; Siddall, 1998). Recent simulation studies have shown that (1) likelihood should estimate the correct tree under conditions where parsimony is subject to LBA, as long as there are sufficient characters and an adequate fit between the observed data and the model of evolution assumed by the likelihood method (e.g., Huelsenbeck, 1995; Yang, 1996), (2) maximum likelihood may estimate the incorrect tree if the taxa with long branches are actually each others' closest relatives (Yang, 1996; Huelsenbeck, 1998; Siddall, 1998). In our study, maximum likelihood analyses do not recover the presumably correct *Cyclura* + *Iguana* clade in separate or combined analyses of the protein-coding genes, suggesting that too few characters have been sampled or that the fit is inadequate between the evolutionary processes assumed by the likelihood methods and the actual processes of character evolution in these data. In the parametric simulations of the combined ND4 and cytochrome *b* data, likelihood recovers the correct (simulated) *Iguana* + *Cyclura* clade in 25% of the replicates, about as often as the two other frequently supported resolutions. Because there is a perfect fit in these analyses between the simulated model of

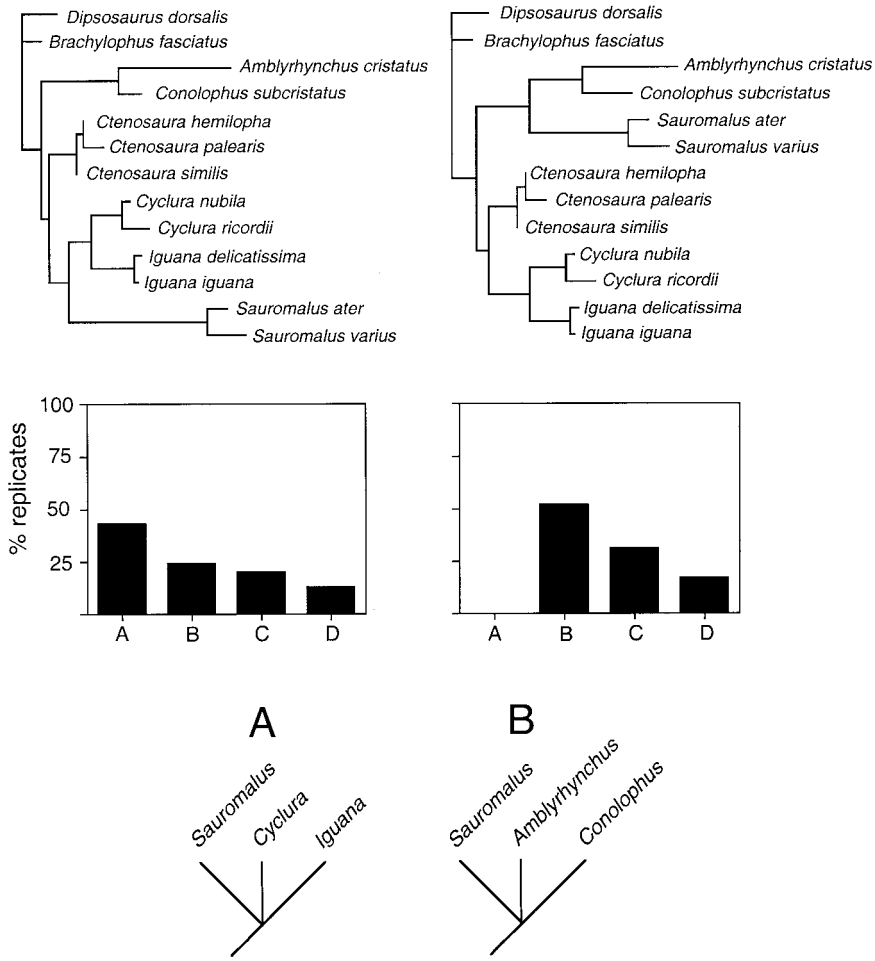

FIGURE 4. Results of a parametric simulation of a potential long-branch problem in the morphological data for iguanid lizards. Two trees were simulated, one in which *Sauromalus* is in a clade with *Cyclura* and *Iguana,* the other in which *Sauromalus* is in a clade with the Galápagos iguanas. The trees shown are the model (true) phylogenies in each set of simulations, with branch lengths estimated by using parsimony. The graphs show how often each placement of *Sauromalus* (A–B) is supported by parsimony in the 100 replicated data sets. C indicates unresolved relationships, and D represents assorted alternative resolutions.

evolution and the model assumed by maximum likelihood, inadequate sampling of characters may be the problem. When the simulated phylogeny has *Cyclura* as basal (and there is potential for long-branch repulsion), maximum likelihood recovers the correct phylogeny in 88% of the replicates. Thus, the simulated phylogenies show that for this data set and branch lengths, there is relatively strong evidence for LBA and weak evidence for long-branch repulsion.

Our study also suggests that long branch problems may be missed entirely if only parsimony is used to infer branch lengths (e.g., Huelsenbeck, 1998). Although the *Cyclura* ancestor is the longest internal branch within the Iguanini, as determined from both parsimony and likelihood analyses for the combined ND4-cytochrome *b* data sets (Fig. 2), there is a clear disparity in the relative length of this branch estimated by the two methods. When branch lengths are estimated by parsimony, the *Cyclura* branch is 1.6 times longer than the next longest internal branch within the Iguanini, and 2.2 times longer than the average length of the other internal branches. For likelihood, the *Cyclura* branch is 3.7 times longer than

the next longest internal branch, and 6.2 times longer than the average length of the other internal branches. Although we do not know what the "correct" lengths are, the tendency for parsimony to underestimate branch lengths is well known (Swofford et al., 1996). Simulation results have now shown that long branches can be problematic for both parsimony and likelihood and will cause both methods to fail under some conditions. Consequently, it is crucial to at least be able to detect long branches and the use of parsimony alone increases the risk of overlooking long branches and the possible failure of either method.

Finally, our study suggests the possibility that LBA may occur in morphological data sets as well; previous examples have involved only molecular data. Unfortunately, the analysis of this phenomenon in morphological data is hampered by uncertainty in estimating branch lengths; the use of parsimony may underestimate branch lengths (i.e., because multiple hits are not accounted for), and the frequent exclusion of invariant and autapomorphic characters by morphologists may also cause branch lengths to be over- or underestimated. Likelihood methods for morphological data are limited. Felsenstein's (1981) continuous maximum likelihood method can be effective on morphological data sets consisting of qualitative characters treated as frequencies (Wiens, 1998a), but current versions of the program CONTML do not allow characters with missing data to be used (a serious problem in many data sets, including that of Hollingsworth, 1998). Finally, use of purely stochastic models for morphology may be inappropriate. For example, many of the unique characters associated with *Sauromalus* and *Amblyrhynchus* may be adaptations to their unusual habitats (rock crevices and rocky seashores), which would indicate that selection is at least partly responsible for the accelerated change in these lineages (de Queiroz, 1987).

Other Potential Sources of Conflict

We acknowledge that LBA is only one of many possible explanations for the incongruence between the molecular and morphological phylogenies of iguanid lizards. Sites et al. (1996) considered three other explanations: (1) mismatch between the gene phylogeny and species phylogeny, (2) saturation in the DNA sequence data, and (3) misdiagnosis of morphological characters. The presence of the *Iguana–Cyclura* clade (albeit weakly supported) in the parsimony trees from the ribosomal genes provides at least some evidence against a mismatch between the gene and species phylogeny involving all three linked mitochondrial genes. Sites et al. (1996) ruled out character saturation as a potential source of error, and character misdiagnosis should be a source of random error rather than systematic error (i.e., the conflicting clades in the morphological tree should therefore not have the high bootstrap values that they do). Another possible source of systematic error would be nonindependence of some of the morphological characters (e.g., Emerson and Hastings, 1998); yet, both the *Iguana–Cyclura* clade and the *Sauromalus*–Galápagos iguana clade are supported by diverse skeletal and external characters (Hollingsworth, 1998), none of which seem obviously nonindependent. Given these arguments, we favor LBA as an explanation for the incongruence in iguanid lizards, at least in the protein-coding genes.

Implications for Conflicting Data Sets

Our results from iguanid lizards have implications for another major debate in phylogenetics: combined versus separate analysis of diverse data sets (see reviews by de Queiroz et al., 1995; Huelsenbeck et al., 1996a). If one accepts our hypothesis that the protein-coding genes have been misled in their placement of the genus *Cyclura* by LBA, then clearly the combined analysis of all data sets (molecular and morphological) has also been misled. This is a disturbing result, in that the combined analysis includes four seemingly independent data sets (although three of these are admittedly linked) and the combined data tree appears to be well-supported at the generic level (i.e., bootstrap values close to or above 70%). The basis for this problematic result appears to be that the source of error (LBA) affects two of the data sets in a similar way, and these two data sets contain more than

two-thirds of the parsimony-informative characters in the combined data. In summary, these results strongly argue against assuming that the combined data tree is always the best estimate of phylogeny and argue in favor of examining trees from the separately analyzed data sets for potential sources of error. However, the combined-data tree for iguanids may still be more accurate overall than any of the trees from the separate analyses, especially given the possibility of systematic error in both molecules and morphology (see Wiens, 1998d).

Our study also shows that real conflict between data sets may be hidden by insufficient character sampling. In the analyses of Sites et al. (1996) the tree based on the reanalysis of the morphological data of de Queiroz (1987) was poorly resolved and poorly supported (e.g., the *Cyclura–Iguana* clade was not even resolved), and the morphological data could not reject the molecular (ND4) tree by the WSR test (although the ND4 data rejected the morphology tree). In the present study, the morphological tree, which is based on the same characters as de Queiroz (1987) plus additional characters, is relatively well resolved, well supported, and significantly in conflict with the trees from molecular data. This example is a useful reminder that differences between trees from different data sets involving weakly supported clades merely indicate the failure to find significant conflict (a negative result). This is an important idea because many methodologies for dealing with diverse data sets implicitly assume that the absence of statistically significant conflict indicates congruence (e.g., Bull et al., 1993; de Queiroz, 1993; Wiens, 1998d). One way to deal with this issue might be to use parametric simulations to test the power of the incongruence tests for a given case study in which data sets produce different topologies but are not significantly in conflict according to a statistical test. By simulating data sets with the same number of characters and same model of evolution as the observed data sets, but with different underlying topologies, one could evaluate whether or not the incongruence test would be able to detect actual differences if they existed.

Rates of Molecular and Morphological Evolution

It is clear that rates of molecular and morphological evolution are decoupled in iguanid lizards, at least in some lineages. There seems to have been accelerated molecular evolution in the protein-coding genes ND4 and cytochrome *b* associated with the genus *Cyclura,* whereas there has been accelerated morphological evolution on the branches associated with *Sauromalus* and the Galápagos iguanas. The decoupling of molecular and morphological evolution in these lineages is interesting in light of a recent review (Omland, 1997), which found rates of change to be generally correlated between these types of data.

The cause of the increased rate of change in the protein-coding genes is unclear. Because cytochrome *b* and ND4 are widely used to infer phylogenies at various hierarchical levels, the possibility that they may give positively misleading results for unknown reasons is of particular concern and should be investigated further. Given this, we urge caution in estimating phylogenies based only on these genes, especially when they conflict with other lines of evidence. Meyer (1994) noted that cytochrome *b* may evolve at unequal rates among distantly related lineages (although he too lacked an explanation as to why this was the case) and that this tendency made the usefulness of the gene problematic at higher taxonomic levels. Our results suggest that this problem may apply to analyses of genera within a single family.

The accelerated rate of change in the morphological characters in *Sauromalus* and the Galápagos iguanas may be associated with adaptations to unusual ways of life. *Sauromalus* are highly specialized rock-crevice dwellers, and *Amblyrhynchus* is the only marine lizard (but is also a rock dweller). de Queiroz (1987) suggested that adaptation was involved in many of the unique morphological features in these two genera.

What is the Best Estimate of Iguanid Lizard Phylogeny?

Because of extensive conflicts between the data sets, we argue that the best esti-

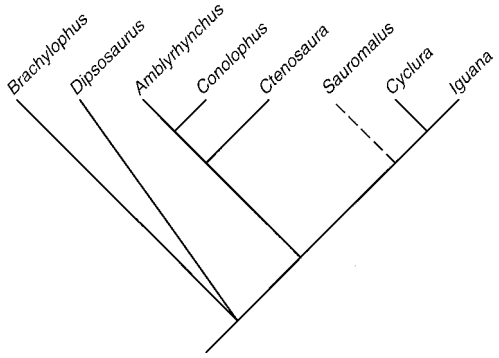

FIGURE 5. Our preferred hypothesis of intergeneric relationships of iguanid lizards. The dashed branch indicates our uncertainty about the phylogenetic placement of *Sauromalus.*

mate of iguanid phylogeny is not represented among any of the trees from the separate or combined analyses (Fig. 5). We hypothesize that the placement of *Cyclura* as basal by the ND4 and cytochrome *b* data sets is erroneous, and that *Cyclura* is instead the sister taxon of *Iguana.* The placement of *Ctenosaura* with the Galápagos iguanas is supported by separate analyses of all three molecular data sets and by the combined molecular and morphological data. The morphological data in this analysis place *Ctenosaura* as the sister taxon of the *Iguana* + *Cyclura* clade, but this is only weakly supported (bootstrap <50%). We tentatively consider the conflict between the molecular and morphological positions of *Ctenosaura* to be the result of stochastic error in the morphological data (e.g., insufficient sampling of characters), and favor placement of *Ctenosaura* with the Galápagos iguanas. The position of *Sauromalus* is less clear. All three molecular data sets agree that *Sauromalus* is associated with *Iguana* or *Cyclura,* or both, and this placement is also supported by the combined analysis. The morphological data strongly support a clade containing *Sauromalus* and the Galápagos iguanas. Although the latter association could be due to convergence or LBA, it is also possible that some unknown source of error affects the linked mitochondrial data sets. We believe that the current evidence is insufficient to rule out either the molecular or morphological placement of *Sauromalus* but think that its position with *Iguana* and *Cyclura* is most likely correct. Additional, unlinked data sets (such as nuclear gene sequences) should provide further insight into all of these problems. Furthermore, increased sampling of species within *Cyclura* for the ND4 and cytochrome *b* genes might break up the long branch associated with this genus and improve the estimate of iguanid phylogeny provided by these data sets.

ACKNOWLEDGMENTS

We are grateful to Jack Sites (and his collaborators), Ken Petren, Ted Case, and Kornelia Rassmann for collecting the molecular data that made our study possible and for providing us with data files and other relevant information. We are especially grateful to John Huelsenbeck for running his Siminator program for our study. We thank Tod Reeder, Cliff Cunningham, and Phil Chu for valuable advice on DNA sequence alignment, the ILD test, and the WSR test, respectively. We thank David Swofford for permission to analyze our data and publish our results by using a test version of his PAUP* program. Larry Buckley, Alan de Queiroz, Richard Etheridge, Brad Livezey, Richard Olmstead, Maria Servedio, Jack Sites, and an anonymous reviewer gave useful comments on various drafts of the manuscript.

REFERENCES

BAKER, R. H., X. YU, AND R. DESALLE. 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. Mol. Phylogenet. Evol. 9:427–436.

BULL, J. J., J. P. HUELSENBECK, C. W. CUNNINGHAM, D. L. SWOFFORD, AND P. J. WADDELL. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42:384–497.

BURGHARDT, G. M., AND A. S. RAND (eds.). 1982. Iguanas of the world: Their behavior, ecology, and conservation. Noyes Publications, Park Ridge, New Jersey.

CHIPPINDALE, P. T., AND J. J. WIENS. 1994. Weighting, partitioning, and combining characters in phylogenetic analysis. Syst. Biol. 43:278–287.

CUNNINGHAM, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14:733–740.

DE QUEIROZ, A. 1993. For consensus (sometimes). Syst. Biol. 42:368–372.

DE QUEIROZ, A., M. J. DONOGHUE, AND J. KIM. 1995. Separate versus combined analysis of phylogenetic evidence. Annu. Rev. Ecol. Syst. 26:657–681.

DE QUEIROZ, K. 1987. Phylogenetic systematics of iguanine lizards. A comparative osteological study. Univ. Calif. Publ. Zool. 118:1–203.

DONOGHUE, M. J., AND M. J. SANDERSON. 1992. The suitability of molecular and morphological evidence in reconstructing plant phylogeny. Pages 340–368 *in* Molecular systematics in plants (P. S. Soltis, D. E. Soltis, and J. J. Doyle, eds.). Chapman and Hall, New York.

EMERSON, S. B., AND P. A. HASTINGS. 1998. Morphological correlations in evolution: Consequences for phylogenetic analysis. Q. Rev. Biol. 73:141–162.

ETHERIDGE, R., AND K. DE QUEIROZ. 1988. A phylogeny of Iguanidae. Pages 283–368 *in* Phylogenetic relationships of the lizard families: Essays commemorating Charles L. Camp (R. Estes and G. Pregill, eds.). Stanford University Press, Stanford, California.

FARRIS, J. S., M. KALLERSJO, A. G. KLUGE, AND C. BULT. 1994. Testing significance of incongruence. Cladistics 10:315–319.

FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

FELSENSTEIN, J. 1981. Evolutionary trees from gene frequencies and quantitative characters: Finding maximum likelihood estimates. Evolution 35:1229–1242.

FELSENSTEIN, J. 1985a. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

FELSENSTEIN, J. 1985b. Confidence limits on phylogenies with a molecular clock. Syst. Zool. 34:152–161.

FROST, D. R., AND R. ETHERIDGE. 1989. A phylogenetic analysis and taxonomy of iguanian lizards (Reptilia: Squamata). Misc. Publ. Mus. Nat. Hist. Univ. Kansas 81:1–65.

GRAYBEAL, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47:9–17.

GU, X., Y.-X. FU, AND W.-H. LI. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol. Biol. Evol. 12:546–557.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating the human–ape split by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.

HEDGES, S. B., AND MAXSON, L. R. 1996. Re: Molecules and morphology in amniote phylogeny. Mol. Phylogenet. Evol. 6:312–314.

HILLIS, D. M. 1987. Molecular versus morphological approaches to systematics. Annu. Rev. Ecol. Syst. 18:23–42.

HILLIS, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47:3–8.

HILLIS, D. M., AND J. J. BULL. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42:182–192.

HOLLINGSWORTH, B. D. 1998. The systematics of chuckwallas (*Sauromalus*) with a phylogenetic analysis of other iguanid lizards. Herpetol. Monogr. 12:38–191.

HUELSENBECK, J. P. 1995. Performance of phylogenetic methods in simulation. Syst. Biol. 44:17–48.

HUELSENBECK, J. P. 1997. Is the Felsenstein Zone a fly trap? Syst. Biol. 46:69–74.

HUELSENBECK, J. P. 1998. Systematic bias in phylogenetic analysis: Is the Strepsiptera problem solved? Syst. Biol. 47:519–537.

HUELSENBECK, J. P., AND J. J. BULL. 1996. A likelihood ratio test to detect conflicting phylogenetic signal. Syst. Biol. 45:92–98.

HUELSENBECK, J. P., J. J. BULL, AND C. W. CUNNINGHAM. 1996a. Combining data in phylogenetic analysis. Trends Ecol. Evol. 11:152–158.

HUELSENBECK, J. P., AND K. A. CRANDALL. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu. Rev. Ecol. Syst. 28:437–466.

HUELSENBECK, J. P., D. M. HILLIS, AND R. JONES. 1996b. Parametric bootstrapping in molecular phylogenetics: Application and performance. Pages 19–45 *in* Molecular zoology: Advances, strategies, and protocols (J. D. Ferraris and S. R. Palumbi, eds.). Wiley-Liss, New York.

JUKES, T. H., AND C. R. CANTOR. 1969. Evolution of protein molecules. Pages 21–132 *in* Mammalian protein metabolism (H. N. Munro, ed.). 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.

KJER, K. M. 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: An example of alignment and data presentation from frogs. Mol. Phylogenet. Evol. 4:314–330.

KLUGE, A. G., AND A. J. WOLF. 1993. Cladistics: What's in a word? Cladistics 9:183–199.

LARSON, A. 1994. The comparison of molecular and morphological data in phylogenetic systematics. Pages 371–390 *in* Molecular ecology and evolution: Approaches and applications (B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds.). Birkauser Verlag, Basel, Switzerland.

MADDISON, W. P., AND D. R. MADDISON. 1992. MacClade Ver. 3.0. Analysis of phylogeny and character evolution. Sinauer, Sunderland.

MEYER, A. 1994. Shortcomings of the cytochrome *b* gene as a molecular marker. Trends Ecol. Evol. 9:278–280.

MICKEVICH, M. F., AND J. S. FARRIS. 1981. The implications of congruence in *Menidia.* Syst. Zool. 30:351–370.

MIYAMOTO, M. 1996. A congruence study of the molecular and morphological data for eutherian mammals. Mol. Phylogenet. Evol. 6:373–390.

NORELL, M. A., AND K. DE QUEIROZ. 1991. The earliest iguanine lizard (Reptilia; Squamata) and its bearing on iguanine phylogeny. Am. Mus. Novit. 2997:1–16.

NORMARK, B. B., AND A. A. LANTERI. 1998. Incongruence between morphology and mitochondrial-DNA characters suggests hybrid origins of parthenogenetic weevil lineages (genus *Aramigus*). Syst. Biol. 47:475–494.

OMLAND, K. E. 1997. Correlated rates of molecular and morphological evolution. Evolution 51:1381–1393.

PATTERSON, C., D. M. WILLIAMS, AND C. J. HUMPHRIES. 1993. Congruence between molecular and morphological phylogenies. Annu. Rev. Ecol. Syst. 24:153–188.

PETREN, K., AND T. J. CASE. 1997. A phylogenetic analysis of body size evolution and biogeography in chuckwallas (*Sauromalus*) and other iguanines. Evolution 51:206–219.

POE, S. 1996. Data set incongruence and the phylogeny of crocodilians. Syst. Biol. 45:393–414.

RASSMANN, K. 1997. Evolutionary age of the Galápagos iguanas predates the age of the present Galápagos islands. Mol. Phylogenet. Evol. 7:158–172.

RODRIGO, A. G., M. KELLY-BORGES, P. R. BERGQUIST, AND P. L. BERGQUIST. 1993. A randomization test of the null hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. N.Z. J. Bot. 31:257–268.

ROHLF, F. J., AND R. R. SOKAL. 1981. Statistical tables, 2nd edition. W. H. Freeman, San Francisco, California.

SIDDALL, M. E. 1998. Success of parsimony in the four-taxon case: Long branch repulsion by likelihood in the Farris Zone. Cladistics 14:209–220.

SITES, J. W., JR., S. K. DAVIS, T. GUERRA, J. B. IVERSON, AND H. L. SNELL. 1996. Character congruence and phylogenetic signal in molecular and morphological data sets: A case study in the living iguanas (Squamata: Iguanidae). Mol. Biol. Evol. 13:1087–1105.

SULLIVAN, J., J. A. MARKERT, AND C. W. KILPATRICK. 1997. Phylogeography and molecular systematics of the *Peromyscus aztecus* species group (Rodentia: Muridae) inferred using parsimony and likelihood. Syst. Biol. 46:426–440.

SULLIVAN, J., AND D. L. SWOFFORD. 1997. Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. J. Mammal. Evol. 4:77–86.

SWOFFORD, D. L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? Pages 295–333 *in* Phylogenetic analysis of DNA sequences (M. M. Miyamoto and J. Cracraft, eds.). Oxford Univ. Press, New York.

SWOFFORD, D. L. 1998. PAUP*: Phylogenetic analysis using parsimony*. Ver. 4.0.0d63. Sinauer, Sunderland, Massachusetts.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogeny reconstruction. Pages 407–514 *in* Molecular systematics, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. Evolution 37:221–244.

THOMPSON, J. D., D. G. HIGGINS, AND T. J. GIBSON. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

TITUS, T. A., AND D. R. FROST. 1996. Molecular homology assessment and phylogeny in the lizard family Opluridae (Squamata: Iguania). Mol. Phylogenet. Evol. 5:49–62.

WHITING, M. F., J. C. CARPENTER, Q. D. WHEELER, AND W. C. WHEELER. 1997. The Strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Syst. Biol. 46:1–68.

WIENS, J. J. 1995. Polymorphic characters in phylogenetic systematics. Syst. Biol. 44:482–500.

WIENS, J. J. 1998a. Testing phylogenetic methods with tree-congruence: Phylogenetic analysis of polymorphic morphological characters in phrynosomatid lizards. Syst. Biol. 47:427–444.

WIENS, J. J. 1998b. The accuracy of methods for coding and sampling higher-level taxa for phylogenetic analysis: A simulation study. Syst. Biol. 47:397–413.

WIENS, J. J. 1998c. Does adding characters with missing data increase or decrease phylogenetic accuracy? Syst. Biol. 47:625–640.

WIENS, J. J. 1998d. Combining data sets with different phylogenetic histories. Syst. Biol. 47:568–581.

WIENS, J. J., AND T. W. REEDER. 1995. Combining data sets with different numbers of taxa for phylogenetic analysis. Syst. Biol. 44:548–558.

WIENS, J. J., AND T. W. REEDER. 1997. Phylogeny of the spiny lizards (*Sceloporus*) based on molecular and morphological evidence. Herpetol. Monogr. 11:1–101.

WIENS, J. J., AND M. R. SERVEDIO. 1997. Accuracy of phylogenetic analysis including and excluding polymorphic characters. Syst. Biol. 46:332–345.

WIENS, J. J., AND M. R. SERVEDIO. 1998. Phylogenetic analysis and intraspecific variation: Performance of parsimony, likelihood, and distance methods. Syst. Biol. 47:228–253.

YANG, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39:306–314.

YANG, Z. 1996. Phylogenetic analysis using parsimony and likelihood methods. J. Mol. Evol. 42:294–307.

YANG, Z., N. GOLDMAN, AND A. FRIDAY. 1995. Maximum likelihood trees from DNA sequences: A peculiar statistical estimation problem. Syst. Biol. 44: 384–399.

Received 1 March 1999; accepted 5 May 1999
Associate Editor: A. de Queiroz