Syst. Biol. 49(4):713–739, 2000

Multiple Data Sets, Congruence, and Hypothesis Testing for the Phylogeny of Basal Groups of the Lizard Genus *Sceloporus* (Squamata, Phrynosomatidae)

OSCAR FLORES-VILLELA,[1,2] KARL M. KJER,[1,3] MIRIAM BENABIB,[1,4] AND JACK W. SITES, JR.[1,5]

[1]*Department of Zoology and M. L. Bean Museum, Brigham Young University, Provo, Utah 84602, USA; E-mail: ofv@hp.fciencias.unam.mx (O.F.V.), kjer@aesop.rutgers.edu (K.M.K.), benabib@servidor.unam.mx (M.B.), Jack_Sites@byu.edu (J.W.S.)*

[2]*Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México, Apdo. Postal 70-399, México D.F. 04510, México*

[3]*Department of Entomology, Cook College, Rutgers University, New Brunswick, New Jersey 08903, USA*

[4]*Instituto de Ecolog´a, Universidad Nacional Autónoma de México, Apdo. Postal 70-275, México D.F. 04510, México*

Abstract.—Several data partitions, including nuclear and mitochondrial gene sequences, chromosomes, isoenzymes, and morphological characters, were used to propose a new phylogeny and to test previously published hypotheses about the phylogenetic positions of basal clades of the lizard genus *Sceloporus* and the relationship of *Sceloporus* to the former genus "*Sator*". In accord with earlier studies, our results grouped "*Sator*" as internal to *Sceloporus*, and both support a hypothesis of transgulfian vicariance for the origin of the former genus "*Sator*" on islands in the Sea of Cortez. Robustness of support for internal nodes in our best tree was established though widely used indices (bootstrap proportions, decay values) but also through congruence among independent data partitions. Several deep nodes in the tree recovered by several methods, including equally weighted and differentially weighted parsimony and maximum likelihood models, are only weakly supported by the traditional indices. This methodological concordance is taken as evidence for insensitivity of the deep structure of the topology to alternative assumptions. [Biogeography; character congruence; maximum likelihood; parsimony; phylogeny; Sceloporus; weighted parsimony.]

With increasingly large data sets for phylogenetic inference, systematists should find reason to feel confident in their results. However, multiple data sets also increase the potential for conflicting hypotheses and lead to questions about how data should be combined, excluded, weighted, and modeled. Even if a consensus of opinion existed for the analysis of data under a particular optimality criterion, we must still choose among equally versus differentially weighted parsimony (EP vs. WP), maximum likelihood (ML), and minimum evolution approaches, each with its own assumptions and limitations.

Simulation studies and investigations of experimentally generated phylogenies have shown that some methods of phylogenetic inference perform better than others for a variety of conditions such as branch lengths, mutation rates, and data set sizes (Hillis et al., 1992, 1994a, 1994b; Huelsenbeck, 1995). However, those studies are often conducted under simplified conditions in comparison with the complexities of multiple data sets that have evolved under different constraints. Consequently, no general consensus has been reached about the "best" approach to phylogeny reconstruction. For example, assume for the sake of argument that under ideal conditions, ML analysis is always the superior approach. In real data sets, the difficulty in modeling insertion/deletion events and morphological characters usually results in their being excluded from phylogenetic analysis. Further, the computational complexity of a likelihood analysis almost always dictates a more cursory search of tree space, at least with large data sets. Is a likelihood approach with these limitations still superior to a parsimony analysis that includes more characters and explores tree space more thoroughly? We do not know the answer to this question.

Potential solutions to these issues include the adoption, on the basis of philosophy, of a single optimality criterion (Miyamoto and Fitch, 1995a; Siddal and Kluge, 1997) or, conversely, the use of a wide range of methods (e.g., parsimony, distance, likelihood; Kim, 1993; Hâstad and Björklund, 1998).

[5]Author for correspondence.

Presumably, analysis and comparison of empirical data sets will influence opinions on methods, just as analyses of simulated data sets have distinguished differences in performance under ideal conditions. Here we present an empirical data set to evaluate a phylogenetic hypothesis for basal clades of the lizard genus *Sceloporus*. In the process, we utilize a variety of data to identify nodes that are strongly supported by character congruence.

The majority of our analyses use some form of parsimony, which permits the simultaneous analysis of different data partitions that lack adequate evolutionary models, is relatively insensitive to missing data and taxa, and is computationally feasible with large data sets (Wiens, 1998a; Wiens and Reeder, 1995, 1997). This computational feasibility also permits a more thorough search strategy, but performance may be strongly influenced by variation in branch length (Felsenstein, 1978; Huelsenbeck and Hillis, 1993; Huelsenbeck, 1995). WP can compensate for some of the drawbacks of EP, however, by reducing the influence of characters that are likely to have had multiple substitutions (Chippindale and Wiens, 1994; Cunningham, 1997a, 1997b; Voelker and Edwards, 1998).

ML methods are also less sensitive to branch length biases than is parsimony but may be sensitive to the model of nucleotide substitution (Hasegawa and Fujiwara, 1993). Likelihood can also be used to test for the effects of nucleotide compositional bias, substitution differences, and among-site rate variation (Huelsenbeck and Crandall, 1997; Huelsenbeck and Rannala, 1997). Both approaches assume character independence, which may not always be met for either molecular or morphological data (Miyamoto and Fitch, 1995b; Emerson and Hastings, 1998).

We explored concordance in tree topology among three different optimality criteria—EP, two methods of WP, and ML—for basal clades of the lizard genus *Sceloporus*. The value of using *Sceloporus* in this study derives from its morphological and ecological diversity. It has been used as a model system for testing a variety of behavioral, demographic, physiological, and evolutionary questions (reviewed by Sites et al., 1992), yet some of the most difficult phylogenetic issues center on relationships among the basal groups (Wiens and Reeder, 1997).

Phylogenetic Hypotheses of Sceloporus

Previous hypotheses of relationships within *Sceloporus*, shown in Figure 1, have been proposed on the basis of a classical taxonomic study of morphological characters (Smith, 1939), and a numerical phenetic study based on chromosomal, distributional, and morphological data (Larsen and Tanner, 1974, 1975). The Smith phylogeny recognized distinct "small-bodied, small-scaled" (SB/SS) and "large-bodied, large-scaled" (LB/LS) radiations, whereas Larsen and Tanner (1975) recognized three major groups, and resurrected the genus *Lysoptychus* Cope, for the species that belonged to their group I (Fig. 1; see also Sites et al., 1992:figs. 7, 8). Hall (1973, 1980, 1983) presented a hypothesis based on chromosome data (summarized by Sites et al., 1992), recognizing monophyletic SB/SS (without any evidence) and LB/LS radiations, and considered *Uta* and other genera to be external to *Sceloporus* (Fig. 1; see also Sites et al., 1992:figs. 19, 26). More recently, Wiens and Reeder (1997) used mitochondrial DNA (mtDNA) sequences (12S and 16S ribosomal RNA [rRNA] gene regions), internal and external morphology, coloration, chromosome, and life history characters to infer a phylogeny for the genus. This hypothesis represents the most complete phylogenetic analysis to date, both in number of characters and number of species sampled, and proposes that the SB/SS radiation is not monophyletic (Fig. 1).

At a higher taxonomic level, hypotheses regarding monophyly of the genus *Sceloporus* with respect to the genus "*Sator*" have been conflicting. Etheridge and de Queiroz (1988) proposed that these two form a monophyletic group within a more-inclusive sceloporine clade ((*Petrosaurus*) ((*Uta*) ((*Urosaurus*) (("*Sator*") + (*Sceloporus*))))), but presented no synapomorphies for *Sceloporus*. Frost and Etheridge (1989), analyzing a similar data set (internal and external morphological characters), concluded that *Sceloporus* was not monophyletic if "*Sator*" was retained as a genus, and therefore "*Sator*" should be synonymized with *Sceloporus*. Wyles and Gorman (1978) also concluded that the genus "*Sator*" should be synonymized with *Sceloporus*, on the basis of immunological distance data. Wiens (1993) presented a cladistic analysis of relationships within the *Sceloporus* group of the family Phrynosomatidae, on the

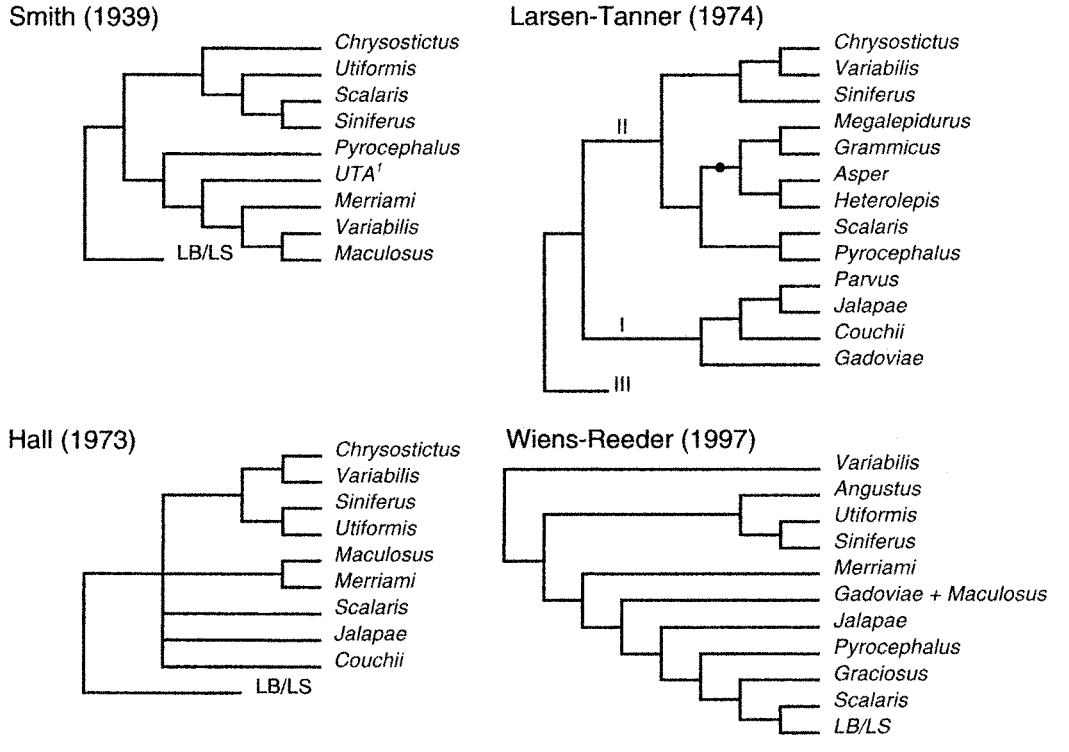

FIGURE 1. Selected alternative hypotheses for basal species group relationships within *Sceloporus*. In the proposal of Smith (1939), UTA[1] designates a group in which the genera *Uta* and *Urosaurus* are nested within the small-bodied, small-scaled (SB/SS) group of *Sceloporus*. In this and all other trees, LB/LS refers to the large-bodied, large-scaled radiation of *Sceloporus*. Smith (1939:29) recognized "*Sator*" as a distinct genus allied to *Sceloporus*. In the proposal of Larsen and Tanner (1974, 1975), I represents the resurrected genus *Lycoptychus* Cope; II represents a reorganized SB/SS group, which contains a clade defined by the inclusion of a monophyletic group (denoted by the solid circle) of four species groups formerly assigned to the LB/LS radiation; and III represents the reorganized LB/LS radiation. In the proposal of Hall (1973, as summarized by Sites et al. 1992), based on chromosome characters, "*Sator*" is considered among the ancestral genera external to *Sceloporus*. In the proposal of Wiens and Reeder (1997), their *Angustus* group includes both species of "*Sator*", in synonomy with *Sceloporus*.

basis of morphological, chromosomal, and behavioral data, and recovered a topology of (("*Sator*") (*Sceloporus merriami* (all other *Sceloporus*))).

More recently, Reeder's (1995) analysis of mtDNA 12S and 16S gene regions recovered *Sceloporus* as a polyphyletic genus in one of three analyses, but this was weakly supported. "*Sator*" was placed in a separate clade with species of the *variabilis* group, when only transitions were used in the analysis. However, when only transversions were analyzed, *Sceloporus* was weakly recovered as monophyletic (Reeder, 1995:fig. 4). Reeder and Wiens (1996:figs. 5, 6) added 155 morphological, chromosomal, and behavioral characters in a further analysis and recovered a monophyletic *Sceloporus* with "*Sator*" as its sister taxon.

The more-inclusive analysis of Wiens and Reeder (1997) recovered a hypothesis that nested "*Sator*" within *Sceloporus* (Fig. 1), but the morphological data alone still (weakly) placed "*Sator*" as the sister taxon of *Sceloporus*, and no tests were performed to distinguish between these rival hypotheses. Schulte et al. (1998) recently proposed an alternative arrangement in which "*Sator*" is presented as the sister taxon to a ((*Petrosaurus*) ((*Sceloporus*) + (*Urosaurus*))) clade, and we return to this issue in the Discussion.

The present study is based on new characters, including regions of mitochondrial protein-coding (ND4) and RNA genes (tRNAs and some new 12S and 16S rRNA sequences), nuclear markers (aldolase intron sequences, allozymes, and chromosomes), and the previously published

morphological/ molecular data set of Wiens and Reeder (1997). We also include molecular data for SB/SS species that were not included by Wiens and Reeder (1997). Our objectives in this study are twofold. First we want to develop a strongly supported phylogeny of (primarily) the SB/SS species of *Sceloporus*. The strength of our hypothesis is based on evidence from character concordance; nodes consistently supported by characters obtained from independent data partitions are considered strongly supported.

Our second objective is to statistically test our hypothesis against various alternatives: (1) the nested position of "*Sator*" within *Sceloporus*, versus the alternative in which "*Sator*" is the sister taxon of a monophyletic *Sceloporus* (the "*Sator* out" hypothesis); (2) the monophyly of the SB/SS radiation (Smith, 1939) (Fig. 1); (3) the monophyly of *Lysoptychus* (Larsen and Tanner, 1975) (Fig. 1); (4) the basal position of *S. merriami* within a monophyletic *Sceloporus* (Wiens, 1993); and (5) the arrangement of the *angustus* (= former "*Sator*"), *utiformis*, and *siniferus* species groups proposed by Wiens and Reeder (1997) (Fig. 1). The issue here is whether alternative tree topologies can be rejected because they are significantly less parsimonious, or have a lower likelihood, than does a strongly supported hypothesis derived from multiple data sets. This is important because if systematists can never define by some criterion the nodes that are probably accurate, and separate them from previous hypotheses that are probably wrong, then additional studies may only contribute more conflict to an increasingly confusing forest of trees. The purpose of this study is to develop a fully tested phylogenetic hypothesis for the SB/SS groups of *Sceloporus*—supporting some nodes, rejecting others, and leaving still others unresolved if support is ambiguous. The placement of "*Sator*" is of biogeographic interest because it represents one of the few island endemics with a potential sister taxon on the Mexican mainland (Grismer, 1994a, b).

Materials And Methods

Collection of Specimens

One hundred eight specimens were collected during 1993 and 1994 from 72 localities in Mexico and the USA. The 33 species (of ~80 in the genus, according to Sites et al., 1992) included all recognized species in the SB/SS radiation, as originally defined by Smith (1939). For some species, two individuals from different parts of the range were included in the sequence data base, because electrophoretic studies revealed fixed allozyme differences between localities. All museum vouchers and localities are listed in the Appendix.

Data Sets

Horizontal starch-gel electrophoresis was performed with samples from 83 individuals belonging to 23 species and subspecies of *Sceloporus*, "*Sator*", and *Urosaurus* (Appendix), following the protocols of Mink and Sites (1996). Three groups of tissues (skeletal muscle, stomach/duodenum, and liver/heart/kidney) were screened separately for different subsets of 27 loci. We used 21 fully resolved isoenzyme loci for the phylogenetic analysis. We followed the stain/buffer protocols described by Murphy et al. (1996) and used buffer systems in several combinations to resolve more charge variation than would be apparent by "single pass" electrophoresis (Coyne, 1982; Barbadilla et al., 1996). All enzymes, gene loci, buffer systems, and tissue combinations studied are listed in Table 1.

We incorporated the matrix of molecular and morphological characters from Wiens and Reeder (1997). This file was used exactly as in Wiens and Reeder (1997), except that four polymorphic multistate morphological characters were deleted because of differences in geographic sampling between the specimens used in that study and the ones we collected. Specifically, Wiens and Reeder (1997) coded characters 55 (number of postrostral scales), 146 (dark nuchal collar pattern), 160 (color pattern on posterior surface of thigh), and 172 (female gular coloration) by the "majority method", in which the condition found in the majority of individuals within a sample was designated as the character state for that sample. Our sampling design did not permit us to derive a defensible coding scheme in the majority of taxa for these four characters. Otherwise the morphological matrix was used exactly as in Wiens and Reeder (1997), and although we refer to EP throughout the text, all analyses of these data included polymorphic characters that were ordered and coded according to the

TABLE 1. Stain and electrophoretic conditions used in this study; enzyme names, Enzyme Commission (EC) numbers, and locus acronyms follow recommendations of the International Union of Biochemistry Nomenclature Committee (1984).

Enzyme	EC no.	Locus	Tissue[a]	Buffer[b]
Aconitate hydratase	4.2.1.3	M-Acoh-A	L	A
		S-Acoh-A	L	A
Adenylate kinase	2.7.4.3	Ak	M	A, B
Adenosine deaminase	3.5.4.4	Ada	S	A, F, I
Aspartate aminotransferase	2.6.1.1	M-Aat-A	L	E, H
		S-Aat-A	L	E, H
Creatine kinase	2.7.3.2	Ck-A	M	A, C
		Ck-C	S	A, D
Fructose-biphosphatase	3.1.3.11	Fbp	L	A
Fumarate hydratase	4.2.1.2	Fumh	L	A, G
Glucose-6-phosphate isomerase	5.3.1.9	Gpi	L	A
Glutamate dehydrogenase	1.4.1.2	Gtdh	L	A, G, H, J
Glycerol-3-phosphate dehydrogenase	1.1.1.8	G3pdh	M	A
Isocitrate dehydrogenase	1.1.1.42	Idh	L	A, J
L-Lactate dehydrogenase	1.1.1.27	Ldh-A	L	A, J
		Ldh-B	L	A, J
Malate dehydrogenase	1.1.1.37	M-Mdh-A	L	A, E, G
		S-Mdh-A	L	A, E, G
α-Mannosidase	3.2.1.24	α-Man	L	G, H
Mannose-6-phosphate isomerase	5.3.1.8	Mpi	L	A
Peptidases	3.4.-.-			
Glycyl-L-leucine		Pep-A	S	A, E
L-Leuclyglycylglycine		Pep-B	S	A, E
L-Leucyl-L-alanine		Pep-?	M	A
Phosphoglucomutase	5.4.2.2	Pgm-A	L	A
Purine-nucleoside phosphorylase	2.4.2.1	Pnp	S	A, F, I
Superoxide dismutase	1.15.1.1	Sod	L	A
Triose-phosphate isomerase	5.3.1.1	Tpi	S	A, F, I

[a] L = liver, heart, and kidney, M = skeletal muscle, and S = stomach and duodenum.

[b] A = Tris–citrate II, pH 8.0, ~10 h at 75 mA; B = phosphate–citrate pH 7.0; C = borate discontinuous, electrode pH 8.0, gel pH 8.6; D = lithium borate, electrode pH 8.1, gel pH 8.3; E = poulik discontinuous, electrode pH 8.2, gel pH 8.7, ~12 h at 200 V; F = Tris–citrate II, pH 8.0, ~10 h at 75 mA; G = Tris–citrate–EDTA, pH 7.0; H = Tris–borate–EDTA II, pH 8.0; I = Tris–citrate III, pH 8.0; J = Tris–borate–EDTA I, pH 8.6.

bins method described by Wiens (1995). This leads to a 24-fold weighting of all fixed character states— allozyme, chromosome, molecular, and morphological (the latter coded as 0 or 1 in this matrix)—relative to the frequency states ordered by bins. Tree lengths estimated from the combined data were then divided by 24 to make them comparable with those calculated from the nonmorphological partitions (see the combined evidence analysis below, for an example).

A total of 37 specimens were sampled from 33 taxa within the genera *Sceloporus*, "*Sator*", *Petrosaurus*, and *Urosaurus* (Appendix) for DNA characters. DNA was extracted by the method of Hillis and Davis (1986), and the polymerose chain reaction (PCR) amplification and sequencing protocols followed those described in Benabib et al. (1997). Sequences were collected for part of the mitochondrial protein-coding ND4 gene; serine, histidine, and leucine tRNAs; regions of the 12S and 16S rRNAs; and parts of the nuclear aldolase-A and aldolase-C introns. Homology of the aldolase introns was inferred from the size of the PCR products (Lessa and Applebaum, 1993:fig. 5) and by matching nucleotide similarities to GenBank records. The external primer pairs for each of these amplified fragments are those of Arévalo et al. (1994) for the ND4 and tRNA genes, Reeder (1995) for the 12S and 16S genes, and Lessa and Applebaum (1993) for the nuclear aldolase genes.

Sequence Alignments

Because there was no length variation in the *Sceloporus* ND4 sequences, alignment of these data with published *Sceloporus* ND4 sequences (Arévalo et al., 1994) was trivial (complete ND4 sequences are available in GenBank under accession numbers AF 210332–210368). The low divergence and conserved nucleotide composition of the

aldolase sequences, together with very few insertion–deletion (indel) differences, permitted manual alignment (see alignments at the Society of Systematic Biologists web site, http: // www. utexas.edu / depts / systbiol /). The tRNAs were aligned on the basis of the vertebrate secondary structure model proposed by Kumazawa and Nishida (1993), whereas 16S and 12S alignments followed the secondary structure models of Gutell (1994) and Gutell et al. (1994). Alignment notations for RNA sequences follow Kjer et al. (1994) and are summarized in Figure 2. Three small regions (46 nucleotides) of the 16S sequence that could not be aligned were excluded from the analysis (Fig. 2). The precise locations of these alignment-ambiguous regions were defined by identifying hydrogen-bonded flanking positions that were confirmed by compensatory base changes (Kjer, 1997). Both manual and computer alignments can be criticized for various reasons (summarized by Kjer, 1995, 1997), but if aligned sequences are available, then results are repeatable and the sensitivity of alignments to different variables can be assessed. Our alignments can be reconstructed from Figure 2 and are provided in the file at the Society of Systematic Biologists website.

Character Coding

We coded chromosome characters previously published and compiled by Sites et al. (1992) for phylogenetic analyses, following general methods suggested by others (Frost and Tim, 1992; Borowik, 1995; Wiens and Reeder, 1997) to create a data matrix for PAUP. Specifically, the six pairs of macrochromosomes for which between-species homology could be reasonably inferred on the basis of chromosome size and shape were scored as unordered multistate characters. Each pair was represented by state 0 (biarmed element), 1 (fissioned), or 2 (pericentric inversion). Five pairs of microchromosomes were coded as 0 (plesiomorphic condition, as seen in the $2n$ = 34 karyotype; see Sites et al., 1992:fig. 13) or 1 (to denote reduction of microelements by fusion). Another independent character, the location of the single pair of nucleolus-organizing regions (inferred from the location of secondary constrictions in the data summarized by Sites et al., 1992; see Porter et al., 1991), was coded either as 0 (present in the long arm of macrochromosome pr 2) or 1 (present in a microchromosome). The sex chromosomes provided another unordered multistate character coded to match several alternative heteromorphisms: 0, XY system with reduced Y; 1, X_1X_2Y (Y element translocation to a microautosome); 2, X_1X_2Y (Y translocation to a macroautosome); 3, pair 7 heteromorphism; 4, XY (large acrocentric Y); and 5, XX (Y indistinct). These heteromorphisms are illustrated by Sites et al. (1992:fig. 18). Finally, the enlarged micro-9 element provided the 14th independent character; it was scored as either present (1) or absent (0). This coding scheme differs slightly from that of Wiens and Reeder (1997) by (1) considering the five microchromosome pairs as separate characters (vs. a single multistate character); (2) including the nucleolus-organizing region location as another character; and (3) scoring the sex chromosome system as an unordered multistate character with five alternative states.

Isoenzyme loci were scored by considering the locus as the character and alternative electromorphs as unordered states. For a given locus, homology was inferred if electromorphs retained identical mobilities after side-by-side comparisons on multiple buffers. Coding for phylogenetic analysis followed Mink and Sites (1996) for conservative markers (i.e., those characterized by low intraspecific but high interspecific variation), and the step matrix methods described by Mabee and Humphries (1993) and Wiens (1995) were used for the few intraspecific polymorphic loci in some of the weighted parsimony analyses. Results for the isoenzyme step matrix approaches were incongruent with each other, and with the majority of other analyses carried out in this study; however, because it is beyond the scope of this paper to explore the reasons for these differences, these results are not considered further.

Nucleotides were treated as unordered characters with four alternative states, and basic summary statistics for all partitioned sequences were calculated with MEGA (Kumar et al., 1993) and PAUP* (test version 4.0.0d52; Swofford, pers. comm.). Gap characters in the RNA sequences were coded as single presence/absence characters, regardless of their length. To accomplish this, indels were treated as missing data in the nucleotide

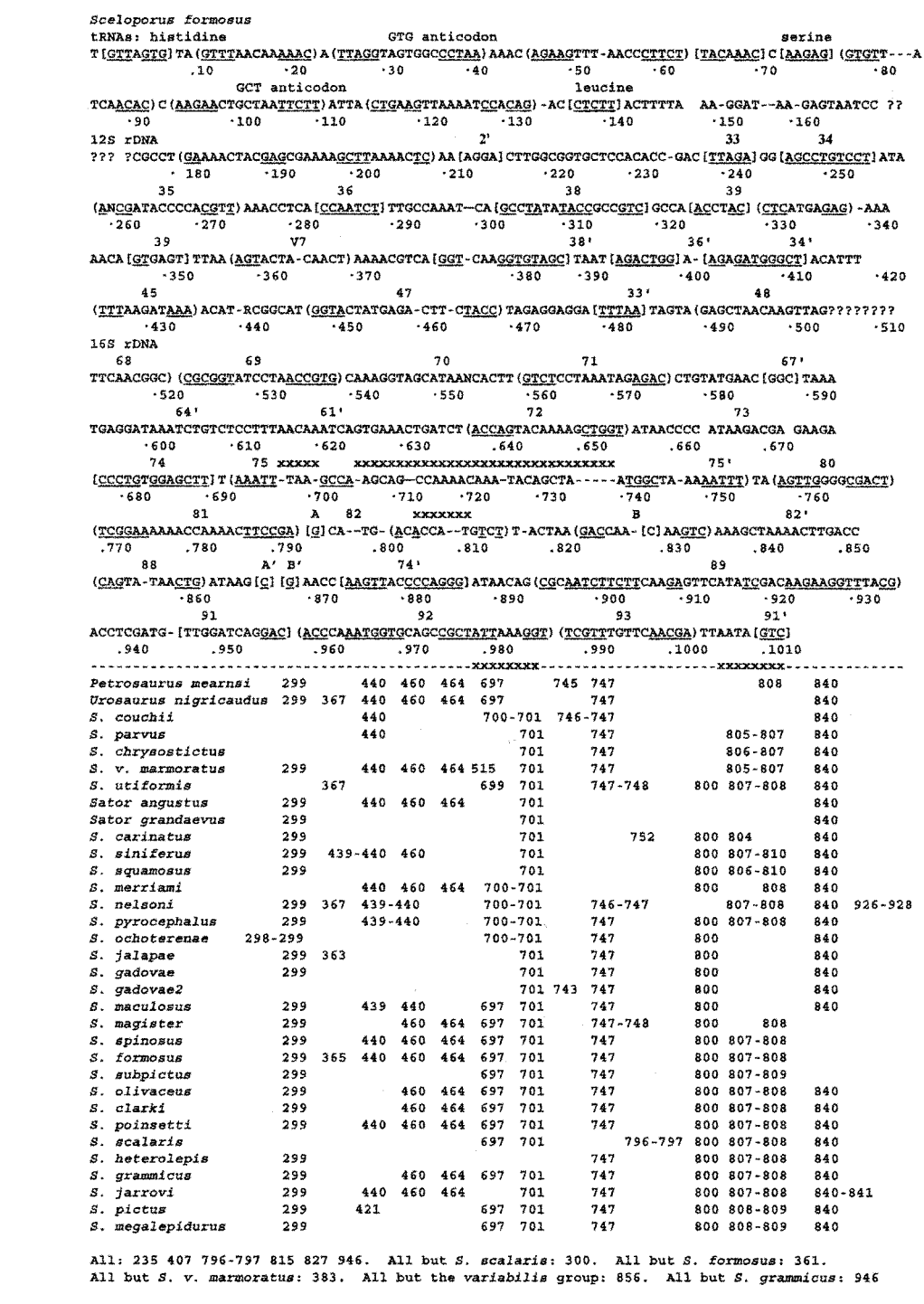

FIGURE 2. Alignment and structure of 12S, 16S, and tRNA sequences, according to the notations recommended by Kjer et al. (1994); parentheses, brackets, and underlining respectively identify stem sequences narrowly separated from their complementary sequences, regions involved in base pairing where regions are separated by other hydrogen-bonded stems, and nucleotides paired in a stem (see also Kjer, 1995:fig. 2). Structural models follow Gutell (1994) and are numbered above the sequence according to the system of van de Peer et al. (1993) for 12S, Larsen (1992) for 16S, and Kumazawa and Nishida (1993) for tRNAs. Nucleotide positions are numbered below the sequence, whereas regions of ambiguous alignment that were excluded from the analysis are identified with reference to secondary structure, and are denoted by "xx" above the positions. Below the *S. formosus* sequence is a summary showing where gaps were inserted into all other taxa, relative to the reference sequence.

data matrix, and then a separate indel matrix was constructed that consisted of gaps coded as 0 and nucleotides as 1. The gaps coded this way were then downweighted according to the inverse of the length of the longest deletion in the region from which they came (Kjer, 1995). This was done because it is impossible to trace the number of events that led to an observed indel, and the regions in which we observed extreme length variations are by the nature of their viability (their existence in living organisms) proven to freely permit length variation.

Outgroup Choice

We chose *Petrosaurus* and *Urosaurus* as the second and first outgroups for this study, respectively based on previous phylogenetic work showing a ((*Petrosaurus*) ((*Urosaurus*)(*Sceloporus* + "*Sator*"))) topology (Wiens, 1993; Reeder and Wiens, 1996; but see Schulte et al., 1998). In some cases of the EP analyses of mtDNA sequence data, and in all analyses of the aldolase sequences, the outgroup topology was constrained as above. This was deemed necessary for aldolase because complete fragments of some of the outgroup taxa could not be amplifed. These constraints also allow PAUP to reconstruct missing data, and the outgroups can be used as separate taxa rather than a hypothetical combination. An unconstrained analysis of the ND4 and combined mtDNA data places *Petrosaurus* inside of *Sceloporus*, and because all previous studies contradict this result, as do all of our unconstrained WP and ML analyses, we consider the inclusion of this constraint justified for exploratory searches (Moritz et al., 1992; Ballard et al., 1998), given that placement of *Petrosaurus* external to *Sceloporus* is not an issue.

Phylogenetic Analysis

EP analyses were performed with PAUP 3.1 (Swofford, 1993) and PAUP* (test version 4.0.0d52; Swofford, pers. comm.). For each data set, 50 heuristic searches were performed, using the options of (random) stepwise addition and tree-bisection reconnection. All data sets were first analyzed separately and then in the following combinations: (mtRNA sequences), (all mitochondrial sequences), (aldolase sequences), (all DNA characters), and (all data, including morphology).

Robustness of the results was evaluated by the decay index of Bremer (1988) and by bootstrap analysis based on either 1,000 pseudoreplicates (Felsenstein, 1985) for the combined data trees or 100 pseudoreplicates with 10 random-addition heuristic searches in each pseudoreplicate for all of the other partitioned data sets. We consider a decay index of 5 and a bootstrap proportion of 95 to be strong support for a clade but recognize that these values are arbitrary.

WP Analyses

Two WP methods were used to accommodate differences in character evolution (Chippindale and Wiens, 1994) and to explore the stability of the proposed phylogeny under a variety of assumptions. One method assumed that transitions, transversions, and nonsilent substitutions occur at different rates; to reflect this without discarding the more rapidly changing sites, we constructed a triple matrix of the data as described by Benabib et al. (1997). Briefly, the entire mitochondrial nucleotide data set was duplicated, and the copy was converted to a matrix that consisted of only binary characters by changing all G's into A's and all C's into T's (so that the altered matrix contains only A's and T's, which represent purines and pyrimidines). The converted copy was then added to the end of the original DNA data. The protein-coding portion of the original data was then copied again, and this time converted into amino acids, which were used as the third section of the same data set. Although this clearly violates the assumption of character independence, a justification for a similar scheme can be found in Agosti et al. (1996). This justification can be summarized by the observation that there are only two possibilities for the second and third (altered) passes of the original data: Additional characters will either support relationships that were supported in the previous pass, in which case the new state is appropriately upweighted, or they will disagree, in which case the new state cannot pose a problem of nonindependence because it is contradictory to its linked alternative. When considered as a weighting scheme, this can be described as follows. Nonsilent substitutions, regardless of their position in the codon, are weighted either as 2 if they resulted from transitions or as 3 if they resulted from transversions.

Silent transversions are weighted as 2 and silent transitions are weighted as 1.

Nucleotide substitution rates also differ among genes, particularly between nuclear and mitochondrial genes. To assess the influence of this heterogeneity, we applied a step matrix–weighting scheme to accommodate different substitution rates in each of the DNA partitions. This required the construction of four step matrixes, one for each of the following partitions: ND4, mtRNAs, aldolase, and indel characters (Table 2). The step matrixes for the first three of these were constructed according to the inverse probability of the different kinds of substitutions (Felsenstein, 1981a). The ML tree was imported into the PAUP file as a starting tree, and the optimality criterion was set to "Likelihood". We then used PAUP to estimate the instantaneous rates of all possible symmetrical substitution types, individually on each of these three DNA partitions, in a 4×4 matrix under a general reversible likelihood model for the imported tree. Table 3 shows that the values of several other trees are very similar to those based on the ML tree, so the step matrix values are relatively tree-independent (Yang et al., 1995).

Once the 4×4 matrices were constructed for each partition, we normalized the values by dividing each by the sum for its respective row. Then, to downweight the more common kinds of substitutions, we took the inverse of these values for use in the step matrix. These numbers were then multiplied by 10 and rounded to integers to obtain the values shown in Table 2. Rather than arbitrarily eliminating indel characters, values were taken from the lowest value in any of the other step matrixes. Finally, we weighted each of the partitions by the inverse of their relative rates, which were measured by using the "charpartition" option of PAUP. This weighting scheme was designed to accommodate heterogeneity among data partitions in a single analysis. Both of these weighting methods allow for heterogeneity of substitution classes within and among molecular partitions, and neither appears strongly influenced by any initial tree topology (in contrast to successive weighting [Farris, 1969]; see Cunningham, 1997b).

TABLE 2. Step matrix used for each DNA partition in the weighted parsimony analysis.

	To:			
From:	A	C	G	T
ND4				
A	0	54	16	57
C	78	0	314	12
G	12	162	0	123
T	83	12	239	0
RNAs				
A	0	50	19	37
C	73	0	279	12
G	12	115	0	241
T	56	12	605	0
Aldolase				
A	0	55	15	73
C	39	0	82	16
G	12	97	0	109
T	48	15	85	0
Gaps				
	present			absent
Present	0			12
Absent	12			0

ML Analyses

ML analyses were performed for all mtDNA and aldolase characters, although we found it not feasible to carry out a full likelihood estimation with what we considered an adequate search strategy. For this reason we took the following shortcuts. First, parameters were determined by a full likelihood estimation from six trees: (1) a tree taken from the non-DNA hypothesis of Reeder and Wiens (1996); (2) the combined data tree of Wiens and Reeder (1997); (3) the shortest tree resolved from the EP analysis of our DNA data; (4) our combined DNA ML tree; (5) a random tree from near the center of the distribution of 10,000 random trees; and (6) a short random tree selected from the shortest 5% of 10,000 random trees. We could then compare random trees of different lengths with optimal trees constructed from either our DNA data or another source we deemed reliable (e. g., the non-DNA data tree of Reeder and Wiens, 1996; or the combined data tree of Wiens and Reeder, 1997).

The model we selected for the likelihood searches was based on the values shown in Table 3; we also implemented the MODELTEST program (ver. 2.0; Posada and Crandall, 1998) to identify the most appropriate substitution model for all DNA partitions combined. MODELTEST applies hierarchical likelihood ratio tests to estimate likelihood scores for nested and nonnested sets of substitution models for a given user tree. We used the default settings in

TABLE 3. Summary of among-site rate variation (α) and likelihood values for the nucleotide composition and rate variation of the three sets of DNA data (RNAs include 12S, 16S, and tRNA sequences). Constant sites and substitution rates were estimated by likelihood in PAUP* (the R matrix).

	Reeder and Wiens, 1996	Wiens and Reeder, 1997	EP tree	ML tree	Random tree	Short random tree
ND4 (687 nts)						
−Ln likelihood	10884.20	10803.82	10807.25	10767.58	11748.45	11677.82
α	0.891	0.929	0.940	0.928	0.595	0.498
Estimated Invariable	0.392	0.393	0.393	0.392	0.398	0.343
A–C	2.945	2.913	2.987	3.065	3.748	3.247
A–G	9.777	10.023	10.248	10.553	9.724	9.327
A–T	2.848	2.823	2.940	2.896	3.812	3.812
G–C	0.688	0.681	0.776	0.760	0.824	0.987
G–T	1.000	1.000	1.000	1.000	1.000	1.000
C–T	19.466	19.823	19.237	20.017	34.757	26.904
RNAs (964 nts)						
−Ln likelihood	6501.91	6433.77	6420.26	6401.14	6950.08	6914.98
α	0.588	0.585	0.601	0.586	0.468	0.502
Estimated Invariable	0.481	0.469	0.471	0.465	0.499	0.504
A–C	8.651	9.233	8.164	8.019	12.419	12.876
A–G	21.804	23.914	21.040	21.061	27.966	26.989
A–T	11.151	12.401	11.083	10.752	18.740	16.823
G–C	2.108	2.420	2.108	2.111	2.279	2.336
G–T	1.000	1.000	1.000	1.000	1.000	1.000
C–T	52.091	55.872	48.483	48.752	77.740	76.218
Aldolase (420 nts)						
−Ln likelihood	1701.48	1692.02	1696.81	1692.12	1779.34	1788.25
α	0.636	0.533	0.704	0.669	0.790	0.694
Estimated Invariable	0.172	0.071	0.179	0.160	0.355	0.336
A–C	2.517	2.495	2.379	2.364	3.061	3.494
A–G	8.828	8.699	8.913	8.753	10.545	10.990
A–T	1.822	1.779	1.813	1.760	2.136	2.092
G–C	1.104	1.105	1.146	1.128	1.097	1.100
G–T	1.000	1.000	1.000	1.000	1.000	1.000
C–T	5.885	5.508	5.819	5.741	7.132	7.352
Combined						
−Ln likelihood	19403.28	19367.33	19201.57	19183.34	20754.72	20635.04
α	0.707	0.724	0.700	0.700	0.569	0.569
Estimated Invariable	0.461	0.464	0.452	0.452	0.479	0.479
A–C	5.454	5.073	5.297	5.310	6.904	6.269
A–G	12.335	11.377	12.326	12.340	12.342	11.252
A–T	5.867	5.231	5.681	5.610	7.961	7.271
G–C	1.114	1.009	1.171	1.150	1.098	1.070
G–T	1.000	1.000	1.000	1.000	1.000	1.000
C–T	30.040	27.061	29.220	29.530	40.646	36.823

nts, nucleotides.

MODELTEST to calculate a neighbor-joining tree as the test topology and began by comparing a simple Jukes–Cantor (1969) model (JC; equal base frequencies) with Felsenstein's (1981b) model (F81), which allows unequal base frequencies. In this paired log-likelihood InL ratio test, the JC69 −lnL = 23258.37 versus −ln L = 23026.73 for F81; and with 3 df, the JC69 model is rejected at $P < 0.000001$. This test was then repeated to test the F81 assumption of equal transition and transversion rates, in comparison with a simple model in which these rates may be unequal (Hasegawa, Kishini, and Yano, 1985; the HKY85 model). Here $-\ln L_{F81} = 23026.73$ and $-\ln L_{HKY85} = 22332.68$; with 1 df we reject the F81 model at $P < 0.000001$. The final test compared HKY85 (which assumes equal rates within both transversions and transitions) against the general time-reversible (GTR) model of Rodríguez et al. (1990), in which all substitution probabilities are assumed to be independent; here $-\ln L_{HKY85} = 22332.68$ versus $-\ln L_{GTR} = 21913.59$, and

with 3 df we reject the HKY85 model at $P < 0.00001$. Further iterations comparing the GTR model with more parameter-rich models did not improve the fit of our data to any of the latter models, so we used the GTR model option in PAUP for all ML searches. To examine heterogeneity in among-site rate variations (ASRV) in our sequence data sets, we estimated the gamma distribution shape parameter (α, using the "tree scores" option in PAUP*) by way of ML analysis, for each of the partitioned DNA data sets.

Because we did not know how long a search would take, we took a second shortcut by constraining all nodes that were supported in the EP analysis by 100% bootstrap and 5+ decay index values. These constraints allowed us to conduct 10 heuristic random-addition likelihood searches. On the second search, we used the parameters from the Wiens and Reeder (1997) combined data tree (Table 3), and relaxed the constraints to conspecific taxa only. Again, this option allowed us to conduct 10 random-addition likelihood searches. To evaluate whether any of our constraints affected the results, we implemented a single heuristic search of all taxa without constraints (this required 8 days of computation on a 240 MHz computer). The searches all resulted in identical tree topologies. We then estimated parameters on our likelihood tree (no. 4 above) and ran another search with parameters input from this tree, constraining only the conspecific taxa. We could thus evaluate the sensitivity of our likelihood searches to the range of parameters used (see Table 3), and could include parameter estimates from the optimal tree under likelihood.

Statistical Tests of Alternative Hypotheses

Our most strongly supported trees were tested against alternatives (topologies in Fig. 1); we also tested the basal position of *S. merriami* as proposed by Wiens (1993). We used the conservative nonparametric Wilcoxon matched-pairs signed-rank test (originally described by Templeton, 1983) as implemented by Larson (1994), to test the fit of constrained alternative topologies to our data. We also evaluated these hypotheses in terms of likelihood by the Kishino and Hasegawa (1989) test. Both tests were implemented with the Tree Scores option of PAUP; each of the parsimony trees was evaluated in terms of likelihood by estimating α, GTR substitution frequencies, and invariant sites. We also tested the phylogenetic position of "*Sator*", using the parametric bootstrap test (Huelsenbeck et al., 1996a,b).

RESULTS

Patterns of Variation

Table 4 summarizes patterns of variation for all data partitions collected in this study. Chromosome and isoenzyme data sets were relatively small and were informative at only a few nodes. Aldolase sequences were informative, but because our primers failed to amplify some taxa, this partition is the least complete. All mtDNA partitions contained large numbers of parsimony-informative sites, although the proportion of these to the total number of variable sites varied among the different partitions (41% for tRNAs to 81% for ND4; Table 4). Tests for stationarity of nucleotide composition (Collins et al., 1994) showed no significant deviations among taxa for any sequence partitions (Table 4).

ASRV, indicated by values of α, was present in all of the partitions but in different amounts (Table 3). Each of our sequence data sets contained many nucleotides that were invariable, some that changed once on a tree, and others that changed many times, but histograms describing this ASRV among different data partitions differed in their precise shape (results not shown). Sullivan (1996) noted that overlap between data sets with different ASRVs could be used to justify a combined analysis of these partitions (i.e., simulated data partitions might have completely nonoverlapping distributions of rates, but biological data partitions are likely to have overlapping distributions). Agreeing with Sullivan (1996) on this issue, we combined the data partitions even though the ASRV values were not identical. Although we accept that combination may not always lead to a superior estimation (Bull et al., 1993), recent empirical studies of these issues show that, under a variety of conditions, combining data partitions frequently improves both character congruence and phylogenetic accuracy (Barrett et al., 1991; Cunningham, 1997a,b; Wiens, 1998a), and in cases where substantial conflict between data partitions is not improved by combining them, phylogenetic accuracy may still be improved by increasing the size of the

TABLE 4. Character variation and nucleotide base composition of the sequences, for the data partitions used in this study, averaged across all taxa; χ^2 and P values are tests for stationarity (Collins et al., 1994).

Sequence	Character no.	No. variable sites (V)	No. informative sites (I)	Ratio I/V sites	% content (and range)				χ^2 (df)	P
					A	G	C	T		
ND4	687	417	338	0.81	33.4 (30.7–36.4)	11.0 (9.3–12.3)	28.9 (26.4–32.0)	26.7 (23.5–30.7)	64.58 (99)	0.99
tRNAs[a]	171	113	51	0.45	36.7 (33.1–39.9)	15.4 (13.4–17.8)	19.6 (17.1–24.0)	28.3 (24.0–31.7)	33.44 (99)	1.00
12S	341	167	83	0.50	35.4 (30.8–39.8)	19.3 (17.2–22.0)	23.0 (19.7–25.5)	22.3 (19.9–25.5)	22.28 (87)	1.00
16S[b]	453	182	118	0.65	35.1 (33.2–38.1)	19.1 (16.9–20.4)	23.3 (21.8–25.2)	22.5 (20.2–24.7)	16.99 (87)	1.00
Aldolase[c]	421	132	54	0.41	23.8 (21.4–27.8)	23.9 (17.0–27.8)	23.2 (20.4–25.3)	29.1 (25.3–33.8)	55.65 (57)	0.53
Chromosomes	14	12	12	1.00						
Isozymes	21	20	20	0.95						
Morphology	191	184	178	0.97						

[a]Includes the serine, histidine, and leucine sequences combined.
[b]Omits 46 unaligned nucleotides.
[c]Includes the A and C introns combined (incomplete for some taxa).

data base (Cunningham, 1997b, Givnish and Sytsma, 1997; Wiens, 1998b).

Data Partitions and Combinations

Individual parsimony analyses were carried out for almost all sequence partitions and compared for conflict. Only the chromosomes and isoenzymes were not analyzed separately, because of the low total number of informative characters. The strict consensus trees from the ND4 and RNA data share many nodes in common, and many have strong bootstrap support in both (not shown). Because there were no nodes in strong conflict (i.e., mutually incompatible clades supported by high bootstrap values), and because linked mitochondrial genes all share the same history, we combined the mitochondrial data.

Analysis of the aldolase sequences resulted in the hypothesis shown in Figure 3. This analysis recovered 13 equally parsimonious trees—tree length = 322, consistency index (CI) = 0.693, and rescaled consistency index (RI) = RI = 0.524—if *Sceloporus* is constrained to be monophyletic. If the *S. variabilis* group was constrained to a (*S. couchii, S. parvus* (*S. variabilis* + *S. chrysostictus*)) topology, two additional monophyletic groups are recovered that are also strongly supported by the mtDNA partition (Fig. 4, see below). This constraint was necessary because missing data in *S. couchii* and *S. parvus* prevented their placement, and the monophyly of this group is not in question (Wiens and Reeder, 1997). The chromosome and isoenzyme data had few parsimony-informative characters (Table 4), but these were not in strong conflict with each other or with the aldolase sequences (results not shown).

Figure 4 reveals that the topology for all DNA data is identical to that for mtDNA, but with all DNA, the bootstrap and decay index values increase slightly for the ("*Sator*" + *S. utiformis*) clade, as well as for all deep clades. Relevant nodes recovered with various amounts of support for SB/SS taxa include the following: (1) the *S. variabilis* group, and a (*S. chrysostictus* + *S. variabilis*) clade within this group (bootstrap 100%, decay

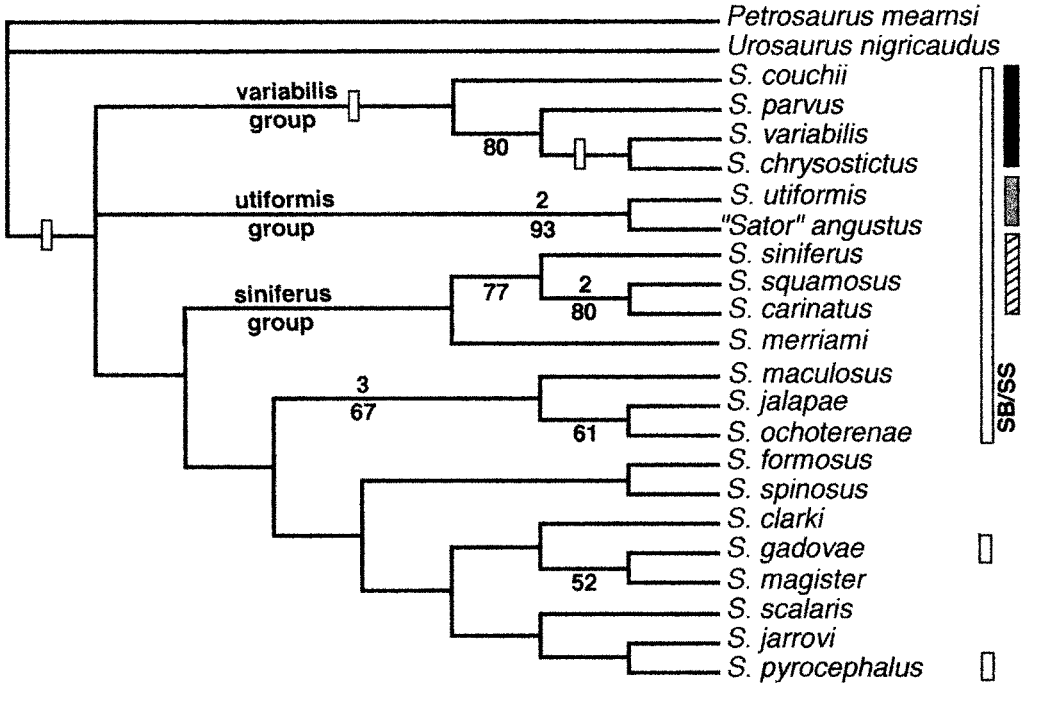

FIGURE 3. Strict consensus of three shortest EP trees constructed from the nuclear aldolase sequences; open bars on the internal branches represent enforced constraints (tree length = 198, CI = 0.808, RI = 0.531). The open vertical bars at the right identify species belonging to the original SB/SS radiation of Smith (1939). See Figures 4 and 5 for explanation of the shaded vertical bars. Decay indexes are shown (above lines) if >1.0, and bootstrap values are given if >50.

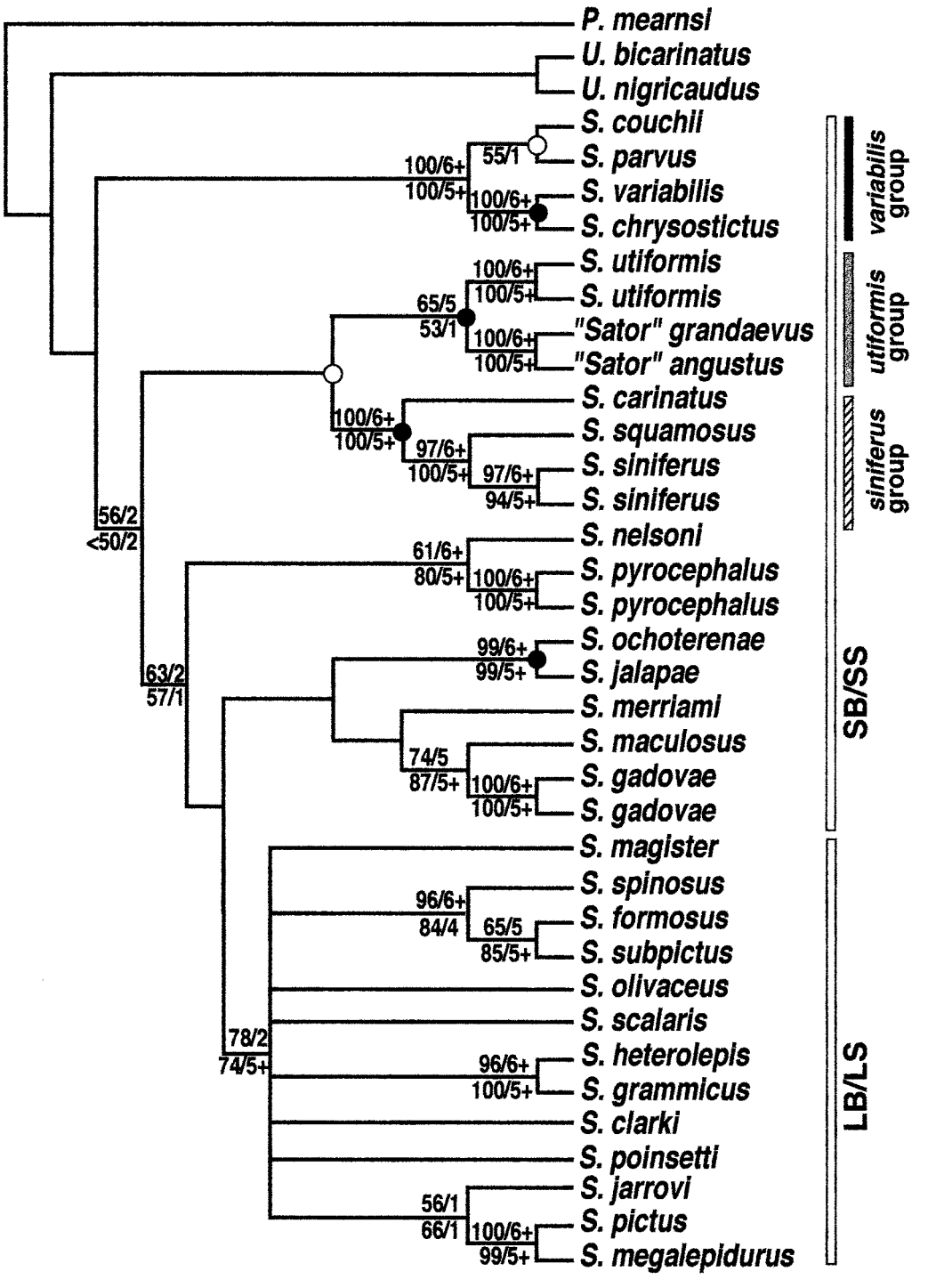

FIGURE 4. Strict consensus of four unconstrained EP trees from the combined mtDNA and nuclear DNA partitions (tree length = 3,950, mtDNA only tree = 3,735, CI = 0.343, RI = 0.379). In this tree, bootstrap proportions are left of the diagonal and decay indexes are to the right (shown only if >50 and >1.0, respectively). Numbers below the internal branches show support based on mtDNA only, and those above are values for total DNA. Circles identify the monophyletic groups recovered from the total nuclear character matrix when the *variabilis* group is constrained to be monophyletic (●) and when it is constrained to a basal position (○) (see Fig. 3).

index 5+ in both mtDNA and total DNA data partitions); (2) a (*S. carinatus* (*S. squamosus* + *S. siniferus*)) clade (bootstrap > 94%, decay index of 5+ in both DNA data partitions); (3) a (*S. utiformis* ("*Sator*" *angustus* + "*Sator*" *grandaevus*)) clade (bootstrap 65% and decay index 5 in the total DNA partition); (4) a (*Siniferus* group + *Utiformis* group) clade (bootstrap < 50%, decay index of 1); (5) a (*S. jalapae* + *S. ochoterenae*) clade (bootstrap > 99% and decay index 5+ in both). The LB/LS clade was also recovered to the exclusion of a paraphyletic SB/SS group (bootstrap > 74% and decay index ≥2) (Fig. 4).

EP Analysis—Combined Data

Figure 5 shows the shortest tree obtained with EP analysis of all available data. The tree recovers a paraphyletic, basal SB/SS radiation, in which "*Sator*" is strongly nested within *Sceloporus*. Further, *S. scalaris* is nested well within the LB/LS radiation (contra Smith, 1939), and other clades recovered include the *variabilis* and *siniferus* groups, a (*S. utiformis* + "*Sator*") group, the *siniferus* and (*S. utiformis* + "*Sator*") groups as sister taxa, a (*nelsoni* + *pyrocephalus*) clade, a (*jalapae* + *ochoterenae*) clade, and the LB/LS clade (Fig. 5).

WP Analyses

The trilevel WP analysis was based on all DNA characters and recovered two shortest trees (Fig. 6a). The basal topology of this tree is ((*variabilis*)(((*siniferus*) + (*utiformis*))(all other *Sceloporus*))); within the (all other *Sceloporus*) clade, a monophyletic LB/LS radiation is recovered (including *S. scalaris*) in one of the two shortest trees. The two shortest trees differ in the placement of *S. magister*; which groups either with the (*S. jalapae* + *S. ochoterenae*) clade or as the basal member of the LB/LS clade.

WP analysis based on the step matrixes (Table 2) generated from likelihood estimates of instantaneous rate changes (Table 3) recovered a single shortest tree (Fig. 6b) in which the relationships among the basal three clades ((*variabilis*)(((*utiformis*) + (*siniferus*))(all other *Sceloporus*))) were congruent with those recovered from the trilevel WP (Fig. 6a). The remaining SB/SS taxa were recovered in more-nested positions (internal to *S. merriami*), as in the EP tree, and the LB/LS radiation was fully resolved (including *S. scalaris*).

ML Analysis

The RNAs and the aldolase genes had high ASRV (as indicated by α values near 0.5), whereas the ND4 gene was relatively more constant, but still had α values < 1.0 (Table 3). We therefore decided that accommodating ASRV was justified in a combined analysis. Differences among classes of transitions and transversions (i.e., T–A transversions were >11 times more frequent than T–G transversions in the RNA sequences; Table 3) and the MODELTEST results led us to select a GTR model. Parameters were estimated from the various trees (nos. 1–6 in Table 3, reading left to right) for the combined data, and after examining the values generated from trees 1–3, 5, and 6, we decided to begin with parameters estimated from the Reeder and Wiens (1996) tree (no. 1 in Table 3), because its values were close to those estimated from our own data (tree no. 3 in Table 3) yet differed from both of the random trees (nos. 5 and 6). We justify this selection with the suspicion that values taken from our own data could harbor some unsuspected circularity, whereas those estimated from another source could not influence our final results in any but a conservative way.

The ML analysis recovered a topology identical to the total DNA tree (Fig. 4), the combined data tree (Fig. 5), and the WP trees (Figs. 6a, b) with regard to the placement of the basal three clades. It was also identical to both WP analyses in placing *S. merriami* as the sister taxon to all remaining *Sceloporus* and a monophyletic LB/LS radiation that included *S. scalaris* (Fig. 7).

Statistical Tests of Alternative Hypotheses

Table 5 summarizes results of Wilcoxon signed-rank tests for the alternative hypotheses presented in Figure 1, relative to the DNA-only (Fig. 4) and combined data (Fig. 5) hypotheses. The sequential Bonferroni-corrected probabilities permit unambiguous rejection of the "monophyly of *Lysoptychus*" hypothesis and the "SS/SB group monophyly" hypothesis in both comparisons, and the DNA-only partition rejects the Wiens and Reeder (1997) placement of the *angustus* group (i.e., both species of "*Sator*", Fig. 1d) basal to the (*siniferus* + *utiformis*) clade ($P = 0.002$). The basal position of *S. merriami* relative to the rest of *Sceloporus*, as proposed by Wiens (1993:fig. 2), is rejected ($P < 0.05$) for the combined data topology,

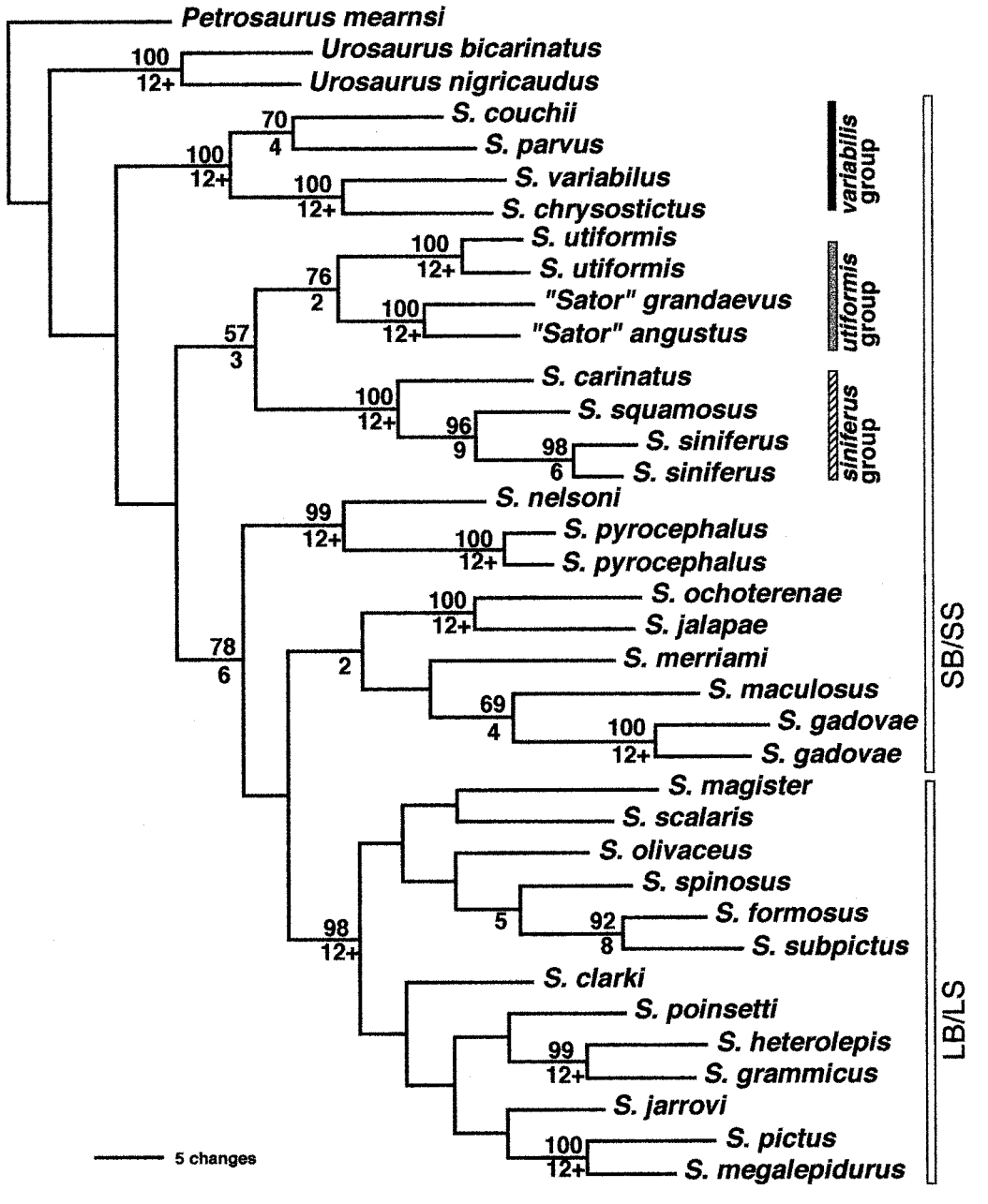

FIGURE 5. Single shortest tree obtained from a EP analysis of all molecular, genetic (chromosomes and isoenzyme), and morphological characters (tree length = 4773.7 [113,945 ÷ 24], CI = 0.3374, RI = 0.3982). Bootstrap values and decay index (1,000 replications, given only if >50% and >1, respectively) are provided above and below internal branches, respectively. Open vertical bars at the right identify all ingroup taxa belonging to the original SB/SS and LB/LS radiations of Smith (1939).

and the "*Sator* out" hypothesis is not rejected by this test in either comparison. Table 6 summarizes the results of the Kishino–Hasegawa (1989) paired likelihood tests. The likelihood tree (Fig. 7) is significantly better than all of the alternative trees tested with sequential Bonferroni-corrected probabilities ($P = 0.05$).

Because of the important biogeographic issues associated with correctly inferring

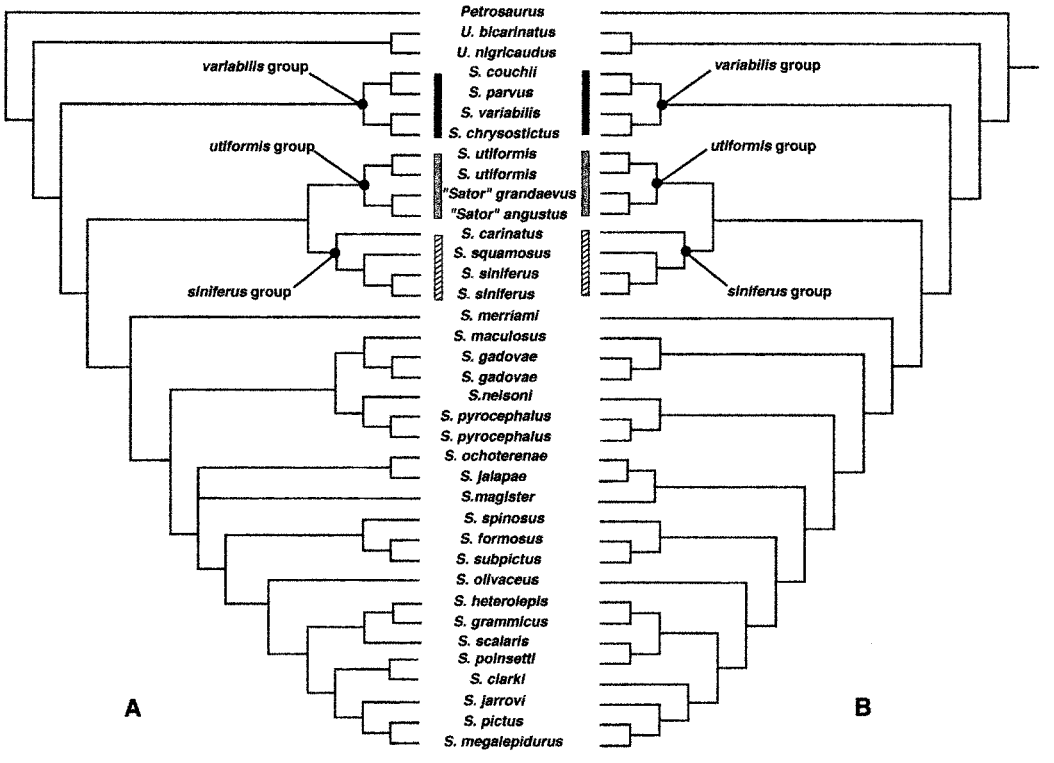

FIGURE 6. Alternative WP analyses of molecular characters. (A) Trilevel-weighted as described in Methods; consensus of two shortest trees, with tree length = 6,453. (B) Frequency step matrix weighting (as given in Table 2). Shaded vertical bars as in Figures 4 and 5.

the phylogenetic position of "*Sator*", we tested our results against the possibility that "*Sator*" is the sister taxon of a monophyletic *Sceloporus* (the "*Sator* out" alternative) by parametric bootstrapping. We constrained a heuristic search of the ND4 data to the "*Sator out*" hypothesis. All equally parsimonious trees generated through the

TABLE 5. Summary of Wilcoxon matched-pairs signed-rank tests (carried out as described by Larson, 1994). The partitions present tree topologies derived from all DNA sequences (Fig. 4) and combined data (Fig. 5), and test these trees against the same set of alternative phylogenetic hypotheses for Sceloporus (Fig. 1).

Constraint tree	Parsimony steps	*n*	Ts	*P*	Rank (*i*)	Sequential Bonferroni[a,b]
DNA only (tree length = 3950; Fig. 4) versus						
"*Sator*"-out	3,966	154	5429.5	0.2809	5	0.2809
SB/SS monophyletic	3,985	134	3430.5	0.0070	3	0.021*
Lysoptychus monophyletic	4,093	181	2393.5	< 0.0001	1	0.0005***
Merriami basal	3,984	151	4669.5	0.0287	4	0.0574
(*siniferus* + *utiformis*) monophyletic	3,986	127	3017.0	0.0005	2	0.002**
Combined data (tree length = 114,608; Fig. 5) versus						
"*Sator*"-out	114,883	86	13062.0	0.3138	5	0.3138
SB/SS monophyletic	115,567	112	9304.5	0.0003	2	0.0012**
Lysoptychus monophyletic	118,456	320	37432.5	< 0.0001	1	0.0005***
Merriami basal	115,436	210	13185.5	0.0132	3	0.0396*
(*siniferus* + *utiformis*) monophyletic	115,137	65	14861.0	0.1059	4	0.2118

[a]The sequential Bonferroni adjustment (Hochberg, 1988; Rice, 1989) corrects each probability by multiplying it by $1 + k - i$, where k is the number of tests ($k = 5$ for each partition), and i is the rank of the tests, given their uncorrected probabilities (from smaller to larger).

[b]* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

TABLE 6. Summary of Kishino–Hasegawa (1989) likelihood variance tests pairing our maximum likelihood tree (Fig. 7) against the alternative topologies illustrated in Figure 1, or as defined below. The probability values are derived from sequential Bonferroni corrections, as described in Table 5.

Tree topology	Parsimony steps	−Ln likelihood	*P*
ML tree (Fig. 7)	3,936	19183.34	
"*Sator*" out	3,966	19242.59	0.05
Wiens and Reeder (1997) combined data	3,986	19249.07	0.0167
(*S. utiformis* (*Siniferus* group))	3,960	19250.64	0.025
(*S. merriami* (all other *Sceloporus*))	3,984	19261.93	0.0125
SB/SS monophyletic	3,985	19317.74	0.01
Wiens and Reeder (1997) morphology	4,061	19367.33	0.0083
Lysoptychus monophyletic	4,093	19533.05	0.0072

10 heuristic-constrained searches of the ND4 data were evaluated by way of likelihood with the PAUP "tree scores" option, and the single best tree and its branch lengths were saved to a file. We chose the ND4 sequences because they had the lowest measure of ASRV (Table 3), because this partition was virtually complete for all taxa, and because as a protein-coding gene, ND4 is less likely to violate the assumptions of the model regarding independence among sites than are the RNA sequences. Adherence to a specified model is particularly important here because in parametric bootstrapping, new DNA matrices are generated through model-dependent simulations.

The tree generated through the heuristic search and its branch lengths were saved to a file. The constrained tree was 10 steps longer than the shortest tree, and we estimated likelihood parameters on this "*Sator* out" tree. We then simulated the evolution of sequence data with the Siminator program (Huelsenbeck et al., 1996a), using the likelihood-estimated parameters of alpha and kappa, and the HKY85 model of the "*Sator* out" tree and its branch lengths, from the previously saved constrained search. In this way, the "known phylogeny" of the simulated data has the topology of ("*Sator*"(*Sceloporus*)), and the branching points in the simulation are based on what they would actually be if this were true. We simulated 100 data sets with 687 nucleotides and then analyzed them with PAUP* (test version 4.0.0d52), using parsimony. Because this was the "known phylogeny" under which the data were simulated, it should be equal to the shortest tree in most cases. A histogram of the distribution of the differences between the tree lengths of our shortest tree minus tree lengths of the ("*Sator*"(*Sceloporus*))–constrained tree is shown in Figure 8. This analysis asks the question: If "*Sator*" is really the sister taxon of a monophyletic *Sceloporus*, how many times in 100 would we observe an "error" of the magnitude we obtained from the actual data (10 steps)? Figure 8 shows that this would occur only once in 100 times, and we reject the "*Sator*-out" hypothesis at $P = 0.01$ with this test.

DISCUSSION

Multiple Data Sets and Character Concordance

Debate on the combinability of different data sets has been extensive (de Queiroz et al., 1995), and a widely held opinion is that combining provides a better chance to recover accurate phylogenies, assuming that data sets meet parsimony assumptions (Chippindale and Wiens, 1994; Wiens and Reeder, 1995; Sullivan, 1996; Cunningham, 1997a,b; Wiens, 1998a,b). This study confirms the advantages of using multiple data sets, even though some taxa are incomplete for some data. We can assess the strength of support for particular clades by evaluating bootstrap and decay values, or the frequency with which any clade is recovered by more than one independent data partition. Figure 7 presents the ML tree and summarizes support for the most consistently recovered positions of the basal clades of *Sceloporus*, in the context of conventional numerical measures of support (bootstrap and decay values) and with congruence among independent data partitions. Conservative measures of congruence require data sets to be independent (Miyamoto and Fitch, 1995a), and here we considered four independent data partitions: combined mitochondrial sequences; aldolase sequences; combined chromosome + isoenzyme loci, and morphology.

None of the individual partitions provided a completely resolved hypothesis, but the EP analysis of all data recovered one fully resolved tree (Fig. 5), and many nodes were supported by multiple independent

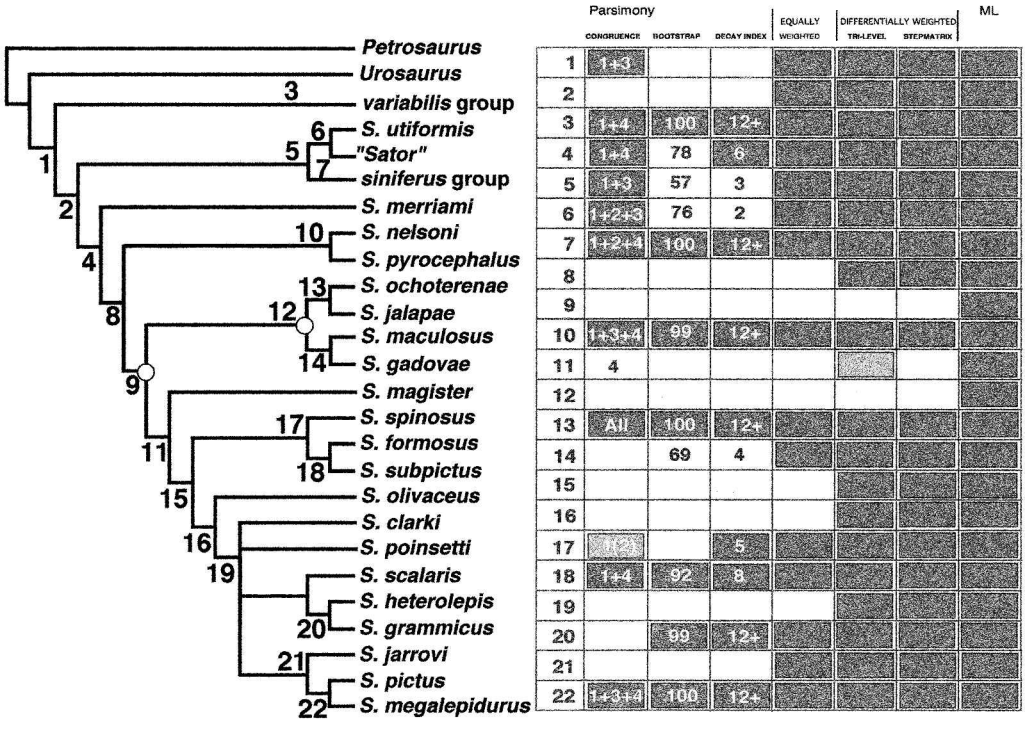

FIGURE 7. Congruence among alternative tree-construction algorithms, and four data partitions, relative to the best-supported topology (the ML tree is shown here). Nodes in the tree correspond to numbers in the left side of the matrix, and the columns in the matrix summarize the following: different data sets recovering the same node (congruence), the bootstrap proportion (if >50%) and decay indexes supporting each node in the tree, and the alternative analysis that recovered (solid box) or failed to recover (open box) a particular node; lightly shaded boxes indicate nodes recovered in some trees for which multiple equally parsimonious solutions were found. Open circles identify nodes with no or very weak support by character congruence or bootstrap (<50%) or decay index values (<1.0). Numbers in the congruence column denote the following data partitions: 1 = mtDNA sequences; 2 = aldolase sequences; 3 = isoenzymes + chromosomes; and 4 = morphological characters.

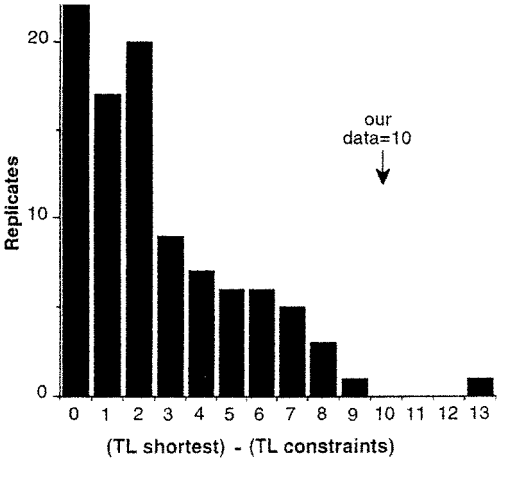

FIGURE 8. Distribution of tree lengths (TL) based on 100 parametric bootstrap replicates of ND4 sequences, under the constraint that "*Sator*" is external to *Sceloporus*. Our topology is significant at $P = 0.01$.

data sets. For example, nodes 1, 3–7, 10, 13, 17, 18, and 22 (Fig. 7) are all supported by at least two data partitions, even though some are only weakly supported by bootstrap and decay indexes (nodes 5 and 6, for example). These are considered strongly supported here because of congruence between mtDNA and chromosome + isoenzyme data (node 5), or between these two and aldolase sequences (node 6).

Sensitivity of Tree Topology to Different Assumptions

The right panel of Figure 7 summarizes the stability of a particular node under EP, WP, and ML assumptions. The WP and ML approaches are different attempts to accommodate rate heterogeneity and probability of homoplasy in the evolution of different

character partitions, and WP methods permit differential character weighting to be assessed in the context of analysis of all partitions combined (Chippindale and Wiens, 1994). ML methods model probabilities of base changes along each tree branch, for each nucleotide in a given set of sequence data, but they are sensitive to the parameters of the specified model (Hasegawa and Fujiwara, 1993; Huelsenbeck and Crandall, 1997) and are currently best suited to analysis of sequence data. The ML parameters used in this study (Table 3) recover a topology that is congruent with those based on EP and WP methods with respect to basal clades (Fig. 7). Our initial expectation was that the step matrix analysis (Fig. 6b) would provide a closer approximation to the ML analysis, because both of these separate substitution frequencies into nine different classes, and both should accommodate nucleotide compositional biases. Yet the trilevel weighting scheme was also strikingly congruent with the ML analysis, surprising concordance considering the difference in their assumptions (trilevel weighting was suspected of being more "coarse" in accommodating rate heterogeneity among molecular partitions). Further, our data set was numerically dominated by molecular characters, some of which were known from the work of Wiens and Reeder (1997) to conflict with morphological data in the placement of some taxa. Under such conditions, EP analyses alone may be misleading for one or the other data partition by long-branch attraction (Felsenstein, 1978), and sequence data may be influenced in a EP analysis by ASRV, compositional bias, or any combination of these factors. Yet agreement among methods suggests that these factors, either singly or in combination, are less likely to be an issue here. This is an important conclusion, because several basal clades (1, 5, and 6) are only weakly supported by conventional measures (bootstrap proportions and decay indexes). We do not claim that congruence among these alternative methods is necessarily evidence for strongly supported nodes but only that the stability of these clades is not sensitive to very different assumptions of character evolution, assumptions designed to correct for problems common to molecular data. Our strongest case for support of most of the clades in Figure 7 is congruence between independent data partitions, and we present this topology as our working hypothesis, with a preference to collapse nodes 9 and 12 (open circles in Fig. 7).

Tests for Alternative Phylogenies

This study confirms the Wiens and Reeder (1997) proposal for many clades, and here we make statistical statements of the strength of our preferred hypothesis (Fig. 7). We reject monophyly of the SS/SB group (Fig. 1a) and monophyly of the *Lysoptychus* group (Fig. 1b) on the basis of the conservative Wilcoxon signed-rank test (Table 5) and confirm the Wiens and Reeder (1997) proposal nesting the *scalaris* group well within the LB/LS radiation (Fig. 1d). The position of *S. merriami* inside of node 4 (Fig. 7) remains uncertain, although we can reject its proposed sister taxon relationship to all other *Sceloporus* (Wiens, 1993) by the Kishino–Hasegawa test (Table 6).

Our hypothesis strongly agrees with the Wiens and Reeder (1997) topology for placement of the two most basal groups formerly recognized as the SB/SS radiation, with one exception: We consistently recover a monophyletic "*Sator*" (their *angustus* group of *Sceloporus*), as the sister clade of *S. utiformis*, and this clade in turn forms a sister group with the *siniferus* group, with a topology of ((*siniferus*) (*S. utiformis* + "*Sator*")) (Figs. 4–7). We reject the Wiens and Reeder topology for this particular clade (Fig. 1d) on the basis of DNA data ($P = 0.002$, Table 5; $P = 0.025$, Table 6), but the nested position of "*Sator*" within *Sceloporus*, as proposed in both studies, confirms a hypothesis previously suggested by many other investigators (Dickerson, 1919; Schmidt, 1922; Wyles and Gorman, 1978; de Queiroz, 1982; Murphy, 1983a; Etheridge and de Queiroz, 1988; Frost and Etheridge, 1989). The recent proposal by Schulte et al. (1998) places "*Sator*" as the sister taxon to a ((*Petrosaurus*) ((*Sceloporus*) + (*Urosaurus*))) clade, on the basis of new mtDNA sequence data (eight tRNAs, ND2, and parts of the ND1 and CO1 protein genes), coupled with morphological characters from Estes et al. (1988), Frost and Etheridge (1989), Reeder and Wiens (1996), and 15 new characters from R. Etheridge (pers. comm.; see Schulte et al., 1998:appendix 1). Our study design does not contain an appropriate sampling of outgroups to evaluate this alternative, and without a description of the new morphological characters, that is difficult to

evaluate. However, we recovered "*Sator*" as the sister taxon to *S. utiformis* in all analyses and on the basis of congruence among three independent data partitions (Fig. 7). Schulte et al. (1998) sampled only a single species of *Sceloporus* (*S. graciosus*, in the LB/LS radiation), and without any species representing basal groups, we suspect that their topology may be biased by sampling artifacts. Further study will need to include more extensive ingroup and outgroup sampling, coupled with the inclusion of nuclear gene sequences.

Most of the groups internal to node 9 in Figure 7, and the position of *S. merriami* with respect to nodes 8–14, need further investigation; because our sampling of the LB/LS radiation species was incomplete, we do not consider it further in this paper. Parts of this group are very poorly resolved in the more extensive phylogeny presented by Wiens and Reeder (1997:figs. 5,8), and will probably require new data for full resolution.

Origin and Biogeography of "Sator" angustus *and* S. grandaevus

Figure 9 shows the distribution of both species of the former "*Sator*" in the Sea of Cortez: *Sceloporus angustus* is endemic to two islands in close proximity (Isla Santa Cruz, Isla San Diego), whereas *S. grandaevus* is known only from Isla Cerralvo, ~150 km south of the range of *S. angustus*. None of these islands have any recent (10,000–15,000 years) connections to the Baja Peninsula; all are thought to have had tectonic origins by block faulting from the eastern edge of the peninsula and are now separated from it by deep water channels (Grismer, 1994a). Grismer (1994a, 1994b) suggested that these two species originated from an ancestor on the Baja Peninsula that has since gone extinct, but neither our phylogenetic hypothesis nor that of Wiens and Reeder (1997) supports this view. All other *Sceloporus* now occurring on the Baja Peninsula and associated islands are nested well within the LB/LS radiation (Wiens and Reeder, 1997) and thus must represent a more recent origin than the two "*Sator*."

An earlier hypothesis proposed that the former genus "*Sator*" originated as a result of a transgulfian vicariance (Murphy, 1975, 1983a, 1983b) associated with the tectonic formation of the peninsula. This original hypothesis has been recently refined by Grismer (1994a, 1994b) on the basis of a better understanding of both the distribution of the peninsular herpetofauna and the geological history of the peninsula itself. He predicts a sister taxon relationship between the species endemic to all or some part of the peninsula and the species confined to a region of western Mexico, which closely approximates the present range of *S. utiformis* (Fig. 9). Grismer (1994a) has also identified several other groups of squamates with similar distributions and calls this group a southern Miocene vicariant complex. The ((*S. angustus* + *S. grandaevus*)(*S. utiformis*)) clade, as a monophyletic group within *Sceloporus*, is consistent with its origin by transgulfian vicariance. In contrast, the Wiens and Reeder (1997) hypothesis (converted to an area cladogram in Fig. 9) recovers *S. utiformis* as the sister group of a clade of taxa distributed further to the south, which is not expected under the transgulfian hypothesis. Our hypothesis and associated area cladogram fits the expectations of the transgulfian vicariance hypothesis, but the Wiens and Reeder phylogeny is also compatible—although it requires an earlier transgulfian event in which the former "*Sator*" diverged first from a more widely distributed ancestor along the west coast of Mexico.

Two caveats to the transgulfian hypothesis must be emphasized. First, a serious concern is that there is currently no definitive support for a vicariant explanation because no well-corroborated phylogenies are available for most of the other groups of the southern Miocene vicariant complex. The single exception is the phylogeny hypothesized by Grismer (1994a) for the *orcutti* group of *Sceloporus*; this group, which is endemic to the Baja Peninsula, was proposed to be the sister clade to the (*nelsoni* + *pyrocephalus*) clade of western Mexico. This is also the relationship predicted by transgulfian vicariance, but Wiens and Reeder (1997:fig. 9) showed that the *orcutti* group is nested well within the LB/LS clade and thus cannot have originated as proposed by Grismer.

Second, an over-water dispersal event cannot be ruled out by either our findings or the Wiens and Reeder (1997) hypothesis. Grismer (1994b:table 2) suggested other examples of this phenomenon on Cerralvo (the snake *Rhinocheilus lecontei*) and Santa Cruz (the rattlesnake *Crotalus atrox*). Such an explanation, however, would require many post hoc assumptions to explain the

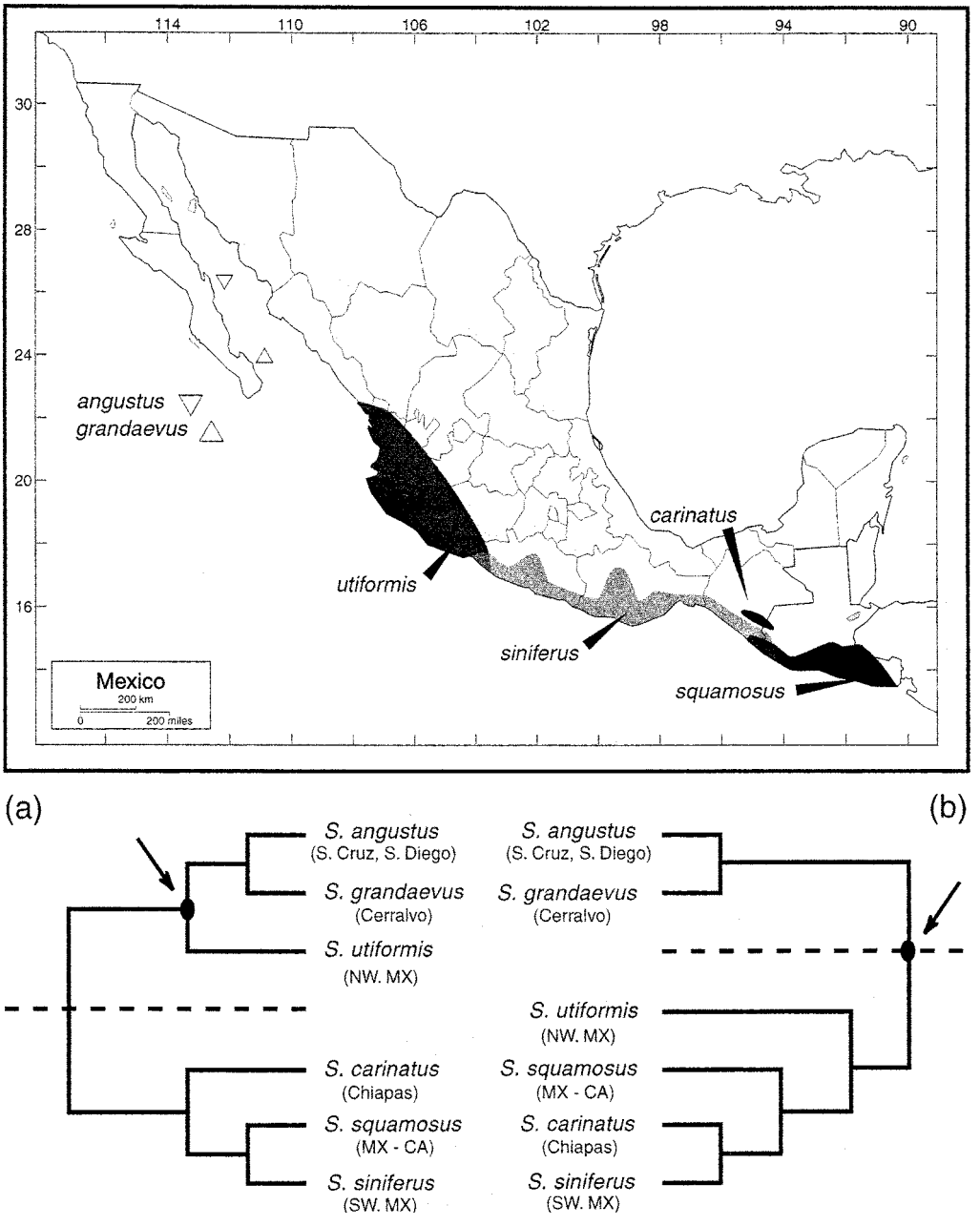

FIGURE 9. Geographic distributions of the *siniferus* and (*S. utiformis* + "*Sator*") groups, and both species of the former "*Sator*". Also shown are alternative area cladograms derived from our phylogenetic hypothesis (a) versus the alternative hypothesis of Wiens and Reeder (1997) (b). The abbreviations S. Cruz, S. Diego, and Cerralvo refer to islands inhabited by *Sceloporus angustus* and *S. grandaevus*; NW.MX and SW.MX refer to northwestern and southwestern Mexico, respectively; MX-CA is extreme southern Mexico and Central America; and Chiapas is the central valley of the state of Chiapas. The dotted line is the first speciation event in each group; and • represents the transgulfian vicariance event postulated in deriving the origin of the former "*Sator*".

current disjunct distribution of *S. angustus* and *S. grandaevus*.

ACKNOWLEDGMENTS

We thank the following persons for their help in various capacities: E. Arellano, D. Busath, D. Cannatella, L. Canseco, K. Crandall, A. Delgadillo, K. de Queiroz, J. Fetzner, D. Frost, F. González, D. Hillis, J. Huelsenbeck, C. Johansson, M. Mancilla, F. Mendoza, S. Muse, A. Nieto Montes de Oca, E. Pérez Ramos, D. Posada, A. Ramos Torres, T. Reeder, A. Rendón, W. Schmidt, J. Wiens, and A. Zaldívar. T. Reeder and J. Wiens kindly sent us the PAUP file of their 1997 morphological and molecular data, and K. Crandall, R. Blahnik, and Z. Swigonova made comments on an earlier drafts. This research was supported by National Science Foundation (NSF) grant DEB 91-19091 to J.W.S. and O.F.V. and by NSF grants DEB 96-32879 and DEB 99-74036, and the New Jersey Agricultural Experiment Station, to K.M.K. Also O.F-.V. and M.B. thank DGAPA, Facultad de Ciencias; Instituto de Ecología, UNAM, and the Department of Zoology, BYU, for support of their leaves.

REFERENCES

AGOSTI, D., D. JACOBS, AND R. DESALLE. 1996. On combining protein sequences and nucleic acid sequences in phylogenetic analysis: the homeobox protein case. Cladistics 12:65–82.

ARÉVALO, E., S. K. DAVIS, AND J. W. SITES, JR. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Syst. Biol. 43:387–418.

BALLARD, J. W. O., M. K. THAYER, A. F. NEWTON, JR., AND E. R. GRISMER. 1998. Data sets, partitions, and characters: philosophies and procedures for analyzing multiple data sets. Syst. Biol. 47:367–396.

BARBADILLA, A., L. M. KING, AND R. C. LEWONTIN. 1996. What does electrophoretic variation tell us about protein variation? Mol. Biol. Evol. 13:427–432.

BARRETT, M., M. J. DONOGHUE, AND E. SOBER. 1991. Against consensus. Syst. Zool. 40:486–493.

BENABIB, M., K. M. KJER, AND J. W. SITES, JR. 1997. Mitochondrial DNA sequence-based phylogeny and the evolution of viviparity in the *Sceloporus scalaris* group (Reptilia, Squamata). Evolution 51:1262–1275.

BOROWIK, O. A. 1995. Coding chromosomal data for phylogenetic analysis: phylogenetic resolution of the *Pan-Homo-Gorilla* trichotomy. Syst. Biol. 44:563–570.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803.

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. 42:384–397.

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

COLLINS, T. M., P. H. WIMBERGER, AND J. P. NAYLOR. 1994. Compositional bias, character–state bias, and character–state reconstruction using parsimony. Syst. Biol. 43:482–496.

COYNE, J. 1982. Gel electrophoresis and cryptic protein variation. *in* Isozymes Curr. Top. Biol. Med. Res. 6:1–32.

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

CUNNINGHAM, C. W. 1997b. Is congruence between data partitions a reliable predictor of phylogenetic accurancy? Syst. Biol. 46:464–478.

DE QUEIROZ, K. 1982. The scleral ossicles of sceloporine iguanids: a reexamination with comments on their phylogenetic significance. Herpetologica 38:302–311.

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

DICKERSON, M. C. 1919. Diagnosis of twenty-three new species and a new genus of lizard from Lower California. Bull. Am. Mus. Nat. Hist. 61:461–477.

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

ESTES, R., K. DE QUEIROZ, AND J. GAUTHIER. 1988. Phylogenetic relationships within squamata. Pages 119–281 *in* Phylogenetic relationships of the lizard families (R. Estes and G. Pregill, eds.). Stanford Univ. Press, Stanford, California.

ETHERIDGE, R. E., AND K. DE QUEIROZ. 1988. A phylogeny of Iguanidae. Pages 283–368 *in* Phylogenetic relationships of the lizard families (R. Estes and G. Pregill, eds.). Stanford Univ. Press, Stanford, California.

FARRIS, J. S. 1969. A succesive approximations approach to character weighting. Syst. Zool. 18:374–385.

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

FELSENSTEIN, J. 1981a. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17:368–376.

FELSENSTEIN, J. 1981b. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. Biol. J. Linn. Soc. 16:183–196.

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

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

FROST, D. R., AND R. M. TIM. 1992. Phylogeny of plecotine bats (Chiroptera: "Vespertilionidae"): summary of the evidence and proposal for a logically consistent taxonomy. Am. Mus. Novit. 3034:1–16.

GIVNISH, T. J., AND K. J. SYTSMA. 1997. Consistency, characters, and the likelihood of correct phylogenetic inference. Mol. Phylogenet. Evol. 7:320–330.

GRISMER, L. L. 1994a. The origin and evolution of the peninsular herpetofauna of Baja California, Mexico. Herpetol. Nat. Hist. 2:51–106.

GRISMER, L. L. 1994b. Geographic origins for the reptiles on islands in the Gulf of California, Mexico. Herpetol. Nat. Hist. 2:17–40.

GUTELL, R. R. 1994. Collection of small subunit (16S and 16S–like) ribosomal RNA structures: 1994. Nucleic Acids Res. 22:3502–3507.

GUTELL, R. R., N. LARSEN, AND C. R. WOESE. 1994. Lessons from an evolving rRNA: 16S and 23S rRNA

structures from a comparative perspective. Microbiol. Rev. 58:10–26.

HALL, W. P. 1973. Comparative population cytogenetics, speciation, and evolution of the crevice-using species of *Sceloporus*. Ph.D. Thesis, Harvard Univ., Cambridge, Massachusetts.

HALL, W. P. 1980. Chromosomes, speciation, and evolution of Mexican iguanid lizards. Natl. Geogr. Soc. Res. Rep. 12:309–329.

HALL, W. P. 1983. Modes of speciation and evolution in sceloporine iguanid lizards. I. Epistemology of the comparative approach and introduction to the problem. Pages 643–679 *in* Advances in herpetology and evolutionary biology (A. G. J. Rhodin and K. Miyata, eds.). Museum of Comparative Zoology, Harvard Univ., Cambridge, Massachusetts.

HASEGAWA, M., AND M. FUJIWARA. 1993. Relative efficiencies of the maximum likelihood, maximum parsomony, and neighbor–joining methods for estimating protein phylogeny. Mol. Phylogenet. Evol. 2:1–5.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating the human–ape split by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21:160–174.

HÅSTAD, O., AND M. BJÖRKLUND. 1998. Nucleotide substitution models and estimation of phylogeny. Mol. Biol. Evol. 15:1381–1389.

HILLIS, D. M., AND S. K. DAVIS. 1986. Evolution of ribosomal DNA: fifty million years of recorded history in the frog genus *Rana*. Evolution 40:1275–1288.

HILLIS, D. M., J. J. BULL, M. E. WHITE, M. R. BADGETT, AND I. J. MOLINEUX. 1992. Experimental phylogenetics: generation of a known phylogeny. Science 255:589–592.

HILLIS, D. M., J. P. HUELSENBECK, AND C. W. CUNNINGHAM. 1994a. Applications and accuracy of molecular phylogenies. Science 264:671–677.

HILLIS, D. M., J. P. HUELSENBECK, AND D. L. SWOFFORD. 1994b. Hobglobin of phylogenetics? Nature 36:363–364.

HOCHBERG, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75:800–802.

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

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., AND D. M. HILLIS. 1993. Success of phylogenetic methods in the four–taxon case. Syst. Biol. 42:247–264.

HUELSENBECK, J. P., AND B. RANNALA. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276:227–232.

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

HUELSENBECK, J. P., D. M. HILLIS, AND R. NIELSEN. 1996b. A likelihood–ratio test of monophyly. Syst. Biol. 45:546–558.

INTERNATIONAL UNION OF BIOCHEMISTRY NOMENCLATURE COMMITTEE. 1984. Enzyme nomenclature, 1984. Academic Press, Orlando, Florida.

JUKES, T. H., AND C. R. CANTOR. 1969. Evolution of protein molecular. Pages 21–123 *in* Mammalian protein metabolism (H. N. Munro, ed.). Academic Press, New York.

KIM, J. 1993. Improving the accuracy of phylogenetic estimation by combining different methods. Syst. Biol. 42:331–340.

KISHINO, H., AND M. HASEGAWA. 1989. Evaluation of maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29:170–179.

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

KJER, K. M. 1997. Conserved primary and secondary structural motifs of amphibian 12S rRNA, domain III. J. Herpetol. 31:599–604.

KJER, K. M., G. D. BALDRIDGE, AND A. M. FALLON. 1994. Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure. Biochim. Biophys. Acta 1217:147–155.

KUMAR, S., K. TAMURA, AND M. NEI. 1993. MEGA: molecular evolutionary genetics analysis, ver. 1.01. Pennsylvania State Univ., University Park, Pennsylvania.

KUMAZAWA, Y., AND M. NISHIDA. 1993. Sequence evolution of mitochondrial tRNA genes and deep–branch animal phylogenetics. J. Mol. Evol. 37:380–398.

LARSEN, K. R., AND W. W. TANNER. 1974. Numeric analysis of the lizard genus *Sceloporus* with special reference to cranial osteology. Great Basin Nat. 34:1–41.

LARSEN, K. R., AND W. W. TANNER. 1975. Evolution of the sceloporine lizards (Iguanidae). Great Basin Nat. 35:1–20.

LARSEN, N. 1992. Higher order interactions of 23S rRNA. Proc. Natl. Acad. Sci. USA 89:5044–5048.

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

LESSA, E. P., AND G. APPLEBAUM. 1993. Screening techniques for detecting allelic variation in DNA sequences. Mol. Ecol. 2:119–129.

MABEE, P. M., AND J. HUMPHRIES. 1993. Coding polymorphic data: examples from allozymes and ontogeny. Syst. Biol. 42:166–181.

MINK, D. G., AND J. W. SITES, JR. 1996. Species limits, phylogenetic relationships, and origins of viviparity in the *Scalaris* complex of the lizard genus *Sceloporus* (Phrynosomatidae: Sauria). Herpetologica 52:551–571.

MIYAMOTO, M. M., AND W. M. FITCH. 1995a. Testing species phylogenies and phylogenetic methods with congruence. Syst. Biol. 44:64–76.

MIYAMOTO, M. M., AND W. M. FITCH. 1995b. Testing the covarion hypothesis of molecular evolution. Mol. Biol. Evol. 12:503–513.

MORITZ, C., C. J. SCHNEIDER, AND D. B. WAKE. 1992. Evolutionary relationships within the *Ensatina* eschscholtzii complex confirm the ring species interpretation. Syst. Biol. 41:273–291.

MURPHY, R. W. 1975. Two new blind snakes (Serpentes: Leptotyphlopidae) from Baja California, Mexico with a contribution to the biogeography of peninsular and insular herpetofauna. Proc. Calif. Acad. Sci. 40:87–92.

MURPHY, R. W. 1983a. Paleobiogeography and genetic differentiaton of the Baja California Herpetofauna. Occ. Pap. Calif. Acad. Sci. 137:1–48.

MURPHY, R. W. 1983b. The reptiles: origins and evolution. Pages 130–158 *in* Island biogeography in the Sea of Cortez (T. J. Case and M. L. Cody, eds.). Univ. California Press, Los Angeles.

MURPHY, R. W., J. W. SITES, JR., D. G. BUTH, AND C. H. HAUFLER. 1996. Proteins: isozyme electrophoresis. Pages 51–120 *in* Molecular systematics. (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.

PORTER, C. A., M. J. HAMILTON, J. W. SITES, JR., AND R. J. BAKER. 1991. Location of ribosomal DNA in chromosomes of squamate reptiles: systematic and evolutionary implications. Herpetologica 47:271–280.

POSADA, D., AND K. A. CRANDALL. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.

REEDER, T. W. 1995. Phylogenetic relationships among phrynosomatid lizards as inferred from mitochondrial ribosomal DNA sequences: substitutional bias and information content of transitions relative to transversions. Mol. Phylogenet. Evol. 4:203–222.

REEDER, T. W., AND J. J. WIENS. 1996. Evolution of the lizard family Phrynosomatidae as inferred from diverse types of data. Herpetol. Monogr. 10:43–84.

RICE, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

RODRÍGUEZ, F. J., J. L. OLIVER, A. MARÍN, AND J. R. MEDINA. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142:485–501.

SCHMIDT, K. P. 1922. The amphibians and reptiles of Lower California and neighboring islands. Bull. Am. Mus. Nat. Hist. 46:607–707.

SCHULTE, J. A., II, J. R. MACEY, A. LARSON, AND T. J. PAPENFUSS. 1998. Molecular tests of phylogenetic taxonomies: a general procedure and example using four subfamilies of the lizard family Iguanidae. Mol. Phylogenet. Evol. 10:367–376.

SIDDALL, M. E., AND A. G. KLUGE. 1997. Probabilism and phylogenetic inference. Cladistics 13:313–336.

SITES, J. W., JR., J. W. ARCHIE, C. J. COLE, AND O. FLORES-VILLELA. 1992. A review of phylogenetic hypotheses for lizards of the genus *Sceloporus* (Phrynosomatidae): implications for ecological and evolutionary studies. Bull. Am. Mus. Nat. Hist. 213:1–110.

SMITH, H. B. 1939. The Mexican and Central American lizards of the genus Sceloporus. Zool. Ser. Field Mus. Nat. Hist. 26:1–397.

SULLIVAN, J. 1996. Combining data with different distributions of among–site rate variation. Syst. Biol. 45:375–380.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, ver. 3.1. Illinois Natural History Survey, Champaign, Illinois.

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

VAN DE PEER, J., M. NEEFS, P. DE RIJK, AND R. DE WACHTER. 1993. Reconstructing evolution from eukaryotic small-subunit RNA sequences: calibration of the molecular clock. J. Mol. Evol. 37:221–232.

VOELKER, G., AND S. V. EDWARDS. 1998. Can weighting improve bush trees? Models of cytochrome b evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). Syst. Biol. 47:589–603.

WIENS, J. J. 1993. Phylogenetic relationships of phrynosomatid lizards and monophyly of the *Sceloporus* group. Copeia 1993:287–299.

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

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

WIENS, J. J. 1998b. 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 spiny lizards (*Sceloporus*) based on molecular and morphological evidence. Herpetol. Monogr. 11:1–101.

WYLES, J. S., AND G. C. GORMAN. 1978. Close relationship between the lizard genus *Sator* and *Sceloporus utiformis* (Reptilia, Lacertilia, Iguanidae): electrophoretic and immunological evidence. J. Herpetol. 12:343–350.

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 12 January 1999; accepted 19 July 1999
Associate Editor: R. Olmstead

APPENDIX

Voucher specimens, museum accession numbers, and localities for all lizards used in this study. Asterisks identify individuals from which both DNA sequence and isoenzyme data were collected.

Taxon	Voucher no.	Locality
DNA		
*Sceloporus carinatus**	MZFC-6508	Buena Vista, 49.5 km S Tuxtla Gutierrez, Villa Flores, Chiapas
S. chrysostictus	MZFC-5512	8 km N Pisté, Yucatán
S. clarki	MZFC-5757	Fatima, 9 km NE Guaymas on Hwy 15, Sonora
S. couchii	MZFC-6676	Cañón de la Huasteca, Sta. Catarina, Nuevo León
S. formosus	IBH-7179	Sierra de Igualatlaco, 0.5–2.0 km W Omiltemi, Chilpancingo, Guerrero
*S. gadoviae**	MZFC-5701	10 km S Mezcala, Guerrero
S. gadoviae	MZFC-5715	Venta Salada, Tehuacán, Puebla
S. grammicus	BYU-38487	San Miguel Ajusco, D.F.
*S. heterolepis**	MZFC-6177	40 km NE Tamazula, Jalisco

APPENDIX. Continued.

Taxon	Voucher no.	Locality
S. jalapae	MZFC-5958	Venta Salada, Rail road Station, Coxatlán, Puebla
S. jarrovi	MZFC-6189	19.7 km E Revolcadero, Hwy Salto-Mazatlán, Durango
*S. maculosus**	MZFC-6783	Francisco Zarco Dam, Durango
S. magister	MZFC-5752	Empalme, 19 km NE Guaymas, Hwy 15, Guaymas, Sonora
S. megalepidurus	MZFC-6423	La Soledad, Tlaxcala
S. merriami	MZFC-6678	11 km S Boquillas del Carmen, Coahuila
*S. nelsoni**	MZFC-6198	12.8 km S of Culiacán, Hwy 15, Sinaloa
S. ochoterenae	MZFC-5703	12 km W of Chilpancingo, Guerrero
S. olivaceus	BYU-42888	Concho Co., Texas, USA
S. parvus	MZFC-5344	10 km NE San Antonio Peña Nevada, Nuevo León
*S. pictus**	MZFC-6425	8 km NW Chapulco, Puebla
S. poinsetti	BYU-42534	Catron Co., New Mexico, USA
S. pyrocephalus	MZFC-6098	Minatitlán, Colima
*S. pyrocephalus**	MZFC-6045	10 km N Nueva Italia, Bridge El Marquéz, Michoacán
S. scalaris	BYU-45474	Carr Peak, Huachuca Mts., Cochise C., Arizona, USA
S. siniferus	MZFC-8531	Sta. Ana del Progreso, Putla, Oaxaca
*S. siniferus**	MZFC-5711	4 km W of Tierra Colorada, Palo Gordo, Guerrero
S. squamosus	MZFC-5441	5 km N of Pijijiapan, "Torre Microondas", Chiapas
S. spinosus	MZFC-6079	10 km E Tehuacán, Sta. Ana, Puebla
*S. subpictus**	MZFC-8552	~115 km from turnoff to Morelos on highway to Tlaxiaco, San Esteban Atlatahuaca, Oaxaca
S. utiformis	MZFC-6091	Boca Iguanas, Hwy Chamela-Barra de Navidad, Jalisco
*S. utiformis**	MZFC-5807	6-7 km E Uruapan, Hwy Pátzcuaro-Uruapan, Michoacán
S. variabilis	MZFC-6839	S side Salado Park, Salado Cr. San Antonio, Bexar Co., Texas, USA
"Sator" angustus	MZFC-6569	Sta. Cruz Island, Gulf of California, Baja California Sur
"Sator" grandaevus	MZFC-6572	Cerralvo Island, Gulf of California, Baja California Sur
Petrosaurus mearnsi	MZFC-6571	Arroud Hotel "La Pinta", Cataviña, Baja California
Urosaurus bicarinatus	MZFC-5700	2 km W Tierra Colorada, Guerrero
Urosaurus nigricaudus	MZFC-6594	San Bartolo, Baja California Sur
ENZYMES		
S. carinatus	MZFC-6508	Buena Vista, 49.5 km S Tuxtla Gutiérrez, Villa Flores, Chiapas
S. chrysostictus	MZFC-6628	5 km ESE Pisté, Yucatán
S. chrysostictus	MZFC-6628	5 km ESE Pisté, Yucatán
S. couchii	MZFC-6675	Cañón de la Huasteca, near Santa Catarina, Nuevo León
S. couchii	MZFC-6675	Cañón de la Huasteca, near Santa Catarina, Nuevo León
S. gadoviae	MZFC-5694	18 km SE Taxco, Guerrero
S. gadoviae	MZFC-5694	18 km SE Taxco, Guerrero
S. gadoviae	MZFC-5953	2 km W Venta Salada, Coxatlán, Puebla
S. gadoviae	MZFC-8542	Zoquiapan, Boca de RÍos, Cuicatlán, Oaxaca
S. gadoviae	MZFC-8543	1 km S Quiotepec, Cuicatlán, Oaxaca
S. gadoviae	MZFC-5953	2 km W Venta Salada, Coxatlán, Puebla
S. gadoviae	MZFC-5701	10 km S Mezcala, Guerrero
S. heterolepis	MZFC-6765	km 19, Hwy Tequila a Volcán Tequila, Est. Microondas, Tequila, Jalisco
S. jalapae	MZFC-5956	Morelos between Tierra Blanca and Tamazulapa, Oaxaca
S. jalapae	MZFC-5957	Morelos between Tierra Blanca and Tamazulapa, Oaxaca
S. jalapae	MZFC-5950	6 km E Tamazulapan, Oaxaca
S. jalapae	MZFC-5946	2 km W Pilas, km 141, Est. Microondas, Oaxaca
S. jalapae	MZFC-5743	Venta Salada, Tehuacán, Puebla
S. jalapae	MZFC-5743	Venta Salada, Tehuacán, Puebla
S. jalapae	MZFC-8545	Cieneguilla, Etla, Oaxaca
S. jalapae	MZFC-8546	Cieneguilla, Etla, Oaxaca
S. jalapae	MZFC-8547	El Venado, Cuicatlán, Oaxaca
S. jalapae	MZFC-8544	El Venado, Cuicatlán, Oaxaca
S. merriami	MZFC-6678	11 km S Boquillas del Carmen, Coahuila
S. merriami	MZFC-6677	11 km S Boquillas del Carmen, Coahuila
S. nelsoni	BYU-45524	20.5 km S Culiacán, Hwy 15, Sinaloa
S. ochoterenae	MZFC-5704	3 km SE Ixcateopan de Cuahutemoc, Guerrero

APPENDIX. Continued.

Taxon	Voucher no.	Locality
S. ochoterenae	MZFC-6845	8 km from turnoff to Izúcar de Matamoros, Epatlán, Puebla
S. ochoterenae	MZFC-5703	12 km W Chilpancingo, Guerrero
S. parvus	MZFC-6663	Cuesta de Malena, Coahuila
S. parvus	BYU-45768	Cuesta de Malena, Coahuila
S. parvus	MZFC-6060	10 km S Tulantongo, Hidalgo
S. parvus	MZFC-5738-2	San Francisco Tecajique, Hidalgo
S. pictus	MZFC-5959	8 km E Chapulco, Puebla
S. pictus	MZFC-6425	8 km E Chapulco, Puebla
S. pyrocephalus	MZFC-6098	9 km W Villa de Alvarez, Colima
S. pyrocephalus	MZFC-6098	Minatitlán, Colima
S. pyrocephalus	MZFC-6098	Arroyo Seco, km 7.5, Hwy Colima-Minatitlán, Colima
S. pyrocephalus	MZFC-6045	10 km N Nueva Italia, Bridge El Marquéz, Michoacán
S. pyrocephalus	MZFC-6045	10.0 km N Nueva Italia, Bridge El Marquéz, Michoacán
S. scalaris	BYU-45475	Carr Peak, Huachuca Mts., Cochise Co., Arizona, USA
S. scalaris	BYU-45478	Chiricahua Mts., Rustles, Cochise Co., Arizona, USA
S. scalaris	BYU-45480	Appleton-White Res. Ranch, Sta. Cruz, Arisona, USA
S. siniferus	MZFC-5963	5 km NW El Camaron, Oaxaca
S. siniferus	MZFC-6500	29 km E Miahuatlán, Oaxaca
S. siniferus	MZFC-6503	7 km N Zanatepec, Oaxaca
S. siniferus	MZFC-6496	Cerro Acaltepec, SW Sta. Maria Ecatepec, Oaxaca
S. siniferus	MZFC-5710	Agua de Obispo, Guerrero
S. siniferus	MZFC-8532	Guadalupe Zacapepec, Putla, Oaxaca
S. siniferus	MZFC-8533	San Isidro Palizada, Putla, Oaxaca
S. siniferus	MZFC-8534	Cerro Piedra Larga, Nejapa, Oaxaca
S. siniferus	MZFC-8535	Río Las Peñas, San Pedro Simiyuri, Putla, Oaxaca
S. siniferus	MZFC-8536	13 km S El Amate, Sta. Ana del Progreso, Putla, Oaxaca
S. squamosus	MZFC-5440	1 km NE towards El Cerezo, from El Triunfo, Chiapas
S. squamosus	MZFC-5439	7 km W El Triunfo, Chiapas
S. spinosus	MZFC-5734	San Francisco Tecajique, Hidalgo
S. subpictus	Missing	~115 km from turn off to Morelos on Hwy to Tlaxiaco, San Esteban Atlatahuaca, Oaxaca
S. subpictus	MZFC-8548	~115 km from turnoff to Morelos on Hwy to Tlaxiaco, San Esteban Atlatahuaca, Oaxaca
S. utiformis	MZFC-6796	1 km W Tepeltictic, Nayarit
S. utiformis	MZFC-5807	6–7 km E Uruapan, Hwy. Pátzcuaro-Uruapan, Michoacán
S. utiformis	MZFC-6092	Turnoff to Boca Iguanas, Hwy Chamela-Barra de Navidad, Jalisco
S. utiformis	MZFC-6093	Bridge Rõo Armerias, Colima
S. utiformis	MZFC-6091	Turnoff to Boca Iguanas, Hwy Chamela-Barra de Navidad, Jalisco
S. utiformis	MZFC-7209	La Calera, Casimiro Castillo, Jalisco
S. variabilis	MZFC-6839	S side Salado Park, Salado Cr. San Antonio, Bexar Co., Texas, USA
S. variabilis	MZFC-6839	S side Salado Park, Salado Cr. San Antonio, Bexar Co., Texas, USA
"Sator" angustus	MZFC-6569	Sta. Cruz Island, Baja California Sur
"Sator" angustus	MZFC-6569	Sta. Cruz Island, Baja California Sur
"Sator" grandaevus	MZFC-6572	Cerralvo Island, Baja California Sur
"Sator" grandaevus	MZFC-6572	Cerralvo Island, Baja California Sur
Urosaurus bicarinatus	MZFC-8550	Cerro Piedra Larga, Nejapa de Madero, Oaxaca