

# Large Subunit Mitochondrial rRNA Secondary Structures and Site-Specific Rate Variation in Two Lizard Lineages

Richard P. Brown

School of Biological & Earth Sciences, Liverpool John Moores University, Liverpool L3 3AF, UK

Received: 16 April 2004 / Accepted: 29 July 2004 [Reviewing Editor: Dr. Yves Van de Peer]

**Abstract.** A phylogenetic-comparative approach was used to assess and refine existing secondary structure models for a frequently studied region of the mitochondrial encoded large subunit (16S) rRNA in two large lizard lineages within the Scincomorpha, namely the Scincidae and the Lacertidae. Potential pairings and mutual information were analyzed to identify site interactions present within each lineage and provide consensus secondary structures. Many of the interactions proposed by previous models were supported, but several refinements were possible. The consensus structures allowed a detailed analysis of rRNA sequence evolution. Phylogenetic trees were inferred from Bayesian analyses of all sites, and the topologies used for maximum likelihood estimation of sequence evolution parameters. Assigning gamma-distributed relative rate categories to all interacting sites that were homologous between lineages revealed substantial differences between helices. In both lineages, sites within helix G2 were mostly conserved, while those within helix E18 evolved rapidly. Clear evidence of substantial site-specific rate variation (covarion-like evolution) was also detected, although this was not strongly associated with specific helices. This study, in conjunction with comparable findings on different, higher-level taxa, supports the ubiquitous nature of site-specific rate variation in this gene and justifies the incorporation of covarion models in phylogenetic inference.

**Key words:** 16S rRNA — Bayesian inference — Covarion — Helix — Lizard — Phylogeny — Skink — Secondary structure

## Introduction

The importance of incorporating secondary structure information in phylogenetic studies of rRNA genes is well established (e.g., Dixon and Hillis 1993), with the need for accurate models increased now that refined models of DNA evolution that incorporate base covariation are available (Huelsenbeck and Ronquist 2001). Inferring secondary structures from sequence data is generally undertaken by either phylogenetic (comparative) studies that search for covarying bases or those that minimize the free energies of putative secondary structures (Zuker 1989), although there are a growing number of hybrid techniques (e.g., Juan and Wilson 1999). One issue with the former is that analyses are generally carried out on quite divergent taxa in order for a reasonable degree of covariation to be observed. This has allowed prediction of the major structures, but identification of all base pairs within those structures may leave some scope for improvement. This can be partly attributed to (1) highly accurate alignments becoming more difficult with increased divergence between taxa and (2) more divergent taxa being more likely to have divergent secondary structures, reducing the likelihood that a comprehensive consensus model (i.e., one containing interacting sites common to most/all taxa in the study) can be obtained.

A phylogenetic comparative study of closely related taxa is unlikely to yield a good secondary structure model through insufficient base covariation (but see Parsch et al. 2000). Fortunately, a large volume of research means that the main structural elements are well-established for some RNA subunits in a wide variety of taxa (Wuyts et al. 2001a). Thus highly refined models applicable to a group of closely related organisms in a phylogenetic study can be obtained by fine-scale analysis of the known structural elements, using information on base complementarities as well as covariation if present (e.g., Espinosa de los Monteros 2003). This approach is feasible for studies of large (LSU; also known as 16S) and small (SSU; also known as 12S) subunit mitochondrial rRNA thanks to the quantity of sequence data now available in genetic databanks.

Over recent years there have been many evolutionary studies of two families within the scincomorph lizard clade that are traditionally named the Lacertidae and the Scincidae. Interest has centered on the systematics of these groups and the interesting phylogeographic patterns that they show (e.g., Fu 1998; Harris et al. 1998a; Honda et al. 2000; Mausfeld et al. 2000). The former is found in Africa and most of Eurasia, while the latter is found on all continents except Antarctica. Secondary structures for the LSU mitochondrial rRNA have been proposed for 14 species of the *Darevskia* genus within the Lacertidae and for *Eumeces egregius* within the Scincidae and are available on a WWW database (see Wuyts et al. 2001a), but how representative these are of other taxa within these groups is unknown. These models are hereafter referred to as the “W&VP” models, with the *Eumeces* model specified as /E and the *Darevskia* models (which are all almost identical) denoted by /D. The main structures are common to both models, with differences restricted to the actual pattern of site interactions that form these elements (note that the term “site interaction” is used here to refer to the pairing between bases at specific positions to form secondary structures). The first objective of this work is to evaluate, compare, and refine these interactions to provide consensus models for each lineage.

Obtaining refined rRNA consensus structures allows detailed and accurate analyses of rRNA evolution, which increases the understanding of functional aspects of rRNA molecules. It also aids alignment of rRNA sequences and contributes to the increasingly sophisticated models of sequence evolution being applied in maximum likelihood and Bayesian approaches. Two of these are of particular interest here: (1) doublet models that take into account the dependent substitutions between interacting sites within an rRNA secondary structure (Savill et al. 2001) and (2) covarion models that allow variation in site-specific rates among lineages, generally known as site-specific

rate variation (SSRV) or covarion-like evolution (Galtier 2001). The former requires application of a good secondary structure model, which is achieved by the first part of this work described above. There has been considerable recent interest in finding empirical evidence of the latter and studies have been carried out that involve diverse groups of taxa (Galtier 2001; Misof et al. 2002). Additional objectives of this work were examination of rate variation between sites and testing for SSRV in the mitochondrial LSU rRNA of the Scincidae and the Lacertidae.

## Materials and Methods

### *Alignment and Assessment of Secondary Structure*

Helix names follow the Wuyts et al. (2001a) nomenclature unless stated otherwise. A total of 99 Lacertidae and 126 Scincidae 16S rRNA sequences were obtained from GenBank (Appendix I). All sequences contained a homologous fragment of the 16S rRNA gene, extending from the terminal two bases of helix E23' (helix 66' in the Gutell model [Gutell et al. 1985; Gutell and Fox 1988; <http://www.rna.icmb.utexas.edu/>]) to the first base of G20' (helix 90' in the Gutell model). This is equivalent to 499 bases in the Scincid *Eumeces egregius* and 496 bases in the Lacertid *Darevskia caucasica*. Four of the scincid sequences were slightly shorter because they lacked up to 13 bases at either the 3' or the 5' end of the sequence, which were treated as ambiguities (although these sites were included in the analyses where possible).

An initial alignment was achieved using ClustalX on the Scincidae and the Lacertidae independently. This provided good alignments due to considerable sequence conservation in these closely related taxa. Sites were then designated to helices or loop regions according to the W&VP/D or W&VP/E secondary structure models, and the initial alignments adjusted manually (similar to the procedure outlined by Kjer [1995]). Some variable-length helices provided many indels, while some very variable-length hairpin loops could not be confidently aligned and so were not used in subsequent analyses.

It was assumed that the W&VP models were accurate unless evidence was found to support alternative models. Mutual information (MI) and potential base pairing (PP) were obtained in half-matrices representing all possible pairwise site comparisons. PP proportions were assessed for the proposed models for each lineage. For values below 0.95, the elements corresponding to neighboring bases were searched within the matrices to investigate whether alternative base pairs could provide higher PP and MI values. The latter were only informative for sites showing above-average variability. Wherever possible, helices ending in hairpin loops were extended under these criteria, conditional on the presence of at least three bases in the terminal loop. G–U wobble pairs were accepted within helices but not as initial or terminal helix pairs.

### *Analysis of Secondary Structure Evolution*

A total of 56 models of sequence evolution were initially assessed by comparison of corresponding likelihoods computed for a neighbor-joining topology based on Jukes Cantor distances (MODELTEST version 3-06; Posada and Crandall 1998). Phylogenies and estimates of sequence parameters were obtained from Bayesian analyses of the sequences (MrBayes version 3.0b4; Huelsenbeck and Ronquist 2001), using a doublet model that specified interacting sites from the consensus models. Other likelihood settings for the sequence evolution model were guided by

the results of MODELTEST. The Bayesian analyses were run for 800,000 generations starting from a random tree, with 1 sample per 100 generations. This was repeated three times on each lizard group, and after removal of the initial 500–2500 samples prior to stationarity being reached (as assessed by the stabilization of the likelihoods), the consistency of the resultant three trees was checked. The 50% majority rule consensus trees were compared among the three runs. This was based on 21,500 trees for the Lacertidae and 17,450 trees for the Scincidae.

Analysis of covarion-like evolution was carried out using PAML version 3.14 (Yang 1997) by comparison of base pairs between the Lacertidae and the Scincidae. The input trees were the consensus trees inferred from the Bayesian analyses. Heterogeneity in site-specific rates within each lineage was assumed to follow a gamma distribution that could be approximated by assigning each site to one of eight rate categories (Yang 1994). Homologous interacting sites in the Lacertidae and Scincidae were analyzed and the rate categories compared, as they provided relative rates within a lineage. Thus, analysis of the categories was comparable between lineages, whereas actual rates themselves would be affected by the degree of divergence within the lineage.

## Results

### Secondary Structures

The findings for the 17 structures in the Lacertidae and Scincidae W&VP models are summarized in Table 1, and consensus models are shown in Figs. 1 and 2, respectively. Further details on structures in which the consensus models differ from the W&VP models are given below.

E28. For Lacertidae, the W&VP/D models contain an additional A–U base pair (compared to W&VP/E) at the hairpin loop end of the helix, but the very low PP value (0.22) suggests that this should not be included in a consensus model. The interaction could be due to helix length extension (and consequent loop reduction) in a small number of taxa, i.e., primarily *Darevskia*. This runs against expectation given that the entire hairpin loop sequence is relatively highly conserved (conserved in >60% of the Lacertidae and >70% of the Scincidae). The W&VP/E model is the best consensus in both cases.

F1. The high PP (0.96) allows a one-base pair extension of the base of the helix in the Scincidae. A similar extension was included for the Lacertidae, despite slightly lower support (PP = 0.91).

G2. An additional terminal A–U base pair (prior to G3) is conserved in all Scincidae and Lacertidae, and this is included in the consensus model.

G3. The extreme variability and high number of A–U pairings make this one of the most difficult secondary structures to align (Buckley et al. 2000). This is the only structure in which the W&VP models for closely-related *Darevskia* differ slightly from one another. The model shown (Fig. 1) provides higher PPs (>0.91

in all cases, as opposed to values as low as 0.51 for the *Darevskia* models) and is supported by compensatory substitutions for some base pairs. The initial three base pairs (all A–U) are conserved for all taxa, while the terminal end of the helix appears to be variable in length. The W&VP/E model does not appear to be entirely appropriate for the Scincidae. Similar to the Lacertidae, the first A–U bond for the proposed structure is invariant for all Scincidae, while the subsequent two bonds (both A–U) are present in 90% of taxa. A further base pair (also supported by MI) is followed by an internal loop. A conserved A–U base pair at/near the terminal end appears to be present.

G15. The W&VP/D model is supported for the Lacertidae, although there is PP and MI support for additional pairing between three more pairs of sites currently situated in the hairpin loop, leading to lengthening of the helix at the terminal end. An alternative to the W&VP/E is strongly supported in the Scincidae, with evidence for a four-base pair helix provided by high PPs and also high MI contents at two sites. None of the W&VP/E base pairs are incorporated and the proposed change makes it more similar to the W&VP/D model.

G18. High PPs support helix extension into the hairpin loop by one G–C base pair in the Lacertidae.

### rRNA Evolution

Some base positions were removed from the analyses because they were in highly variable loops and so difficult to align and/or in areas of putative helix length variation. These were found in the following structures: E25 (terminal loop and helix), noninteracting sites between G2 and G3, G3 (terminal loop and helix), some noninteracting sites between G8 and G13 and between G13 and G15, most of the terminal G15 loop, one indel site within G16, and one (Scincidae) or two (Lacertidae) indel sites between G17' and G19'. For the Lacertidae, this left 446 bases for analysis, of which 176 were involved in secondary structure interactions, while corresponding values for the Scincidae were 450 and 184, respectively.

Independent of the criterion used, comparisons of likelihoods for a Jukes Cantor tree suggested that a general time reversible model (Rodriguez et al. 1990) incorporating invariant sites (I) and gamma-distributed site rate heterogeneity (G) provided the best model of sequence evolution for the Scincidae. The same model was also selected for the Lacertidae under the Aikake information criterion, although the TrN model (Tamura and Nei 1993) with I and G was favored by hierarchical likelihood ratio testing. Thus all Bayesian analyses allowed different rates for all substitution types and gamma-distributed rate variation

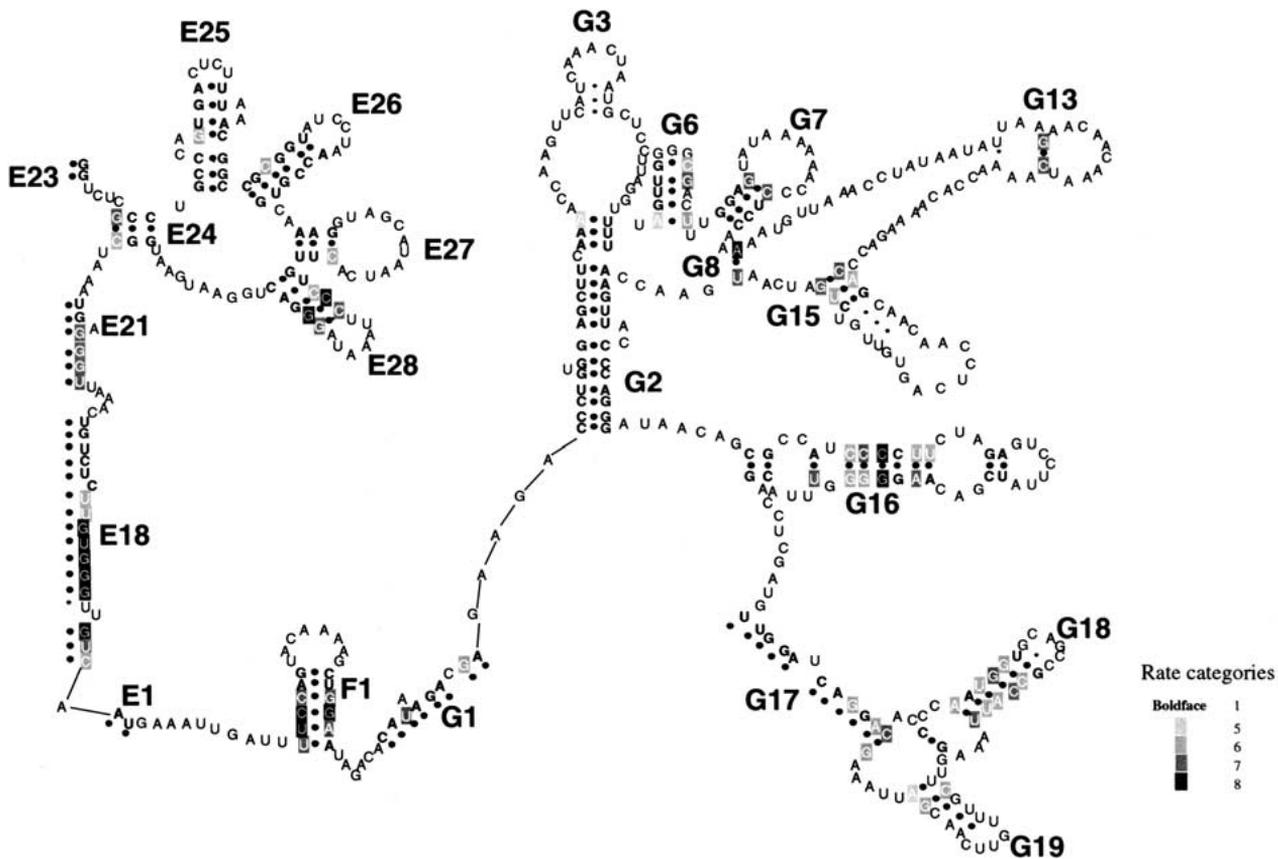
**Table 1.** Evaluation of helix base-pairing in the Lacertidae and Scincidae relative to Wuyts et al. (2001a) models, with information on hairpin loop variability

Wuyts et al. helix no. (Gutell no.)	Comparison of W&VP models for Lacertidae and Scincidae	Helices: Support for length variation and/or modifications to W&VP model		
		Lacertidae	Scincidae	Hairpin loop variation
E24 (67)	Same model	No	No	—
E25 (68)	Same	Helix extension by one base pair supported but rare ( <i>Ophisops elegans</i> & <i>Heliobolus spekii</i> only)	Helix may be extended by at least 1 base pair in > 50% and by 2 base pairs in < 10% of sequences	Considerable variability in Scincidae (3–7 bases). Less so in Lacertidae (mostly 4 bases), with CTCT motif conserved in > 80% of sequences
E26 (69)	Same	No	No	Length invariant (6 bases). AUYCUA conserved (>90%) within both lineages
E27 (70)	Same	No	No	Length invariant (12 bases). GUAGCRUAAUCA conserved (> 90%) within both lineages
E28 (71)	Different	Terminal couplet in model not present in majority of sequences, indicating length variation	No	Hairpin loop sequence highly conserved (> 70%) and completely length-conserved, except for < 30% of Lacertidae that follow W&VP/D model
F1 (72)	Same	Potential pairing lacking at initial 2 bases in 13–15% of sequences, but model supported. Some support for extension of base of helix by 1 base pair	Putative W&VP potential pairing lacking at helix base in 17% of sequences. Support for extending helix by 1 base pair at base end, as in Lacertidae	Loop of 7 bases in all sequences containing conserved central AAA motif in both lineages
G2 (74)	Same	Extend model by 1 base pair at terminal end	Extend model by 1 base pair at terminal end	—
G3 (75)	Different	Alternative helix proposed. Highly variable—consisting of 5–8 interactions interrupted by highly length-variable symmetrical loop	Alternative helix proposed. Highly variable—consisting of 8–11 interactions interrupted by highly length-variable symmetrical loop	Highly variable hairpin loop: 8–25 bases under proposed model for Scincidae and 4–12 bases for Lacertidae
G6 (80)	Same	No	No	Invariant: GGG in all taxa studied
G7 (81)	Same	Additional potential pairing (1 base pair) at base of model is found in 76% of sequences, suggesting possible length variation	Potential pairing and mutual information indicates extension of base of model by 1 base pair	Length conserved in both lineages (11 bases); high A composition
G8 (82)	Same	Potential pairing is only 0.85 for this single base pair, suggesting that alternative pairings may occur in some taxa	No	—
G13 (84)	Different	No	No	Differs between, but conserved within, lineages. Central ACA motif conserved in > 90% of sequences from both lineages
G15 (88)	Different	Potential pairing and mutual information support incorporation of internal bulge loop (1 base) and 3 or more additional base pairs to helix	Same: 4-base pair helix beginning at first base of Lacertidae model is more strongly supported.	Highly variable in Lacertidae (3–12 bases); less so in Scincidae (3–7 bases; modal length = 3)

(Continued)

Table 1. Continued

Wuyts et al. helix No. (Gutell Na.)	Comparison of W&VP models for Lacertidae and Scincidae	Helices: Support for length variation and/or modifications to W&VP model		
		Lacertidae	Scincidae	Hairpin loop variation
G16 (89)	Same	No	No	Length conserved and 5 of 7 sites (Scincidae) or 6 of 7 sites (Lacertidae) conserved in 90% of sequences
G17 (90)	Same	Only final base pair of helix covered in this study and supported	As for Lacertidae	—
G18 (91)	Same	Model supported but also support for extending terminal end of helix by 1 base pair	Extension of terminal end of helix as for Scincidae is only applicable in 76% of taxa and therefore not applied in the consensus model	Length and motif conserved in >70% of sequences within each lineage
G19 (92)	Same	No	No	Loop is UGUUC in all but one sequences studied



**Fig. 1.** Consensus secondary structure model for the Lacertidae shown for *Darevskia caucasica* (AF206187), based on potential pairing and mutual information analysis (note that this is based on site interactions for the majority of taxa, explaining the presence of

non-Watson–Crick/GU bonds and why some potentially bonding bases do not appear to be paired). Paired sites that were included in the analysis (i.e., those homologous with the Scincidae) are connected by a large dot; other paired sites are joined by a small dot.

across sites (six categories), with a proportion of the sites invariable. The doublet model was applied to interacting sites, with the “4by4” model for other sites. The Bayesian strict consensus trees (see Appendix II)

provided the topologies for subsequent analyses using PAML.

For the Lacertidae the assigned site rate categories were either 1, 5, 6, 7, or 8, while in the Scincidae the

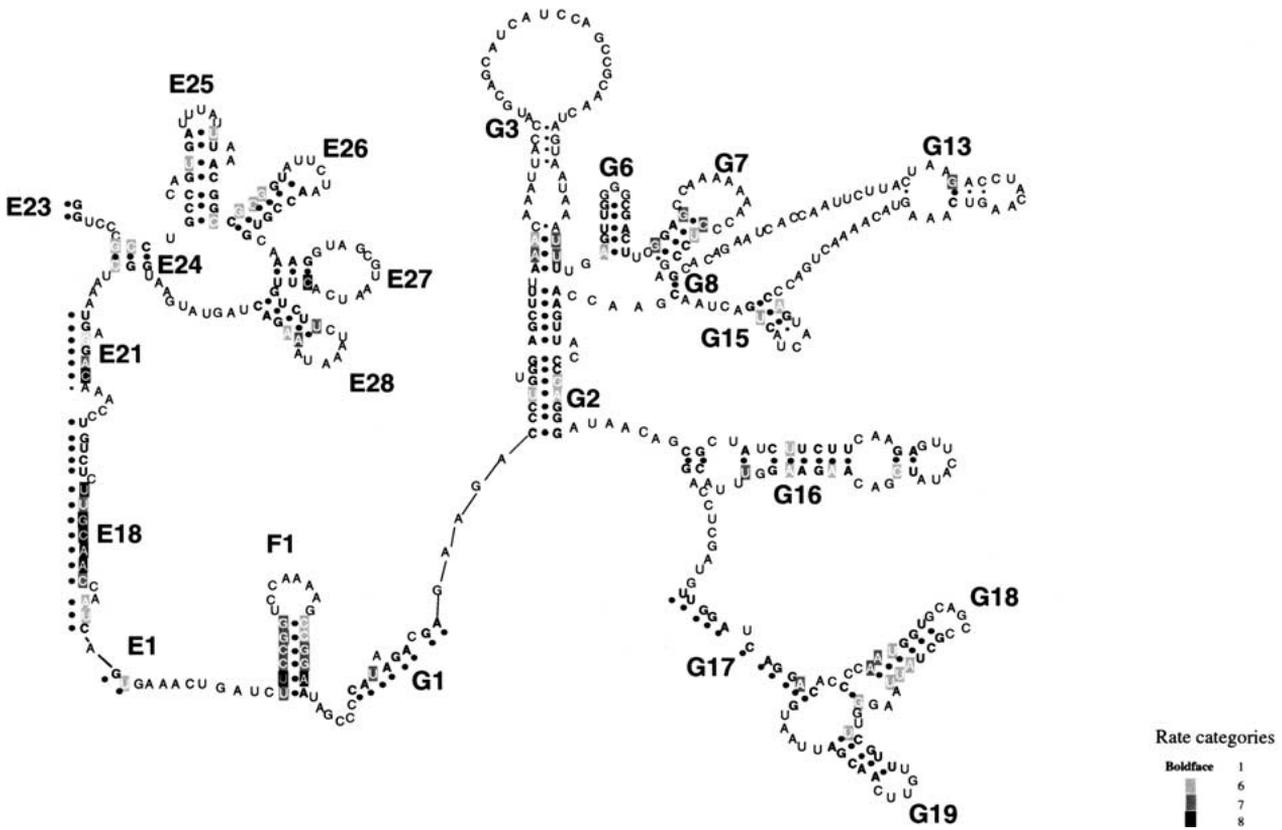
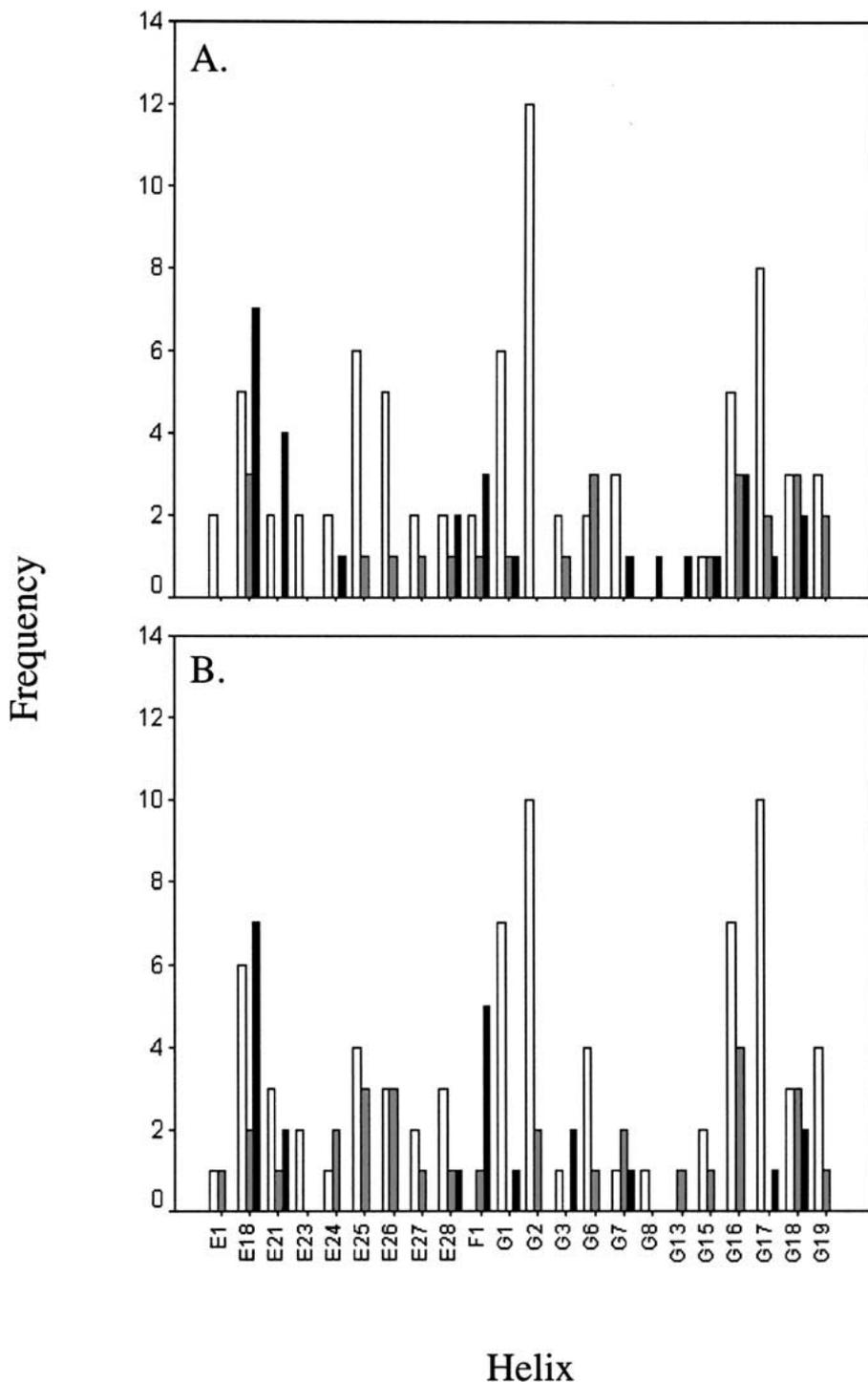


Fig. 2. Consensus secondary structure model for the Scincidae shown for *Eumeces egregius*. See Fig. 1 legend for details.

rates were either 1, 6, 7, or 8. Site rate assignment to categories does have an error term associated, although this decreases with the number of categories used and is less of a problem when the number of rate categories is greater than three, as in this case (Misof et al. 2002). As in previous studies (Misof et al. 2002), rate categories sometimes differed between sites within a pair (Figs. 1 and 2). Several explanations of this are possible, including deviations from the consensus structures, incorrect assignment, and substitutions in only one of the two sites that did not affect base pairing (e.g., G–C to G–U). Evidence of the latter was found, but examination of the original sequences revealed a more persistent cause: conserved sites with only a single substitution (potentially because of an error in the original sequence or an aberrant individual) could be assigned to rate category 5 or 6, while the corresponding site may be completely conserved and thus assigned to category 1.

Site-specific rate categories obtained for homologous interacting sites were used to obtain rate category values from the mean rate of the site pair (although a single value was used when only one of the paired sites was present) (Fig. 3). Deviations between Scincidae and the Lacertidae rate category assignments allows an assessment of the degree of covarion-like evolution between the two lineages

(Fig. 4). To overcome the potential problem of differing rate categories between paired sites outlined above, additional conservative analyses considered rate variation to exist only when one site pair was invariant, (i.e., rate category 1) while the other site pair averaged more than two substitutions each (i.e., corresponding to rates above 6). The proportion of between-lineage deviations that differ from zero, as expected when covarion-like evolution occurs, is 0.43 (i.e., 54). Using the normal approximation (for large sample sizes) for proportions, the 95% confidence interval for this is (0.34, 0.51), suggesting that approximately one-third or more of sites could show covarion-like evolution. The corresponding estimate for the more conservative analysis is 0.07 (i.e., 9), with corresponding 95% confidence intervals of (0.04, 0.13), which substantiates the overall finding. An analysis of whether covarion-type evolution is more concentrated in internal parts of the helices rather than terminals was investigated using a likelihood ratio test. The presence or absence of substantial differences in rate assignment was not significantly contingent on whether the interacting sites were internal (five deviations of  $\leq 5$  in 90) or not (four deviations  $> 5$  in 36) at the 5% significance level:  $G = 1.11$ ,  $p$  (Exact) = 0.44. The median deviation did not differ significantly from zero (Wilcoxon signed ranks procedure: smallest sum of



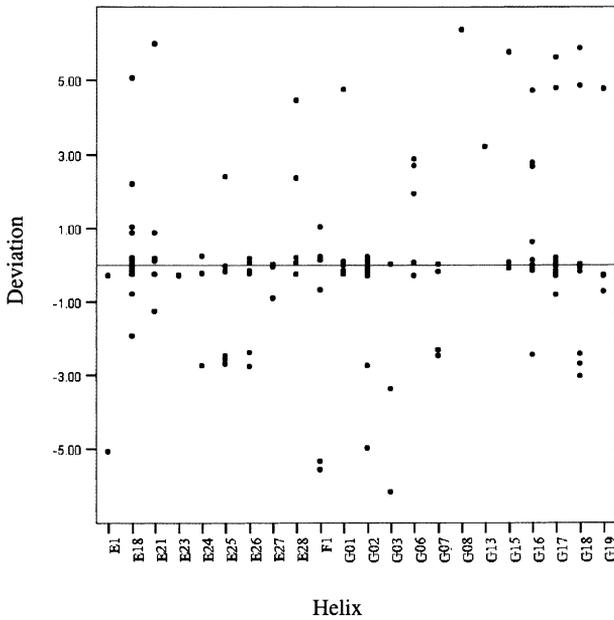
**Fig. 3.** Summary of frequencies of paired site rate categories by helix. The rate categories are simplified to fast (black bars; rate categories > 5), medium (gray bars; rate categories 2–5), and slow (open bars; rate category 1) rate categories for (A) Lacertidae and (B) Scincidae.

ranks = 637,  $p[\text{Exact}] = 0.34$ ), which clearly shows that there was no increase in the frequency of higher or lower rate category sites in one lineage relative to the other.

### Discussion and Conclusions

All secondary structure elements present in the *Eumeces* and *Darevskia* models from the LSU database (Wuyts et al. 2001a) were confirmed. This is not too

surprising, as these models are now known to be robust, being originally based on large-scale comparative analysis and confirmed by recent fine resolution of the crystal structure of the LSU from an archaebacterium (see Wuyts et al. 2001a). However, adjustments to the pattern of site interactions were possible and this provided more accurate specification for doublet models in phylogenetic analyses. Early comparative studies of secondary structure tended to be limited by the trade-off between using related sequences with



**Fig. 4.** Rate category deviation between homologous interacting sites in the Lacertidae and the Scincidae, by helix. Scincidae rates were subtracted from Lacertidae rates, so positive deviations indicate a higher rate category for the latter, and vice versa. Note that points are jittered.

nearly identical patterns of site interactions but highly conserved sites that provide little mutual information as support for base pairing and using highly divergent sequences with potentially high mutual information but different site interactions. Here and elsewhere (Buckley et al. 2000) it has been shown that the latter have provided good overall models but that they can be further refined by comparative analyses of taxa with lower levels of divergence.

In general, helices with the highest evolutionary rates (i.e., those with highest proportions of rapidly evolving base pairs in both lineages) were E18, E21, and F1 (Figs. 1, 2, and 3). G16 and G18 also contained significant proportions of rapidly evolving sites. It may appear surprising that the highest proportions of rapidly evolving base pairs were not detected within G3 and G15, as these are possibly the most variable helices. However, this might be explained by the conservative nature of the remaining sites once sites that were difficult to align had been removed. Higher rates in E18 and E21 are largely expected, as these correspond to domain IV, which comprises most of the subunit-subunit interface surface (Ban et al. 2000). Although these helices are not part of the prominent active cleft site, they are known to be either quite or extremely variable, respectively (Wuyts et al. 2001b). F1 is also adjacent to domain IV. Domain V encompasses F1 and subsequent helices and is involved in peptidyl transferase activity. It is found in the middle of the subunit, where variability is expected to be lower (Wuyts et al. 2001b). It helps stabilize the elongation factor-binding region of

the ribosome and is generally quite conserved across higher taxa (Ban et al. 2000; Wuyts et al. 2001b).

Evidence for covarion-like evolution is provided, even though very large differences in rates between the two lizard families were only observed at approximately 7% of homologous base pairs. This represents one of the first observations of such a pattern between such closely related clades, indicating that SSRV is widespread. The distribution pattern of the major covarion sites is difficult to compare with those described in insects (Misof et al. 2002) due to the relatively small number that were detected. However, all of the helices in which SSRV was detected in this study also showed SSRV in insects, except for G17 and G18 (and G19 if all potential covarion sites are considered). Why some sites within these generally conserved helices should evolve at different rates between lizard lineages is difficult to ascertain, although it is possible that the factors determining SSRV differ at different taxonomic levels.

The impact of nonindependent evolution of bases within a base pair and site-specific rate variation on phylogenetic tree construction was not explicitly examined by the current study. Bayesian analyses that did not implement the doublet model were performed on both data sets and found to provide similar trees to those described, suggesting that the impact of correct site pair specification may not be that large. Despite this, it is clearly preferable to account for these known sequence evolution constraints and specify them as correctly as possible in any analysis. Furthermore, a detailed analysis of RNA sequence evolution shows that models incorporating secondary structure information can provide significantly improved log likelihoods for the data (under a given topology) than simpler models that ignore compensatory substitutions (Savill et al. 2001). The impact of specifying covarion-like models for LSU and SSU rRNA evolution has also been examined recently (Galtier 2001). Again, incorporation of the more sophisticated models was found to lead to a significant increase in log likelihood. The present study provides additional empirical support for use of the latter even at lower levels of sequence divergence. Future studies should be directed toward obtaining further empirical evidence of covarion-like evolution, the interdependencies of these rate changes, and the corresponding functional implications for the rRNA itself.

Finally, the phylogenetic trees used here were based on a single mtDNA fragment and so unlikely to be as well-resolved or robust as those based on larger data sets: they were constructed solely for use in subsequent analyses of rRNA evolution (this is not a systematics study). However, it is worth commenting on inferred phylogenetic relationships that either deviate from or include taxa not present in other molecular phylogenetics studies (Fu 1998, 2000; Harris et al. 1998a,b; Lin et al. 2002; Whiting et al. 2003). For the Lacertidae

(Appendix IIA), the (*Zootoca/Lacerta vivipara*, *Lacerta/Archeolacerta bedriagae*) lineage was well supported, as were the (*Takydromus sexlineatus*, *Lacerta cappadocica wolteri*) and (*Archaelaceta mosorensis*, *Algyroides fitzingeri*) lineages. The grouping of *Psammodromus algirus* and *Lacerta andreanskyi* within the smaller of the two major lineages (corresponding to the Eremiainae [see Harris et al. 1998a]) is also surprising. For the Scincidae (Appendix IIB), the recently inferred interesting phylogeographical relationships within the large *Mabuya* genus (i.e., (Asia, (Cape Verde, (Africa/Madagascar, South America))) with *Mabuya atlantica* from South America within the Africa clade (see Mausfeld et al. 2003) was still strongly supported, despite my inclusion of more taxa and a doublet model of sequence evolution. However, the very well-supported clade comprising *Carlia*, *Emoia*, *Eugonglyus*, two *Eumeces*, and *Morethia* has not been observed previously. Honda et al. (2003) included representatives of the first four of these but their analysis grouped only the first three genera within a well-supported clade, possibly because *Eumeces* was used as an outgroup in the analysis. Future studies could further explore some of these relationships by testing for congruence across multiple DNA markers.

**Acknowledgments.** I wish to thank two anonymous reviewers for very helpful comments and Liverpool John Moores University for supporting this work.

## References

- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:878–879
- Buckley TR, Simon C, Flook PK, Misof B (2000) Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect large subunit rRNA gene. *Insect Mol Biol* 9:565–580
- Dixon MT, Hillis DM (1993) Ribosomal RNA secondary structure: Compensatory mutations and implications for phylogenetic analysis. *Mol Biol Evol* 10:256–267
- Espinosa de los Monteros A (2003) Models of the primary and secondary structure for the 12S rRNA of birds: A guideline for sequence alignment. *DNA Sequence* 14:241–256
- Fu J (1998) Toward the phylogeny of the family Lacertidae: Implications from mitochondrial DNA 12S and 16S gene sequences (Reptilia: Squamata). *Mol Phylogenet Evol* 9:118–130
- Fu J (2000) Toward the phylogeny of the family Lacertidae—Why 4708 base pairs of mtDNA sequences cannot draw the picture. *Biol J Linn Soc* 71:203–217
- Galtier N (2001) Maximum-likelihood phylogenetic analysis under a covarion-like model. *Mol Biol Evol* 18:866–873
- Gutell RR, Fox GE (1988) A compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Res* 6:R175–R269
- Gutell RR, Weiser B, Woese CR, Noller HF (1985) Comparative anatomy of 16S-like ribosomal RNA. *Prog Nucleic Acid Res* 32:155–216
- Harris DJ, Arnold EN, Thomas RH (1998a) Relationships of lacertid lizards (Reptilia: Lacertidae) estimated from mitochondrial DNA sequences and morphology. *Proc Roy Soc Lond B* 265:1939–1948
- Harris DJ, Arnold EN, Thomas RH (1998b) Rapid speciation, morphological evolution, and adaptation to extreme Environments in South African sand lizards (*Meroles*) as revealed by mitochondrial gene sequences. *Mol Phylogenet Evol* 10:37–48
- Hickson RE, Simon C, Perrey SW (2000) The performance of several multiple-sequence alignment programs in relation to secondary-structure features for an rRNA sequence. *Mol Biol Evol* 17:530–539
- Honda M, Ota H, Kobayashi M, Nabhitabhata J, Yong H-S, Hikida T (2000) Phylogenetic relationships, character evolution and biogeography of the subfamily (Reptilia: Scincidae) inferred from mitochondrial DNA sequences. *Mol Phylogenet Evol* 15:452–461
- Honda M, Ota H, Köhler G, Ineich I, Chirio L, Chen S-L, Hikida T (2003) Phylogeny of the lizard subfamily Lygosominae (Reptilia: Scincidae), with special reference to the origin of the New World taxa. *Genes Genet Syst* 78:71–80
- Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:154–755
- Juan V, Wilson C (1999) RNA secondary structure prediction based on free energy and phylogenetic analysis. *J Mol Biol* 289:935–947
- Lin SM, Chen CA, Lue KY (2002) Molecular phylogeny and biogeography of the grass lizards genus *Takydromus* (Reptilia: Lacertidae) of east Asia. *Mol Phylogenet Evol* 22:276–288
- Kjer KM (1995) Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: An example of alignment and data presentation from frogs. *Mol Phylogenet Evol* 4:314–330
- Mausfeld P, Vences M, Schmitz A, Veith M (2000) First data on the molecular phylogeography of Scincid lizards of the genus *Mabuya*. *Mol Phylogenet Evol* 17:11–14
- Mausfeld P, Schmitz A, Bohme W, Misof B, Vrcibradic D, Rocha CFD (2002) Phylogenetic affinities of *Mabuya atlantica* Schmidt 1945, endemic to the Atlantic ocean archipelago of Fernando de Noronha (Brazil): necessity of partitioning the genus *Mabuya* Fitzinger, 1826 (Scincidae, Lygosominae). *Zool Anzeiger* 241:281–293
- Misof B, Anderson CL, Buckley TR, Erpenbeck D, Rickert A, Misof K (2002) An empirical analysis of mt 16S rRNA covarion-like evolution in insects: Site specific rate variation is clustered and frequently detected. *J Mol Evol* 55:460–469
- Parsch J, Braverman JM, Stephan W (2000) Comparative sequence analysis and patterns of covariation in RNA secondary structures. *Genetics* 154:909–921
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818
- Rodriguez F, Oliver JF, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitutions. *J Theor Biol* 142:485–501
- Savill NJ, Hoyle DC, Higgs PG (2001) RNA sequence evolution with secondary structure constraints: Comparison of substitution rate models using maximum-likelihood methods. *Genetics* 157:399–411
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512–526
- Whiting AS, Bauer AM, Sites JW Jr (2003) Phylogenetic relationships and limb loss in Saharan African scincine lizards (Squamata: Scincidae). *Mol. Phylogenet Evol* 29:582–598
- Wuyts J, De Rijk P, Van de Peer Y, Winkelmans T, De Wachter R (2001a) The European large subunit ribosomal RNA database. *Nucleic Acids Res* 29:175–177
- Wuyts J, Van de Peer Y, De Wachter R (2001b) Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA. *Nucleic Acids Res* 29:5017–5028

- Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol* 39:306–314
- Yang Z (1997) PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556
- Zuker M (1989) On finding all suboptimal foldings of an RNA molecule. *Science* 244:48–52

## Appendix I. EMBL Accession Numbers for the Sequences Used in the Current Study

Note that the species/subspecies names correspond to those on the database, even though this might differ from the current usage.

### Scincidae

AB016606, *Eumeces egregius*; AB028812, *Eumeces schneideri*; AB028813, *Emoia caeruleocauda*; AB028814, *Eugongylus* sp.; AB028815, *Isopachys anguinoides*; AB028816, *Lipinia vittigera*; AB028817, *Lygosoma koratense*; AB028818, *Lygosoma quadrupes*; AB028819, *Mabuya bladingii*; AB028820, *Sphenomorphus indicus*; AB028821, *Sphenomorphus maculatus*; AB028822, *Sphenomorphus praesignis*; AB028823, *Tropidophorus berdmorei*; AF054540, *Eumeces algeriensis*; AF153561, *Mabuya affinis*; AF153562, *Mabuya binotata*; AF153563, *Mabuya bistriata*; AF153564 and AY028888, *Mabuya capensis*; AF153565, *Mabuya comorensis*; AF153566, *Mabuya* cf. *dumasi*; AF153567 and AF153568, *Mabuya elegans*; AF153569, *Mabuya hoeschi*; AF153570, *Mabuya homalocephala*; AF153571, *Mabuya irregularis*; AF153572, AY070359, and AY169607, *Mabuya longicaudata*; AF153573 and AY070353, *Mabuya macularia*; AF153574 and AY070356, *Mabuya maculilabris*; AF153575, *Mabuya margeritifera*; AF153576, *Mabuya multifasciata*; AF153577 and AY028890, *Mabuya occidentalis*; AF153578, *Mabuya perroteti*; AF153579, *Mabuya quinquetaeniata*; AF153580, *Mabuya spilogaster*; AF153581, *Mabuya striata striata*; AF153582, *Mabuya striata wahlbergi*; AF153583, *Mabuya sulcata*; AF153584, *Mabuya varia*; AF153585, *Mabuya variegata*; AF153586, *Scelotes mirus*; AF202628 and AF202629, *Mabuya frenata*; AF202630, *Mabuya cochabambae*; AF202631 and AY070363, *Mabuya dorsivittata*; AF215234, *Chalcides ocellatus*; AF215235, *Pygomeles braconnieri*; AF215236, *Sphenops* sp.; AF215237, *Voeltzkowia ferinensis*; AF215238, *Voeltzkowia lineata*; AY028869 and AY028870, *Acontias gracilicauda*; AY028872 and AY028873, *Acontophiops lineatus*; AY028874 and AY028875, *Acontias lineatus*; AY028876 and AY028877, *Acontias tristis*; AY028878–AY028880, *Acontias meliagrís*; AY028881–AY028884, *Acontias orientalis*; AY028885–AY028887, *Acontias percevali*; AY028889, *Mabuya striata*; AY028891, *Scelotes anguena*; AY028893, *Scelotes mirus*; AY070347, *Mabuya agilis*; AY070348, *Mabuya spinalis salensis*; AY070349, *Mabuya heathi*; AY070350, *Mabuya albilabris*; AY070351, *Mabuya macrorhyncha*; AY070352, *Mabuya nigropunctata*; AY070354, *Mabuya carinata*; AY070355, *Mabuya boettgeri*; AY070357, *Mabuya mabouya*; AY070358, *Mabuya gravenhorstii*; AY070360, *Mabuya spinalis spinalis*; AY070361, *Mabuya delalandii*; AY070362, *Mabuya*

*atlantica*; AY141000, *Acontias percevali*; AY169601, *Carlia* sp.; AY169602, *Egernia whitii*; AY169603, *Emoia physicae*; AY169604, *Eugongylus rufescens*; AY169605, *Lamprolepis smaragdina*; AY169606, *Lygosoma fernandi*; AY169608, *Euprepis perroteti*; AY169609, *Morethia butleri*; AY169610, *Tiliqua adelaidensis*; AY169611, *Tribolonotus gracilis*; AY169612, *Anomalopus mackayi*; AY169613, *Anomalopus swansonii*; AY169614, *Calyptotis scutiostrium*; AY169615, *Coeranoscincus reticulatus*; AY169616, *Ctenotus leonhardii*; AY169617, *Ctenotus pantherinus*; AY169618, *Ctenotus robustus*; AY169619, *Eremiascincus richardsonii*; AY169620, *Eulamprus amplus*; AY169621, *Eulamprus murrayi*; AY169622, *Eulamprus quoyii*; AY169623, *Glaphyromorphus arnhemicus*; AY169624, *Glaphyromorphus isolepis*; AY169625, *Glaphyromorphus gracilipes*; AY169626, *Gnypetoscincus queenslandiae*; AY169627, *Hemiergis peroni*; AY169628, *Lerista bipes*; AY169629, *Lerista bougainvillii*; AY169630, *Nangura spinosa*; AY169631, *Notoscincus ornatus*; AY169632, *Ophisoscincus ophisoscincus*; AY169633, *Prasinochaema virens*; AY169634, *Saiphos equalis*; AY169635, *Scincella lateralis*; AY169636, *Sphenomorphus muelleri*.

### Lacertidae

AF206187, *Darevskia caucasica*; AF206185, *Darevskia alpina*; AF206183, *Darevskia brauneri*; AF206190, *Darevskia clarkorum*; AF206182, *Darevskia lindholmi*; AF206184, *Darevskia saxicola*; AF206186, *Darevskia praticola*; AF206188, *Darevskia daghestanica*; AF206189, *Darevskia mixta*; AF206191, *Darevskia derjugini*; AF206192, *Darevskia raddei*; AF206193, *Darevskia rudis*; AF206194, *Darevskia portschinskii*; AF206195, *Darevskia parvula*; AY184996, AY184997, and AF206599, *Podarcis sicula*; AY185009–AY185019, *Podarcis melisellensis*; AF206593, *Lacerta monticola*; AY278187–AY278192, and AF440616, *Archaeolacerta oxycephala*; AY190306, AF206600, and AF440618, *Podarcis muralis*; AF149938, *Darevskia rudis bischoffi*; AF149943, *Iberolacerta horvathi*; AF149946, *Lacerta cappadocica wolteri*; AF206589, *Takydromus sexlineatus*; AF206590, *Lacerta media*; AF206591, *Lacerta schreiberi*; AF206592, *Lacerta bedriagae*; AF206594, *Lacerta vivipara*; AF206595 and AF378948–AF378954, *Lacerta lepida*; AF206596 and AF378958, *Lacerta pater*; AF206597, *Darevskia valentini*; AF206598, *Algyroides fitzingeri*; AF206601, *Podarcis hispanica*; AF206602, *Lacerta perspicillata*; AF206603, *Lacerta andreanskyi*; AF206604, *Eremias velox*; AF206605, *Ophisops elegans*; AF206606 and AY035841, *Mesalina brevirostris*; AF206607, *Acanthodactylus erythrurus*; AF206608, *Heliobolus spekii*; AF206609, *Latastia longicaudata*; AF206610, *Meroles ctenodactylus*; AF20661, *Meroles suborbitalis*; AF206612, *Nucras tessellata*; AF206613, *Pedioplanis namaquensis*; AF206614, *Adolfus vauereselli*; AF206615, *Adolfus jacksoni*; AF206616, *Tropidosaura gularis*; AF316170, *Zootoca vivipara sachalinensis*; AF378955–AF378957, *Lacerta tangitana*; AF440604, *Iberolacerta monticola monticola*; AF440605, *Iberolacerta monticola cantabrica*; AF440606, *Iberolacerta monticola cantabrica*; AF440607, *Iberolacerta cyreni cyreni*; AF440608, *Iberolacerta cyreni castiliana*; AF440609, *Iberolacerta cyreni martinezricai*; AF440610, *Iberolacerta aurelioi*; AF440611 and AF440612, *Iberolacerta aranica*; AF440613, *Iberolacerta bonnali*; AF440614, *Archaeolacerta bedriagae*; AF440615, *Archaeolacerta mosorensis*; AF440617, *Lacerta graeca*; AY035834, *Mesalina balfourii*; AY035837, *Eremias arguta*; AY035838, *Eremias pleskei*; AY035839, *Mesalina olivieri*; AY035842, *Mesalina guttulata*; AF206588, *Psammotromus algirus*.



