

**REASSESSMENT OF COMPARATIVE GENETIC DISTANCE IN REPTILES FROM THE MITOCHONDRIAL CYTOCHROME *b* GENE**

D. J. HARRIS

*Unidade de Genética Animal e Conservação, Campus Agrário de Vairão, Portugal*

*Key words:* base composition, Cytochrome *b*, genetic distance, reptiles

Mitochondrial cytochrome *b* is one of the most widely sequenced genes in vertebrates. As such it is commonly used as a genetic “yardstick”, especially for comparing congeneric species for which it is typically phylogenetically informative and not severely saturated (Moritz, Dowling & Brown, 1987). Recently, Johns & Avise (1998) presented data culled from Genbank to make comparisons of genetic distances between congeneric species and confamilial genera across the major vertebrate classes. They also presented percentage base compositions across all sites, and for the third positions only. However, for those interested in reptiles this study

TABLE 1. Base composition in third position sites of cytochrome *b* sequences.\*Chi squared tests indicate significant heterogeneity of base frequencies across taxa within the Iguanidae.

Taxon	Percentage base composition in third position sites			
	A	C	G	T
Teiidae	41.6	34.5	3.2	20.7
Xantusiidae	39.1	34.7	3.9	22.3
Scincidae	37.7	38.7	4.3	19.3
Gekkonidae	36.1	44.1	5.3	14.6
Chelonidae	44.6	42.2	1.8	11.4
Emydidae	43.2	39.7	3.4	13.6
Testudinidae	41.0	38.8	2.4	17.8
Colubridae	35.5	39.1	3.3	22.1
Elaphidae	37.8	34.3	3.8	24.1
Viperidae	31.8	44.3	4.9	18.9
Anguidae	40.6	35.2	4.0	20.2
Iguanidae*	40.0	39.1	3.5	17.5
Lacertidae	31.2	38.4	3.3	27.1
Mean	38.5	38.7	3.6	19.2
Johns & Avise (1998)	29.7	28.1	18.0	24.2
All positions average	28.2	28.8	13.7	29.2
Johns & Avise (1998)	27.7	27.6	16.1	28.6

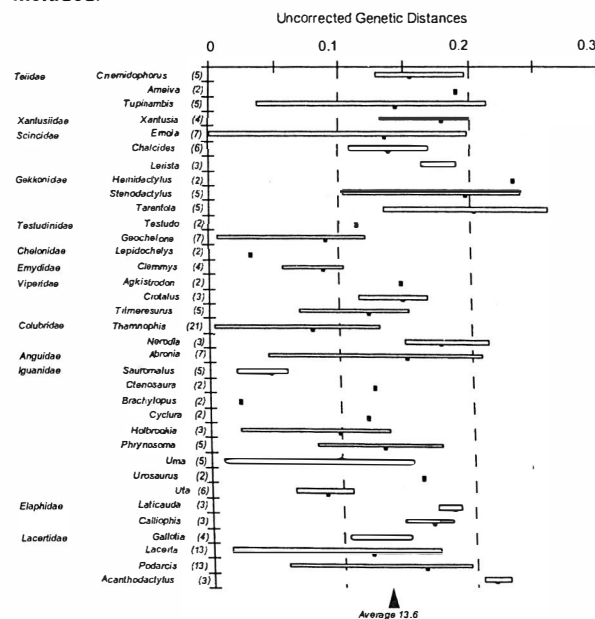
*Correspondence:* D. J. Harris, Unidade de Genética Animal e Conservação, Campus Agrário de Vairão, R. Monte-Crasto, 4485-661 Vila do Conde, Portugal. *Email:* james@mail.icav.up.pt

was limited by the data available at the time – of the 16 genera examined, seven were snakes and three iguanids, thus possibly causing distortion towards averages from these groups. Also, some researchers unfortunately fail to follow the standard convention of submitting the sense strand for protein coding sequences, and instead submit the reverse complement to Genbank. In at least one case these were included in the previous analysis in the reverse direction relative to the other sequences (Johns, pers. comm.). This affects the base composition estimates for all positions. Here I include data for reptile groups that were not previously available, increasing the number of families examined from eight to 13, and the number of genera from 16 to 35. I also calculate the correct base compositions as compared to those previously published (Johns & Avise, 1998).

All sequences were downloaded from Genbank and aligned by eye. The alignment is available on request from the author. There were no indel insertions or deletions. All analyses were carried out in PAUP\* 4.0 (Swofford, 2000). Average overall, and third position only, base compositions were calculated for each family, and uncorrected genetic distances were estimated for all pairwise comparisons within genera (Table 1, Fig. 1).

Base composition in the third positions is now very different from that previously reported, and much more similar to that obtained for other vertebrate groups. In particular, the very low proportion of guanines (3.4-5.5% in other vertebrates) is similar in the reptiles (3.6%; not the 18% reported by Johns & Avise, 1998). Overall base compositions are also significantly different. Assessment of the correct base composition is very important, as it is often used as an indication that mito-

FIG. 1. Uncorrected genetic distances in cytochrome *b* gene sequences between reptile species within 35 genera. Numbers of species assayed are given in parentheses. Mean genetic divergence estimates and ranges are shown. Only a monophyletic subset of sequences from *Lacerta* was included.



chondrial DNA has been amplified for phylogenetic studies, as opposed to nuclear-integrated copies (e.g. Macey *et al.*, 1999). There is significant variation in base composition between genera within the Iguanidae, something which will be problematic for researchers using cytochrome *b* to assess phylogenetic relationships within this family (e.g. Wiens & Hollingsworth, 2000). The mean cytochrome *b* genetic distance for congeneric reptiles is higher than the previous estimate (13.6%, compared to 12%).

This reassessment, like the study of Johns & Avise (1998), highlights the important role Genbank can play in large scale analyses, but it also exposes some problems with the present system. Despite the fact that Johns & Avise (1998) rejected two sequences as probable pseudogenes, three of the published and analysed sequences (from *Crotalus* and *Agkistrodon*, Cullings *et al.*, 1997) require single base pair insertions to align them with those of other reptiles. A similar error has been reported for the analysed *Gallotia* sequences, where one of the published sequences has an erroneous single base pair insertion (Carranza *et al.*, 1999). These errors will affect any analyses in which they are not detected. Another problem is that some assessment of taxonomy is required. For example, the analysed genus *Lacerta* is known to be paraphyletic (Arnold, 1989; Harris, Arnold & Thomas, 1998). Further, there is little agreement on the generic status of some lizards – sequences from the lizard endemic to Madeira have been deposited in Genbank simultaneously under *Teira dugesii*, *Lacerta dugesii* and *Podarcis dugesii*. Any attempt to use Genbank to assess congeneric levels of variation will run into difficulties in this situation unless researchers are aware of these problems.

While cytochrome *b* is still the most obvious genetic yardstick for most vertebrate groups, the inclusion of new data has altered both the previously reported base compositions (Johns & Avise, 1998), and average congeneric divergences. Clearly, as more data are submitted to Genbank both can be quantified across a wider range of taxa, allowing more accurate calibration.

*Acknowledgements.* This work was funded by grants from FCT: SFRH/BPD/5702/2001 and POCTI/34547/BSE/2000.

#### REFERENCES

- Arnold, E. N. (1989). Towards a phylogeny and biogeography of the Lacertidae: relationships within an Old-World family of lizards derived from morphology. *Bulletin of the British Museum (Natural History). Zoology Series* **55**, 209-257.
- Carranza, S., Arnold, E. N., Thomas, R. H., Mateo, J. A. & Lopez-Jurado, F. (1999). Status of the extinct giant lacertid lizard *Gallotia simonyi simonyi* assessed using mtDNA sequences from museum specimens. *Herpetol J.* **9**, 83-86.
- Cullings, K. W., Morafka, D. J., Hernandez, J. & Roberts, J. W. (1997). Reassessment of phylogenetic

relationships among pitviper genera based on mitochondrial cytochrome *b* gene sequences. *Copeia* **1997**, 429-432.

- Harris, D. J., Arnold, E. N., & Thomas, R. H. (1998). Relationships of the lacertid lizards (Reptilia: Lacertidae) estimated from mitochondrial DNA sequences and morphology. *Proc. R. Soc. London B.* **265**, 1939-1948.
- Johns, G. C. & Avise, J. C. (1998). A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Mol. Biol. Evol.* **15**, 1481-1490.
- Macey, J. R., Schulte II, J. A., Larson, A., Tuniyev, B. S., Orlov, N. & Papenfuss, T. J. (1999). Molecular Phylogenetics, tRNA Evolution, and Historical Biogeography in Anguid Lizards and Related Taxonomic Families. *Mol. Phyl. Evol.* **12**, 250-272.
- Moritz, C., Dowling, T. E. & Brown, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annu. Rev. Ecol. Syst.* **18**, 269-292.
- Swofford, D. L. (2000). PAUP\*: phylogenetic analysis using parsimony (and other methods). Version 4.0b1. Sinauer, Sunderland, Mass.
- Wiens, J. J. & Hollingsworth, B. D. (2000). War of the Iguanas: Conflicting molecular and morphological phylogenies and long-branch attraction in Iguanid lizards. *Syst. Biol.* **49**, 143-159.

*Accepted: 29.1.02*