

9-1-2011

Nonadaptive evolution of mitochondrial genome size

Bastien Boussau
University of California, Berkeley

Jeremy M. Brown
University of California, Berkeley

Matthew K. Fujita
Harvard University

Follow this and additional works at: https://digitalcommons.lsu.edu/biosci_pubs

Recommended Citation

Boussau, B., Brown, J., & Fujita, M. (2011). Nonadaptive evolution of mitochondrial genome size. *Evolution*, 65 (9), 2706-2711. <https://doi.org/10.1111/j.1558-5646.2011.01322.x>

This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Digital Commons. For more information, please contact ir@lsu.edu.

NONADAPTIVE EVOLUTION OF MITOCHONDRIAL GENOME SIZE

Bastien Boussau,^{1,2,3} Jeremy M. Brown,¹ and Matthew K. Fujita⁴

¹Department of Integrative Biology, University of California—Berkeley, 3060 Valley Life Sciences Building, Berkeley, California, 94720–3140

²Laboratoire de Biométrie et Biologie Evolutive, Université de Lyon, Université Lyon 1, CNRS, UMR5558, Villeurbanne, France

³E-mail: bastien.boussau@univ-lyon1.fr

⁴Museum of Comparative Zoology, Harvard University, 26 Oxford St., Cambridge, Massachusetts, 02138

Received January 13, 2011

Accepted March 31, 2011

Genomes vary greatly in size and complexity, and identifying the evolutionary forces that have generated this variation remains a major goal in biology. A controversial proposal is that most changes in genome size are initially deleterious and therefore are linked to episodes of decrease in effective population sizes. Support for this hypothesis comes from large-scale comparative analyses, but vanishes when phylogenetic nonindependence is taken into account. Another approach to test this hypothesis involves analyzing sequence evolution among clades where duplications have recently fixed. Here we show that episodes of fixation of duplications in mitochondrial genomes of the gecko *Heteronotia binoei* (two independent clades) and of mantellid frogs (five distinct branches) coincide with reductions in the ability of selection to purge slightly deleterious mutations. Our results support the idea that genome complexity can arise through nonadaptive processes in tetrapods.

KEY WORDS: D_n/D_s , genome evolution, lizards, mantellid frogs, mitochondria, parthenogenetic duplication.

Lynch and Conery (2003) proposed that gene duplication events and other changes that alter genome sizes are initially deleterious even if they ultimately become adaptive. The initial survival of duplicate genes is thus dependent on small effective population size, in which selection is weakened. Support for this hypothesis originally came from the observation of a strong negative correlation between genome size and $N_e u$, a product of effective population size (N_e) and mutation rate (u) that can be estimated from silent-site diversity. However, this strong negative correlation may have been confounded by the phylogenetic relatedness between species. Although the first phylogenetically corrected analysis supported Lynch and Conery's hypothesis (based on 33 fish species, Yi and Strelman 2005, results later disputed by Gregory and Witt 2008), the most recent studies found no significant correlation between genome size and $N_e u$ when accounting for phylogenetic structure in a dataset of 205 seed plants

(Whitney et al. 2010) and in Lynch and Conery's (2003) original dataset (30 taxa, Whitney and Garland 2010). Such negative results may arise from the lack of a substantive linear relationship between genome size and $N_e u$, from a lack of power in Lynch and Conery's original dataset (Whitney and Garland 2010), or from the inability of Brownian motion based models used in standard comparative methods to adequately capture the true process of genome size evolution.

An alternate avenue to test this hypothesis is to examine the ratio of nonsynonymous to synonymous substitutions (D_n/D_s) among recently diverged taxa that differ in genome size. According to Lynch and Conery's (2003) hypothesis, D_n/D_s should increase as genome size changes. In bacteria, Kuo, Moran and Ochman (2009) found that D_n/D_s increased as genome size decreased. It appears that, in prokaryotes, a strong deletion bias leads to the fixation by drift of a large excess of deletions compared to

insertions. Therefore, as the effective population size decreases, D_n/D_s increases, but genome size decreases. Because this deletion bias is not as strong in eukaryotes as it is in prokaryotes (Kuo and Ochman 2009), it remains to be seen whether Lynch and Conery's (2003) hypothesis can explain increases in genome size among other groups of organisms.

Tetrapod mitochondrial genomes can be used to test Lynch and Conery's (2003) hypothesis due to their dense sampling among some recently diverged taxa. Small, tandem duplications in the control region of mitochondrial genomes have often occurred along the branches of the tetrapod phylogeny. Larger duplications in other parts of the mitochondrial genome are much less frequent and are often found on terminal branches, suggesting that they are generally deleterious. Such mitochondrial duplications occurred recently in the gecko *Heteronotia binoei* (Moritz 1991) and in mantellid frogs of Madagascar (Kurabayashi et al. 2008), enabling us to address how such changes evolve on short timescales.

There are several mechanisms that can generate duplications in mitochondrial genomes. Duplications in parthenogenetic *Heteronotia* likely occurred by a slipped-strand mispairing mechanism, whereby the leading strand dissociates during DNA replication and re-anneals at a downstream location, causing the re-synthesis (and therefore duplication) of a portion of the genome (Fujita et al. 2007). In the mantellid frogs, several of the duplications are thought to have occurred by intramolecular recombination, which buds off a mini-circle that can reincorporate at any position back into the genome, creating nontandem duplications (Kurabayashi et al. 2008; see also Mueller and Boore 2005 for examples in salamanders).

In *H. binoei*, two independent lines of evidence support the idea that increases in mitochondrial genome size arose nonadaptively. First, large duplications are restricted to two independent parthenogenetic lineages (Fujita, Boore and Moritz 2007; Fig. 1A, B), where smaller mitochondrial effective population size reduces the ability of selection to purge these large duplications (the theory behind changes in mitochondrial N_e for parthenogenetic lineages is explained further in Paland and Lynch (2006)). More generally, such duplications have occurred frequently in multiple asexual, but not sexual, lineages across squamates (Fujita, Boore and Moritz 2007). Second, in line with Paland and Lynch's result in *Daphnia* (2006), we find a significantly higher D_n/D_s in two independent parthenogenetic lineages compared to sexual lineages (Table 1). This pattern exists both in a thoroughly sampled dataset of 215 sequences of the mitochondrial gene NAD2 and in a large concatenated alignment of 10 genes sampled in 13 individuals (see the Methods section). These two lines of evidence confirm that large genomic duplications are restricted to lineages where the efficiency of selection has been reduced.

To explore the generality of the relationship between mitochondrial duplications and selective efficacy, we analyzed mitochondrial genomes from a separate, distantly related tetrapod group for which appropriate data were available: sexually reproducing mantellid frogs of Madagascar. As in *H. binoei*, duplications tend to co-occur with decreases in selection efficiency, in five branches along the mantellid phylogeny (Fig. 1C). Across the mantellid phylogeny, we find a higher D_n/D_s on branches where a duplication is predicted to have occurred (based on the scenario outlined in Kurabayashi et al. 2008) than on branches where no duplication is predicted to have occurred (Table 1; Fig. 1C). The genes included in the analysis were not themselves involved in the duplication events (Kurabayashi et al. 2008), so these increased D_n/D_s cannot be linked to neo- or subfunctionalization events. Additionally, the association between duplications and increased D_n/D_s cannot simply be a byproduct of a change in reproductive mode (e.g., altered mutational or ecological pressures associated with parthenogenesis), because all these mantellids are sexual.

Our analyses show that genomic duplications occur concomitantly with lower estimated selection strength in the mitochondria of two distantly related tetrapod clades. To our knowledge, this is the first demonstration of a link between genome structure evolution in tetrapods and selective efficiency over short evolutionary timescales. These findings support a view of tetrapod genome evolution where major alterations first fix nonadaptively, but may ultimately contribute to an increase in genomic complexity (Lynch and Conery 2003).

Methods

We analyzed mitochondrial genome evolution in clades with sufficient sequence data, large duplications, and low divergence (i.e., D_s is small enough to accurately estimate). Two tetrapod clades were found to meet these criteria. (1) *Heteronotia binoei* is a lizard species complex with multiple chromosomal lineages, including two asexuals (3N1 and 3N2). We used two datasets to investigate sequence evolution. Dataset 1 contains *nad2* gene sequences from 688 individuals, including 320 from the parthenogenetic lineages (3N1: 231, 3N2: 89). These sequences are available from GenBank, or were newly collected according to protocols in Fujita et al. (2010) (new GenBank accession numbers HQ132796-HQ132922). Dataset 2 contains 10 mitochondrial protein-coding genes from 13 individuals, including 10 CA6-type individuals (nine 3N1 and one CA6 sexual; Fujita, Boore and Moritz 2007), one 3N2, and one SM6 sexual. The 3N2 and SM6 sequences were collected according to methodology described in Fujita, Boore, and Moritz (2007), and have been deposited into GenBank (accession numbers HQ153035 and HQ153036). Mitochondrial genomes for sexual individuals have a size of 17 kb, but several parthenogenetic lineages contain large duplications (Fujita,

A

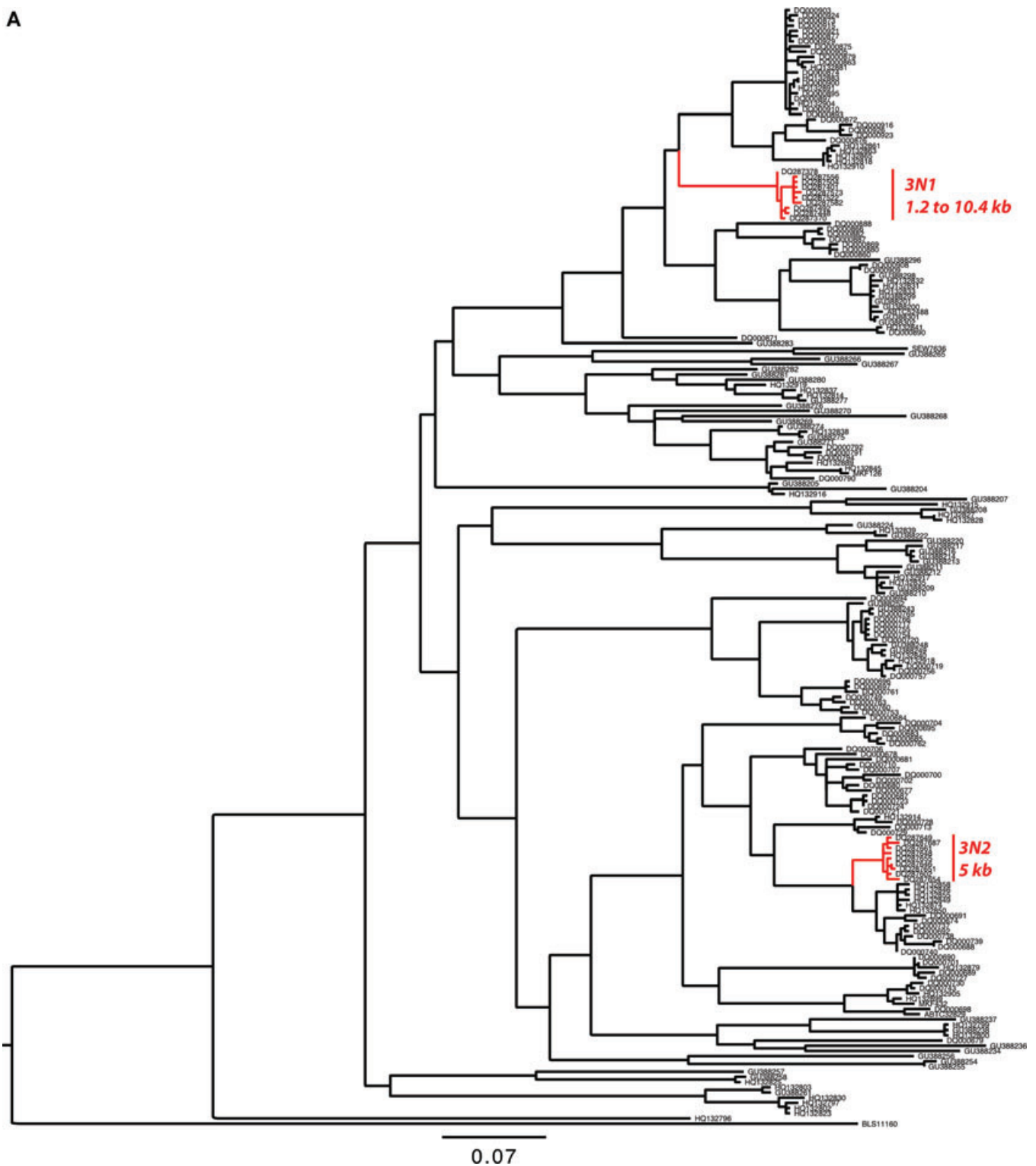


Figure 1. Maximum likelihood phylogenies used in the analyses. (A) *Heteronotia* dataset 1, containing 215 NAD2 sequences. Sizes of duplicated regions are indicated after sequence names where duplications have been found. (B) *Heteronotia* dataset 2, from an alignment of 10 mitochondrial protein-coding genes. Sizes of duplicated regions are indicated after sequence names where duplications have been found. (C) Mantellid frogs, with branches labeled with the estimated sizes of inferred duplication events, in bases, based on the scenario from Kurabayashi et al. (2008). Red branches correspond to asexual lineages (A and B) or predicted duplication events (C).

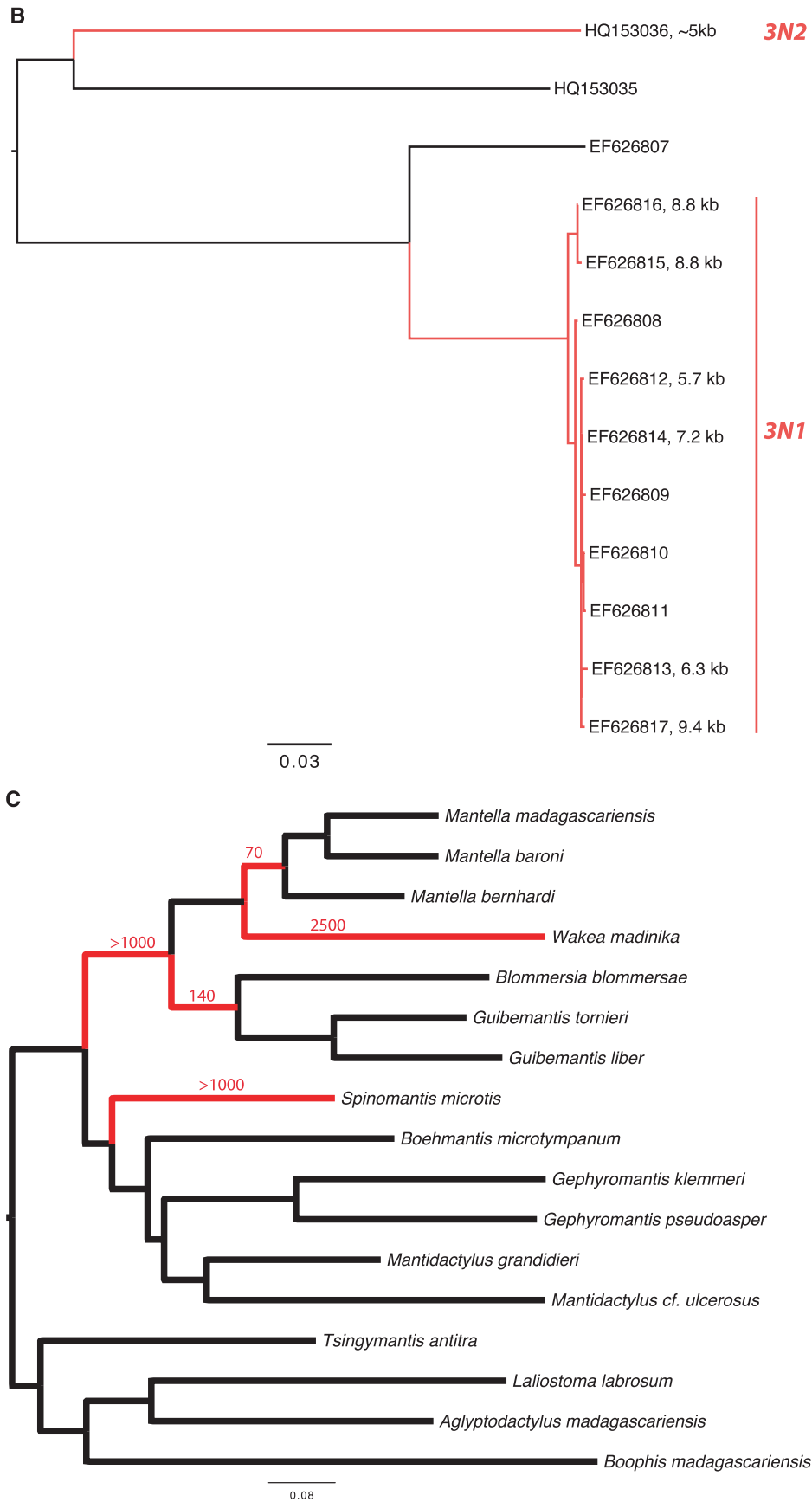


Figure 1. Continued.

Table 1. Datasets and results of the tests for relaxation of selection on focal branches. The first two rows correspond to two different sets of *Heteronotia binoei* sequences. The first is an alignment of *nad2* sequences from 215 individuals, whereas the second is an alignment of 10 concatenated gene sequences from 13 individuals.

Taxon	Number of species	Number of individuals	Number of genes	Log Likelihood (M0) ¹	Log Likelihood (M1) ²	LRT <i>P</i> -value	D_n/D_s (M1) ³
<i>H. binoei</i>	1	215	1	−18323	−18320	0.02	0.18 and 0.35
<i>H. binoei</i>	1	13	10	−22724	−22717	9×10^{-5}	0.09 and 0.14
Mantellid Frogs	17	17	4	−47045	−47040	5×10^{-4}	0.10 and 0.13

¹All branches constrained to evolve under one D_n/D_s .

²One D_n/D_s for sexual lineages (*Heteronotia binoei*) or for branches without duplications (mantellids), and another for asexual lineages (*H. binoei*) or for branches where a duplication is predicted to have occurred (mantellids). Model M1 has one more parameter than model M0.

³In each case, D_n/D_s is significantly higher in asexual lineages or on branches with duplications.

Boore, and Moritz 2007). More specifically, all 3N2 parthenogens possess a duplication of approximately 5 kb in length (Zevering et al. 1991), whereas only some lineages of 3N1 experienced independent and recent duplications that range in size from 1.2 to 10.4 kb (Moritz 1991; Fujita et al. 2007). In certain cases, one copy of a duplicated gene has become a pseudogene (Fujita, Boore, and Moritz 2007). When pseudogenization was detected, we only used the functional copy. (2) Mantellid frog sequence data used in this study come from a group of 17 closely related species, with genome sizes ranging between 9.5 kb and 19 kb (median 13 kb, Kurabayashi et al., 2008). Available mitochondrial sequences were downloaded using accession numbers provided by Kurabayashi et al. (2008), who also provide great detail on putative duplication events across sampled mantellids. Protein-coding DNA sequences of cytochrome *b* and *nad* subunits 1, 2, and 5 were extracted and clustered as orthologs. In both *Heteronotia* and mantellid frogs, duplications are up to several kilobases and involve several genes, which can include tRNA, rRNA, and protein-coding genes, as well as the control region in the frogs.

Sequences were translated and aligned based on amino-acids in Seaview (Gouy, Guindon and Gascuel, 2010) using the muscle algorithm (Edgar 2004) with default parameters. For multi-gene datasets, individual gene alignments were then concatenated. Identical sequences were removed from dataset 1 using bppPhySamp (Dutheil and Boussau 2008). A total of 215 sequences were kept for subsequent analyses, including 10 from 3N1 and 9 from 3N2.

Maximum likelihood (ML) phylogenetic trees were inferred from nucleotide sequences with PhyML (Guindon and Gascuel 2003), using a GTR + I + Γ model. The inferred ML topology for mantellid frogs was congruent with the published estimate (Kurabayashi et al. 2008). Trees and alignments are available from Treebase (<http://purl.org/phylo/treebase/phyloWS/study/TB2:S11392>).

Ratios of nonsynonymous to synonymous substitution rates (D_n/D_s) were estimated using bppml (Dutheil and Boussau 2008).

Topologies were fixed (Fig. 1), but branch lengths and D_n/D_s were reestimated by bppml. Two analyses were performed per topology with either one (M0) or two (M1) free D_n/D_s across branches. For *H. binoei*, different ratios corresponded to asexual and sexual branches. For mantellids, different ratios corresponded to branches with, and without predicted gene duplications, according to the scenario described by Kurabayashi et al. (2008).

ACKNOWLEDGMENTS

We thank J. Bull, C. Moritz, the associate editor, and two anonymous reviewers for helpful comments on the manuscript. BB was supported by a postdoctoral fellowship from the Human Frontier Science Program and the CNRS. JMB and MKF were supported by National Science Foundation postdoctoral research fellowships in biology (Award Nos. DBI-0905867 and DBI-0905714, respectively).

LITERATURE CITED

- Dutheil, J., and B. Boussau. 2008. Non-homogeneous models of sequence evolution in the Bio++ suite of libraries and programs. *BMC Evol. Biol.* 8:255.
- Edgar, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* 5:113.
- Fujita, M. K., J. L. Boore, and C. Moritz. 2007. Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (*Heteronotia binoei*; Squamata, Gekkonidae). *Mol. Biol. Evol.* 24:2775–2786.
- Fujita, M. K., J. A. McGuire, S. Donnellan, and C. Moritz. 2010. Diversification and persistence at the arid-monsoonal interface: Australia-wide biogeography of the Bynoe's gecko (*Heteronotia binoei*; Gekkonidae). *Evolution* 64:2293–2314.
- Gouy, M., S. Guindon, and O. Gascuel. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27:221–224.
- Gregory, T. R., and J. D. S. Witt. 2008. Population size and genome size in fishes: a closer look. *Genome* 51:309–313.
- Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696–704.
- Kuo, C. H., and H. Ochman. 2009. Deletional bias across the three domains of life. *Genome Biol. Evol.* 1:145–152.

- Kuo, C. H., N. A. Moran, and H. Ochman. 2009. The consequences of genetic drift for bacterial genome complexity. *Genome Res.* 19:1450–1454.
- Kurabayashi, A., M. Sumida, H. Yonekawa, F. Glaw, M. Vences, and M. Hasegawa. 2008. Phylogeny, recombination, and mechanisms of step-wise mitochondrial genome reorganization in mantellid frogs from Madagascar. *Mol. Biol. Evol.* 25:874–891.
- Lynch, M., and J. S. Conery. 2003. The origins of genome complexity. *Science* 302:1401–1404.
- Moritz, C. 1991. Evolutionary dynamics of mitochondrial DNA duplications in parthenogenetic geckos, *Heteronotia binoei*. *Genetics* 129:221–230.
- Mueller, R. L., and J. L. Boore. 2005. Molecular mechanism of extensive mitochondrial gene rearrangement in plethodontid salamanders. *Mol. Biol. Evol.* 22:2104–2112.
- Paland, S., and M. Lynch. 2006. Transitions to asexuality result in excess amino acid substitutions. *Science* 311:990–992.
- Whitney, K. D., E. J. Baack, J. L. Hamrick, M. J. W. Godt, B. C. Barringer, M. D. Bennett, C. G. Eckert, C. Goodwillie, S. Kalisz, I. J. Leitch, et al. 2010. A role for nonadaptive processes in plant genome size evolution? *Evolution*. 64:2097–2109.
- Whitney, K. D., and T. Garland, Jr. 2010. Did genetic drift drive increases in genome complexity? *PLoS Genet.* 6:e1001080.
- Zevering, C.E., C. Moritz, A. Heideman, and R. A. Sturm. 1991. Parallel origins of duplications and the formation of pseudogenes in mitochondrial DNA from parthenogenetic lizards (*Heteronotia binoei*; Gekkonidae). *J. Mol. Evol.* 33:431–441.

Associate Editor: H. Innan

Supplementary Material

The following supporting information is available for this article:

Table S1. A list of Genbank accession numbers for *Heteronotia* lizard sequences used in the analyses.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.