Inverted Replication of Vertebrate Mitochondria

Miguel M. Fonseca,† David Posada,‡ and D. James Harris*†

*CiBiO/UP, Centro de Investigação em Biodiversidade e Recursos Genéticos, Vairão, Portugal; †Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal; and ‡Departamento de Bioquímica, Genética e Imunologia, Faculdade de Biologia, Universidade de Vigo, Vigo, Spain

After analyzing the base composition asymmetry of coding regions in vertebrate mitochondria, we identified 2 fishes, Albula glossodonta and Bathygadus antrodes, with inverted compositional patterns. Both species appear to have an unusual control region (CR), and in B. antrodes, it has switched from the light strand to the heavy strand. To our knowledge, this is the first report in vertebrates of inverted mitochondrial replication, caused by an inversion of the CR. These findings support the strand-asymmetric model of mtDNA replication and suggest that vertebrate mtDNA can tolerate globally reversed mutational pressures. In addition, we propose that nucleotide bias is not strand specific but that it depends on the location of the CR.

The organization and evolution of the mitochondrial DNA (mtDNA) are strongly related to its replication. Two models have been proposed to explain this process (Falkenberg et al. 2007). The “strand-asymmetric model” proposes that replication is unidirectional, starting from the origin of heavy strand replication located in the noncoding control region (CR) and displacing the parental heavy strand (H-strand), which becomes temporarily single-stranded DNA (ssDNA) (Clayton 1982). When the H-strand synthesis reaches two-thirds of the genome, it exposes the origin of light strand replication (O_L), initiating the synthesis of a new light strand (L-strand) in the opposite direction. On the other hand, the “strand-coupled model” proposes that mtDNA is synthesized from multiple, bidirectional origins (Reyes et al. 2005). This model describes ribonucleotide incorporation on the L-strand, acting as replication intermediates. The replacement of this RNA by DNA may initiate at dispersed sites, or around the O_L, or near the origin of the leading strand replication (Yasukawa et al. 2006). Importantly, although both models predict a delay between the synthesis of both strands, only the asymmetric model predicts a significant single-stranded exposure of the H-strand.

Although there is still no consensus about which of these models is correct, several attributes of mtDNA are only predicted by the strand-asymmetric model. Within these, the most important is the strand-specific bias (skew) in nucleotide composition (i.e., the H-strand is rich in G and T, whereas the L-strand is rich in A and C), explained by the exposition of the H-strand as ssDNA. The ssDNA is more prone to mutation/damage, especially to deaminations C → T and A → G (Frederico et al. 1990; Lindahl 1993). This different exposure to chemical damage between strands is confirmed by the strand-specific substitution patterns observed at 4-fold redundant sites, which should closely reflect the underlying patterns of mutation (Faith and Pollock 2003; Fonseca et al. 2006). In addition, the fact that redundant sites show the strongest bias suggests that selection may have a minor influence in the differences in nucleotide composition. On the other hand, it could be argued that this bias results mainly from errors that occur during transcription and not during replication. In transcription, both strands become ssDNA, but deamination should occur more often in the L-strand (the “least-transcribed or “nontranscribed” strand). However, the observed mutational pattern is the opposite (Faith and Pollock 2003). Moreover, the asymmetrical model predicts the observed positive correlation between compositional bias and distance to the O_L (Reyes et al. 1998; Faith and Pollock 2003; Broughton and Reneau 2006). During asymmetrical replication, the parental H-strand becomes ssDNA while the nascent H-strand is synthesized, and genes closer to the O_L, in the direction of the L-strand synthesis, should remain exposed to mutation for less time.

Another important feature of the mtDNA is the CR, which regulates transcription and the initiation of replication. In most vertebrates and some invertebrates, the CR contains 3 conserved sequence blocks (CSB I, CSB II, and CSB III), which may be involved in the initiation of replication (Cao et al. 2004; Falkenberg et al. 2007). In mammalian mtDNA, the majority of the 5′ ends of the nascent H-strand are located within or immediately downstream of CSB I (Clayton 1991), whereas CSB II increases the stability of the H-strand synthesis initiation (Xu and Clayton 1995). However, the absence of CSBs or their lack of conservation in some taxa suggests that other sequences may function as alternative regulatory elements. The location of the CSBs determines the leading strand during replication. Typically, CSBs are located in the L-strand, implying that the H-strand is the leading strand. Importantly, if CSBs switch strands, the L-strand becomes the leading strand and therefore the paternal L-strand will be temporarily single stranded and exposed to DNA damage during replication. Thus, the mutational scenario between strands will switch, and given enough time, the compositional bias and gradient may also revert, especially at synonymous positions.

In invertebrates, CR inversion has been proposed to be the cause of the reversion of the strand bias at 4-fold redundant sites in coding mtDNA (Scouaras et al. 2004; Hassanin et al. 2005; Kilpert and Podsiadlowski 2006). In vertebrates, despite representing 80% of the available mtDNA sequences, CR inversions have never been detected. However, the increasing number of gene arrangements (San Mauro et al. 2006, Fonseca and Harris 2007), partial genome duplications (Fujita et al. 2007), and gene inversions (Amer and Kumazawa 2007) reported, suggest that vertebrate mtDNA should not be much constrained.

Key words: mtDNA, control region, inversion, strand bias, replication mechanism, vertebrates.

E-mail: mmfonseca@mail.icav.up.pt.

doi:10.1093/molbev/msn050
Advance Access publication February 21, 2008

© The Author 2008. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org
Here, we report the first evidence for reversed mtDNA replication in vertebrates. We downloaded all 748 complete vertebrate mtDNA genomes available in GenBank (http://www.ncbi.nlm.nih.gov) in June 2007 and calculated the GC skew \( \frac{G}{C} = \frac{G}{C + C} \) and the AT skew \( \frac{A}{C} = \frac{A}{A + T} \) (Perna and Kocher 1995) at 4-fold redundant sites using BACA (Antão et al. 2007). The redundant codons examined were alanine (GCN), proline (CCN), serine (TCN), threonine (ACN), arginine (CGN), glycine (GGN), leucine (CTN), and valine (GTN). Two Teleostei fishes, *Albula glossodonta* and *Bathygadus antrodes*, showed significant and independent reversion of the GC and AT skews (figs. 1 and 2; if the complete genome constraint is relaxed and if all the complete coding segment for the mitochondrial genes are considered [multiplying sample size by 20], *A. glossodonta* and *B. antrodes* are still the 2 clear outliers [anonymous reviewer personal communication]). In both species, the CR presented atypical features. In *A. glossodonta*, the CR appears to be nonfunctional, as it is relatively short, very dissimilar from other CRs, and it has no CSBs, although it can form secondary structures. In addition to showing reversed skews and from the fact that the CR reverse complement sequence is more similar to a typical CR and may contain a putative CSB II (AACCCCCACTCCTTTCC), we did not find any additional features that could support its inversion. On the other hand, we found strong evidence for a CR inversion in *B. antrodes*. In this species, an H-strand noncoding fragment (flanked by 2 L-tRNAs, fig. 3), originally described as an ND1-pseudogene (Satoh et al. 2006), is in fact the main CR: 1) it has significant sequence similarity with other CRs of closely related species (fig. 4), 2) it is capable of forming secondary structures, 3) it contains CSBs I and II in the same relative position (fig. 2; Saccone et al. 1991; Lee et al. 1995), and 4) the fragment located in the standard CR location is short and it appears to have no origin of replication (it does not have CSBs).

Although, ultimate proof can only be achieved with in vitro experiments, to the best of our knowledge, this constitutes the first known example in vertebrates of inversion of the mitochondrial replication mechanism, through an inversion of the CR.

These findings provide further support for the asymmetric model of replication and may be important to decipher, which conserved motifs of the CR are in fact crucial for an efficient replication/transcription. They also suggest that it is not necessarily correct to affirm that gene skew is strand specific. In these 2 examples, genes reverted their skew without changing the coding strand, and in *B. antrodes* only the CR inverted. Instead, we hypothesize that compositional bias depends much on which strand the CR (first origin of replication) is located. In addition, our results suggest that even vertebrate mtDNA may experience extreme structural changes and tolerate reversed mutational pressures, which may strongly influence the codon usage and amino acid composition (Min and Hickey 2007).
Indeed, it is important for accurate phylogenetic estimation to know the nucleotide bias across different taxa, as it has been shown that unrelated taxa with similar nucleotide composition tend to spuriously group together (Rosenberg and Kumar 2003; Hassanin et al. 2005).

Acknowledgments

We acknowledge R. Faria, R. Vasconcelos, and the anonymous reviewer for their helpful comments and suggestions. M.M.F. and D.J.H. are supported by FCT (Fundação para a Ciência e a Tecnologia) grants SFRH/BD/31158/2006 and POCI/65745/2006, respectively. D.P. is funded by the Spanish Ministry of Education and Science.

Literature Cited


Cao L, Kenchington E, Zouroso E, Rodakis GC. 2004. Evidence that the large noncoding sequence is the main control region of maternally and paternally transmitted mitochondrial genomes of the marine mussel (Mytilus spp.). Genetics. 167:835–850.


Kilpert F, Podsiadlowski L. 2006. The complete mitochondrial genome of the common sea slater, Ligia oceanica (Crustacea,
Isopoda) bears a novel gene order and unusual control region features. BMC Genomics. 7:241.


Franz Lang, Associate Editor
Accepted February 15, 2008