Note

Multiple Events Are Responsible for an Insertion in a Paternally Inherited Mitochondrial Genome of the Mussel *Mytilus galloprovincialis*

Athanasia Mizi,* Eleftherios Zouros† and George C. Rodakis*,1

*Department of Biochemistry and Molecular Biology, National and Kapodistrian University of Athens, Panepistimioupolis, 15701 Athens, Greece and †Department of Biology, University of Crete, 71409 Heraklion, Crete, Greece

ABSTRACT

In a sperm-transmitted mtDNA of *Mytilus galloprovincialis* we found an insertion that is not present in the typical genome and whose origin can be explained by a sequence of three events: a tandem duplication, a nonhomologous recombination, and a deletion. Unless such events are extremely rare in this species, the identical gene arrangement of the two gender-specific genomes should imply strong selection for same gene order and size.

SPECIES of the mussel family Mytilidae have two mtDNAs, one transmitted through the egg (the F genome) and one transmitted through the sperm (the M genome) (SKIBINSKI *et al.* 1994a,b; ZOUROS *et al.* 1994a,b). The system is known as doubly uniparental inheritance (DUI). In a previous study we published the complete sequence of the two genomes (MIZI *et al.* 2005). The sequences differ by 21.5%, yet their gene contents and arrangements are identical and their lengths differ only by 118 bp. The specific M genome we sequenced contained a 1045-bp insertion that was not found in 19 other genomes that were examined for this purpose (MIZI *et al.* 2005). The insertion consisted of the first 11 bp of *tRNA^{Asn}* and the last 56 bp of *tRNA^{Gly}* separated by 978 bp that contained a large part of the main control region, a copy of the *tRNA^{Tyr}*, and the first 88 bp of *Cyt b* (Figure 1B, *d*). The typical M sequence, the sequence of the insertion, and the insertion's position in the typical M genome can be retrieved from GenBank (accession no. AY363687, bases 14968–16012). Here we trace the molecular mechanism that might have produced this insertion and show that no less than three events were necessary for it.

The first event involved a duplication (Figure 1A) in a typical genome (*a*) that produced a *tRNA^{Gly}* + *tRNA^{Asn}* tandem repeat (genome *b*). Tandem repeats are usually caused by replication slippage mediated by short repeated sequences (LEVINSON and GUTMAN 1987). MACEY *et al.*

(1997) suggested that the asymmetrical replication of mtDNA (CLAYTON 1982) facilitates tRNA duplication by exposing the parental heavy strand to a single-strand state and allowing it to form stem-and-loop structures. When DNA polymerase encounters such structures, it may become detached from one and reattached to another, thus causing tRNA duplications or deletions in the nascent light strand.

The second event was a recombination (Figure 1B) that involved two genomes (*a'* and *b'*), was nonhomologous, and had two crossing-over points. The first crossing over occurred between the control region (CR) of *a'* and the first *tRNA^{Asn}* copy of *b'* and the second between the *Cyt b* of *a'* and the second *tRNA^{Gly}* copy of *b'*. One recombinant genome (*c*) lacked the *tRNA^{Tyr}* gene, had a truncated CR and a truncated *Cyt b* gene, and was eliminated as nonfunctional. The other (*d*) was the genome with the 1045-bp insertion. Alignment of *a'*, *b'*, and *d* can identify the exact position of each crossing over and the sequence similarity between *a'* and *b'* that might have mediated the crossing-over events (Figure 1C). As expected, the recombinant points were found in the CR and the *Cyt b* of *a'* and in the N1 and G2 of *b'*. An interesting observation is that the switch points are not single nucleotides, but rather small sequences (seven and nine nucleotides, respectively). That these identical motifs were found exactly at the switch points cannot be a coincidence, but it must signify their involvement in the catalysis of recombination.

In three of the four exchanges, the switch from one parental sequence to the other occurs immediately after the common oligonucleotide motif (Figure 1C). The

¹*Corresponding author:* Department of Biochemistry and Molecular Biology, National and Kapodistrian University of Athens, Panepistimioupolis, 15701 Athens, Greece.   E-mail: grodakis@biol.uoa.gr
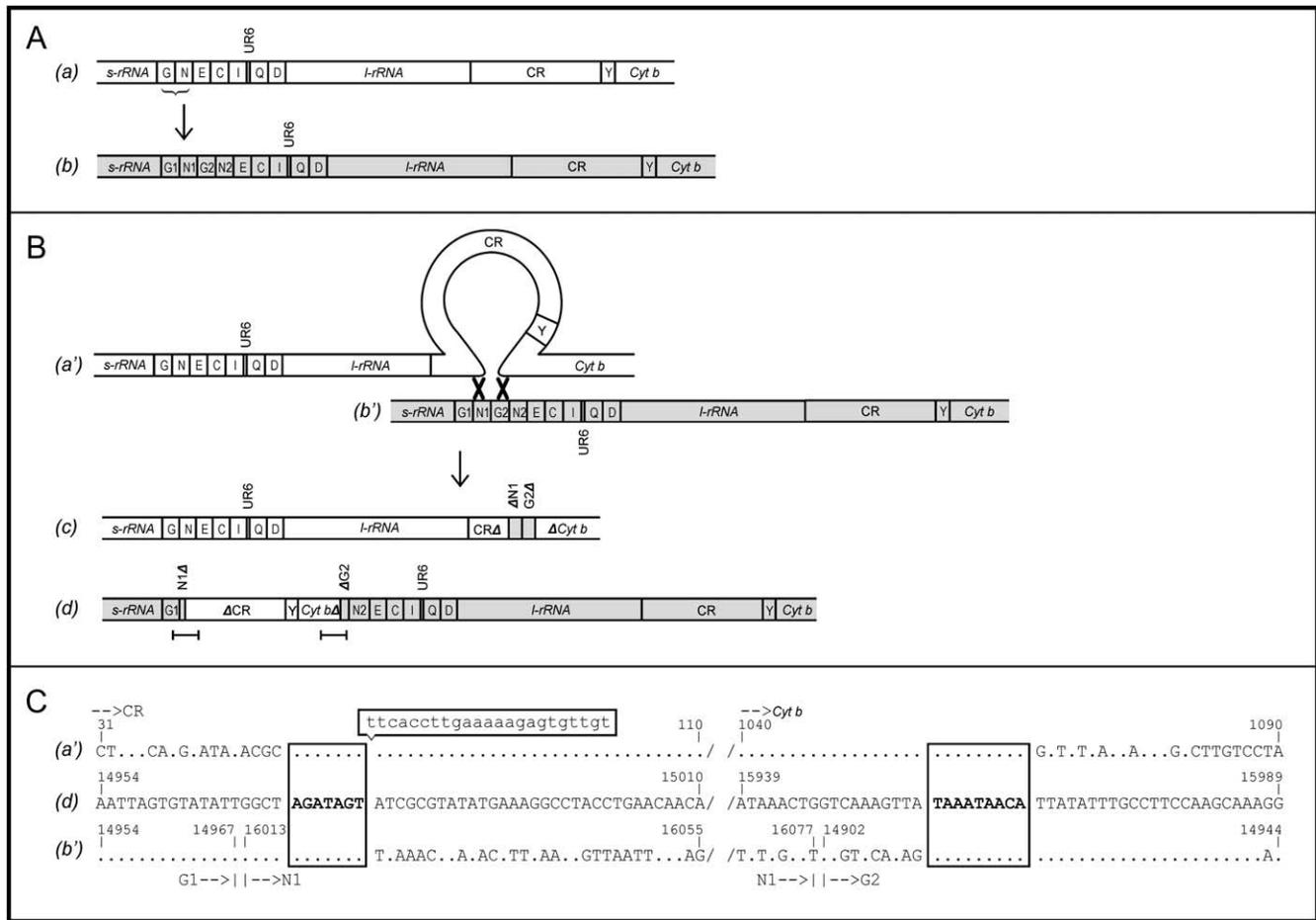
FIGURE 1.—The three events that produced the insertion. tRNA genes are denoted with the single-letter amino acid code; CR is the major control region; unassigned region 6 (UR6) is a 16-bp noncoding sequence; and *s-rRNA*, *l-rRNA*, and *Cyt b* stand for small rRNA, large rRNA, and cytochrome-*b* genes. (A) Tandem duplication. *a* is the standard gene order; *b* is the gene order resulting from tandem duplication of two tRNA genes marked by a brace in *a*. (B) Double-point recombination. The parental genomes are denoted as *a′* and *b′* to indicate that they may not be identical to *a* and *b*, Δ denotes deletion at the 5′ or 3′ region of a gene, open and shaded parts indicate parts from different genomes, genome *c* lacks a complete *tRNA^{Tyr}* and *Cyt b* and is assumed to be nonviable, genome *d* is the gene order of the examined genome, and bars mark the regions whose sequences are shown in C. (C) Deletion. *a′*, *d*, and *b′* are the nucleotide sequences of the respective genomes around the points of recombination; dots denote nucleotide identity with the sequence of the recombinant genome; boxes with boldface type mark motifs shared by all sequences; boxed lowercase nucleotides represent the sequence that is present in the full CR, but is missing from ΔCR of *d*; and numbers refer to nucleotide positions in the complete sequence of the M genome of *M. galloprovincialis*.

exception is in the first crossing over, where identity of *d* with *a′* occurs 23 bp downstream from the common motif. These 23 nucleotides are present in the complete CR but absent from the ΔCR of *d*. Cao *et al.* (2004) divided the CR of the Mytilus mtDNA into three domains: variable domain 1 (VD1), conserved domain (CD), and variable domain 2 (VD2). The 23-nucleotide stretch is part of VD1, which in the M genome is characterized by dispersed indels (Burzynski *et al.* 2003; Cao *et al.* 2004). The 23-bp stretch might not have been present in the CR of *a′*, might have been excised during recombination, or might have been removed from ΔCR after genome *d* was produced.

Several arguments can be offered in support of the proposed scenario for the insertion. The presence in the same genome of a complete and a partial copy of

tRNA^{Asn} and *tRNA^{Gly}* could be the result of recombination between two typical genomes. However, no single recombination event, with either one or two crossing-over points, between two such genomes could explain the order in which partial and complete tRNA sequences were found. The tandem duplication solves this problem. It could be produced by the mechanism suggested by Macey *et al.* (1997), as it involved two tRNAs located next to each other with no intervening sequences. The two-point recombination explains the nucleotide sequence similarities and differences between the two extant genomes (*a* and *d*). The required state of heteroplasmy for two different M genomes might have arisen in two ways. Genome *b* might have survived in the population and came into heteroplasmy—now as *b′*—with *a′* in some future generation. This would

require co-inheritance of two M genomes, which we do not know how common might be in species with DUI. Alternatively, it is possible that genomes *a* and *b* recombined immediately after the duplication and that the 23-bp deletion occurred during recombination or afterward. This would eliminate the need to account for the presence of two different M genomes in the same mitochondrion. Less doubtful is the possibility of recombination, which has been demonstrated in male mussels (LADOUKAKIS and ZOUROS 2001). In conclusion, the scheme we propose makes use of events that are known to occur in the mussel mtDNA and is the most parsimonious with regard to the number of required steps.

All putative functional elements of the main control region (CAO *et al.* 2004) are located downstream of the 23-bp deletion and are present in ΔCR. The mtDNA replication and transcription mechanism in mollusks, as in the vast majority of invertebrates, is not known in sufficient detail to allow one to speculate how it would be affected by the presence of two putatively functional control regions. The insertion includes only one complete gene (*tRNA^{Tyr}*) that may be functional. There are only six nucleotide differences between the entire 1045 bp of the insertion and the corresponding duplicated parts in the genome (genome *d*): one between ΔG2 and G, one between *Cyt b*Δ and *Cyt b*, and four between ΔCR and the conserved central domain of the CR. Kimura's two-parameter *K*-distance between the typical (*a*) and the insertion-bearing genome (*d*) is 0.005 (SE = 0.002). For the CR alone this distance is also 0.005 and can be compared to 0.087 (SE = 0.009) that is obtained from pairwise comparisons of the three typical *M. galloprovincialis* genomes whose sequences at the compared region are known (GenBank accession nos. AY350793, AY350794, and AF188280). Assuming that the accumulation of nucleotide substitutions is proportional to the time of divergence, we conclude that the insertion is much younger than the average time of separation of two randomly chosen M genomes.

Multiple events of the type we describe here must be fairly common, but are detected mostly as fixed among-taxa differences (EBERHARD *et al.* 2001; LAVROV *et al.* 2002 and references therein) and can, therefore, serve as landmarks for inferring deep branches of animal phylogeny (BOORE 1999). Cases of standing intraspecific variation due to such events are also known, but appear to be rare (MORITZ and BROWN 1987; GACH and BROWN 1997). It is clear that the animal mitochondrial genome is not protected from events that can disrupt its gene order and insert nonfunctional sequences. These events may be even more common between gender-specific genomes, if judged from the high level of recombination in male mussel gonads (LADOUKAKIS and ZOUROS 2001). The insertion we describe here is an example of such an event. Nothing can be said about its ultimate fate given its apparently recent origin and the lack of information about its frequency in natural populations. Yet its mere detection makes it more remarkable that the typical egg-transmitted and sperm-transmitted mitochondrial genomes of this species have maintained identical gene arrangements and differ only by ∼100 bp in total length, even though they have diverged by >20% in primary sequence. It is difficult to put an estimate on the time of separation of these genomes because of their different rates of evolution (SKIBINSKI *et al.* 1994b; ZOUROS *et al.* 1994b) and the possibility of recurrent invasions of the M lineage by the F that may reset the amount of divergence to zero (HOEH *et al.* 1997). Regardless of this, sequence variation among various F molecules and among various M molecules (*e.g.*, LADOUKAKIS *et al.* 2002) is large enough to suggest that the F/M separation has been sufficiently long for events like the one we describe here to occur multiply and cause extensive differences in gene order and size. That this has not happened implies that selection has maintained the same gene organization in the two conspecific genomes of *M. galloprovincialis*.

LITERATURE CITED

BOORE, J. L., 1999 Animal mitochondrial genomes. Nucleic Acids Res. **27:** 1767–1780.

BURZYNSKI, A., M. ZBAWICKA, D. O. SKIBINSKI and R. WENNE, 2003 Evidence for recombination of mtDNA in the marine mussel *Mytilus trossulus* from the Baltic. Mol. Biol. Evol. **20:** 388–392.

CAO, L., E. KENCHINGTON, E. ZOUROS and G. C. RODAKIS, 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.

CLAYTON, D. A., 1982 Replication of animal mitochondrial DNA. Cell **28:** 693–705.

EBERHARD, J. R., T. F. WRIGHT and E. BERMINGHAM, 2001 Duplication and concerted evolution of the mitochondrial control region in the parrot genus *Amazona*. Mol. Biol. Evol. **18:** 1330–1342.

GACH, M., and W. M. BROWN, 1997 Characteristics and distribution of large tandem duplications in brook stickleback (*Culaea inconstans*) mitochondrial DNA. Genetics **145:** 383–394.

HOEH, W. R., D. T. STEWART, C. SAAVEDRA, B. W. SUTHERLAND and E. ZOUROS, 1997 Phylogenetic evidence for role-reversals of gender-associated mitochondrial DNA in *Mytilus* (Bivalvia: Mytilidae). Mol. Biol. Evol. **14:** 959–967.

LADOUKAKIS, E. D., and E. ZOUROS, 2001 Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. Mol. Biol. Evol. **18:** 1168–1175.

LADOUKAKIS, E. D., C. SAAVEDRA, A. MAGOULAS and E. ZOUROS, 2002 Mitochondrial DNA variation in a species with two mitochondrial genomes: the case of *Mytilus galloprovincialis* from the Atlantic, the Mediterranean and the Black Sea. Mol. Ecol. **11:** 755–769.

LAVROV, D. V., J. L. BOORE and W. M. BROWN, 2002 Complete mtDNA sequences of two Millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. Mol. Biol. Evol. **19:** 163–169.

LEVINSON, G., and G. A. GUTMAN, 1987 High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. Nucleic Acids Res. **15:** 5323–5338.

Macey, J. R., A. Larson, N. B. Ananjeva, Z. Fang and T. J. Papenfuss, 1997   Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol. Biol. Evol. **14:** 91–104.

Moritz, C., and W. M. Brown, 1987   Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. Proc. Natl. Acad. Sci. USA **88:** 7183–7187.

Mizi, A., E. Zouros, N. Moschonas and G. C. Rodakis, 2005   The complete maternal and paternal mitochondrial genomes of the Mediterranean mussel *Mytilus galloprovincialis*: implications for the doubly uniparental inheritance mode of mtDNA. Mol. Biol. Evol. **22:** 952–967.

Skibinski, D. O. F., C. Gallagher and C. M. Beynon, 1994a   Mitochondrial DNA inheritance. Nature **368:** 817–818.

Skibinski, D. O. F., C. Gallagher and C. M. Beynon, 1994b   Sex limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. Genetics **138:** 801–809.

Zouros, E., A. O. Ball, C. Saavedra and K. R. Freeman, 1994a   Mitochondrial DNA inheritance. Nature **368:** 818.

Zouros, E., A. O. Ball, C. Saavedra and K. R. Freeman, 1994b   An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. Proc. Natl. Acad. Sci. USA **91:** 7463–7467.

Communicating editor: M. Aguadé