Homologous Recombination between Highly Diverged Mitochondrial Sequences: Examples from Maternally and Paternally Transmitted Genomes

Emmanuel D. Ladoukakis,* Ioannis Theologidis,1 George C. Rodakis,2 and Eleftherios Zouros1
1Department of Biology, University of Crete, Iraklio, Greece
2Department of Biochemistry and Molecular Biology, National and Kapodistrian University of Athens, Athens, Greece

*Corresponding author: E-mail: ladoukakis@biology.uoc.gr
Associate editor: Dr. Richard Thomas

Abstract
Homologous recombination is restricted to sequences of low divergence. This is attributed to the mismatch repairing system (MMR), which does not allow recombination between sequences that are highly divergent. This acts as a safeguard against recombination between nonhomologous sequences that could result in genome imbalance. Here, we report recombination between maternal and paternal mitochondrial genomes of the sea mussel, whose sequences differ by >20%. We propose that the strict maternal inheritance of the animal mitochondrial DNA and the ensuing homoplasy has relieved the MMR system of the animal mitochondrial DNA from the pressure to tolerate recombination only among sequences with a high degree of similarity.

Key words: mtDNA recombination, sequence divergence, doubly uniparental mtDNA inheritance.

Introduction
Homologous DNA recombination is widespread in all kingdoms of life. Even though from an evolutionary perspective the importance of recombination remains largely unresolved, its participation in several important functions is unquestionable: It maintains genome stability, promotes genetic diversity by generating new gene combinations (Neale and Keeney 2006), facilitates natural selection in purging deleterious mutations, and increases effective population size (Charlesworth 2009). Genome stability, in particular, is served by assuring accurate repair of double-strand breaks of DNA (West 2003) and by allowing the correct DNA partners to recombine. Recombination between nonhomologous chromosomal regions may result in chromosomal rearrangements with deleterious consequences in most cases (Opperman et al. 2004). The need of high sequence similarity between the recombination substrates is, therefore, assumed to act as a means for preventing recombination between nonhomologous regions. Indeed, the recombination rate between two sequences falls rapidly as their degree of divergence increases (Zawadzki et al. 1995; Datta et al. 1997; Vulic et al. 1999; Opperman et al. 2004; Eppley et al. 2007). This appears to be a general trend from prokaryotes to higher eukaryotes, even though the details of the mechanism may vary among major taxa. Recombination between highly divergent sequences is very rarely observed in nature.

The mismatch DNA repairing (MMR) mechanism, whose primary function is the correction of mismatches occurring during DNA replication, has been implicated in the prevention of recombination between divergent sequences. Antirecombinational activity of MMR has been observed in prokaryotes (Majewski and Cohan 1999; Vulic et al. 1999; Majewski et al. 2000), eukaryotic nuclear DNA (Chen and Jinks-Robertson 1999; Smith et al. 2007; Li et al. 2009) and plant mitochondrial DNA (mtDNA; Shedge et al. 2007). The MMR enzymatic machinery contains several proteins with homologs in all kingdoms of life (Iyer et al. 2006; Li 2008). Defects of MMR proteins may cause a substantial increase in the recombination rate between divergent sequences (Iyer et al. 2006; Li 2008). Besides MMR, sequence divergence itself may mechanically impede the formation of a heteroduplex at the crossover points, thus suppressing recombination (Datta et al. 1997; Majewski and Cohan 1998). Once a heteroduplex is formed, recombination may occur even between highly divergent sequences. For a heteroduplex to be formed, short fragments of perfect homology are needed at one or both crossover points (Shen and Huang 1986; Datta et al. 1997; Majewski and Cohan 1999; Majewski et al. 2000; Opperman et al. 2004). The length of these fragments, known as minimal efficient processing segments (MEPS), is usually >20 bp (Shen and Huang 1986). The role of MEPS has been studied more thoroughly in prokaryotes (Shen and Huang 1986; Zahrt and Maloy 1997; Majewski and Cohan 1999; Majewski et al. 2000; Costechareyre et al. 2009). But short identical fragments similar to MEPS have been also implicated in recombination in yeast (Datta et al. 1997) and Arabidopsis (Opperman et al. 2004). In plants and fungi, mtDNA recombination has been mainly studied in the context of gene transfer and gene rearrangement (for references see Kmiec et al. 2006; Bonnefoy and Fox 2007).

Unlike nuclear and bacterial DNA, recombination in animal mtDNA has not been extensively studied, partly because animal mtDNA recombination was in doubt until 10
years ago. Animal mtDNA recombination is now uncontroversial and includes two lines of evidence. The first line (real-time recombination) comes from the recovery of recombinant sequences from the mtDNA pool of individual organisms together with one or both parental sequences (Ladoukakis and Zouros 2001a; Kraysberg et al. 2004; Zsurka et al. 2005; Burzynski and Smietanka 2009). The second line (historical recombination) comes from the detection of recombinant mtDNA molecules that have been fixed in different yet closely related species, after their divergence from a common ancestor, in which the recombination presumably occurred (Ladoukakis and Zouros 2001b; Piganeeu et al. 2004; Tsousis et al. 2005).

The strict maternal inheritance of animal mtDNA results in inherently homoplasmic individuals (i.e., individuals with a single type of mtDNA). In a pool of identical sequences, recombinant molecules cannot be directly detected because recombinants are identical to parental sequences. An exception occurs in regions that contain repeats, which upon unequal crossing over produce recombinants whose length is different from that of the parental sequences. In contrast to standard maternal inheritance, species with doubly uniparental inheritance (DUI) of mtDNA (Zouros 1997 contrast to standard maternal inheritance, species with unequal crossing over produce recombinants whose length is different from that of the parental sequences. In contrast to standard maternal inheritance, species with doubly uniparental inheritance (DUI) of mtDNA (Zouros et al. 1994a) provide a natural experiment for the detection of mtDNA recombination. The main characteristic of DUI is the stable presence in the same species of two types of mtDNA, one with maternal (known as F) and the other with paternal (known as M) inheritance. The F genome is transmitted to both female and male progeny through the egg. The M genome is transmitted through the sperm’s enlarged mitochondria (in Mytilus the number of mitochondria in a sperm is five or, rarely, six [Longo and Dornfeld 1967; Cogswell et al. 2006]) and has different fates in female and male embryos. In the former, it remains a small minority in the mtDNA pool of the fertilized egg and is either lost stochastically during successive cell divisions or can be detected in small amounts in some tissues of adult females (Stewart et al. 1995; Garrido-Ramos et al. 1998; Dalziel and Stewart 2002). In male embryos, the sperm mitochondria form an aggregate (Cao et al. 2004; Obata and Komaru 2005; Cogswell et al. 2006) that, through a developmental mechanism that remains unknown, enters the primordial germ cells and becomes the sole mtDNA occupant of the male’s germ line. The aggregation of sperm mitochondria is not perfect. One or more mitochondria may fail to join the aggregate and end up in somatic tissues. Thus, males are obligatory heteroplasmic mosaics. Their germ line— and eventually the sperm—contains the M genome (Venetis et al. 2006; Ghiselli et al. 2011), and the somatic tissues are dominated by the F genome with the occasional presence of small amounts of M genome. DUI has been first observed in mussels (Skibinski et al. 1994; Zouros et al. 1994a, 1994b) but is now known to occur in more than 30 bivalve species (Theologidis et al. 2008).

The two mtDNA genomes are highly diverged, from as low as ~20% in some species of sea mussels of the family Mytilidae (Mizi et al. 2005) to about 50% in fresh water mussels of the family Unionidae (Doucet-Beaupre et al. 2010). A twist of DUI, one that has so far been observed only in species of the genus Mytilus, is “masculinization”. It has been observed that, in addition to F and M genomes, populations of these species may contain a third mitochondrial genome, normally in low frequencies (Hoeh et al. 1997; Quesada et al. 1999; Ladoukakis et al. 2002). This genome’s coding sequence resembles that of the F genome, but its control region (the region that contains the elements for the control of replication and transcription) is a mosaic of sequences from the control regions of the F and M genomes (Venetis et al. 2007; Cao and et al. 2009). This genome is paternally inherited and assumed to have resulted from an F genome that reversed transmission route after incorporation of sequences from the M genome at the control region (thus, referred to as “masculinized”; Saavedra et al. 1997; Burzynski et al. 2003, 2006; Venetis et al. 2007). With time, a masculinized genome will diverge from the maternal genome from which it originated. A genome of this type has been fully sequenced (genome C, Venetis et al. 2007) and found to have diverged from the maternal genome by about 3%. This divergence is large enough to allow for the detection of mtDNA recombination in male mussels that have inherited a masculinized genome from their father along with the maternal genome from their mother. This idea was exploited by Ladoukakis and Zouros (2001a), who produced direct evidence of recombination at the cox3 gene in male mussels.

In this paper, we have extended the search for recombination in male mussels and have shown that recombination occurs also between “typical” maternal and paternal genomes, that is, between genomes that differ by about 20%. This result demonstrates that animal mtDNA recombination may involve sequences whose degree of divergence is comparable with the highest known for nuclear and bacterial DNA.

Materials and Methods
We used preserved DNA from animals that were used for a phylogeographic survey of Mytilus galloprovincialis (Ladoukakis et al. 2002). We randomly selected nine male individuals (see table 1 for animal code and locality of origin) from among those in which the DNA was still in good condition and the presence of the typical F and M mtDNA genomes could be detected in their gonadal tissue. This tissue is a mixture of somatic and germ cells, so it contains both F and M genomes. It is, therefore, the most suitable tissue to detect recombination. From these individuals, total DNA was extracted using the salt method as described in Miller et al. (1988) and modified by Ladoukakis et al. (2002). For this study, we used polymerase chain reaction (PCR) to amplify segments from three mitochondrial genes, namely, cox3, cox1, and nad5. For cox3, we used the primers COIII-F (5’-TAT GTA CCA GGT CCA AGT CCG TG-3’) and COIII-REV (5’-ATG CTC TTC TTG AAT ATA AGC GTA CC-3’); for cox1, the primers COI-F (5’-GGT AAR GAT ATA ATT YCCR CGG-3’) and COI-R (5’-TGT GCT ACM ACR TAR TAA GTA TCA T-3’); and for

---

1848
The number of clones checked for each type, and potential recombinants). Positive clones were digested the PCR product from each individual. Restriction profiles from each individual were sequenced from both strands in a PT100 sequencer (MG Research).

Results

Restriction Patterns

We surveyed three mtDNA regions that contained parts of the cox3, nad5, and cox1 genes. In the absence of recombination, clones of PCR products from an individual male are expected to be of two types, one corresponding to the maternal (F) and the other to the paternal (M) genome for each amplified region. New types would result from mutation or recombination. First, we screened the cloned fragments for new types by RFLP (fig. 1). Clones with RFLP patterns different from the parental patterns were considered as putative recombinants. The number of recovered patterns is given in table 1, along with their characterization as parental or recombinant. In total, we screened 433 clones and recovered 11 nonparental restriction patterns in four of the nine individuals: one for the cox3 fragment, one for the nad5, and nine for the cox1. This gives

### Table 1. Different Sequences Recovered among Clones Screened for Recombinants.

<table>
<thead>
<tr>
<th>Animal Code</th>
<th>cox3</th>
<th>nad5</th>
<th>cox1</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2</td>
<td>27/2/0</td>
<td>13/2/0</td>
<td>12/2/0</td>
</tr>
<tr>
<td>H8</td>
<td>24/3/1</td>
<td>13/2/0</td>
<td>11/3/1</td>
</tr>
<tr>
<td>H25</td>
<td>28/2/0</td>
<td>9/1/0</td>
<td>10/2/0</td>
</tr>
<tr>
<td>H28</td>
<td>22/2/0</td>
<td>10/2/0</td>
<td>32/4/5</td>
</tr>
<tr>
<td>H33</td>
<td>20/2/0</td>
<td>12/1/0</td>
<td>10/2/0</td>
</tr>
<tr>
<td>BS29</td>
<td>31/2/0</td>
<td>10/1/0</td>
<td>33/4/3</td>
</tr>
<tr>
<td>BS45</td>
<td>42/2/0</td>
<td>8/1/0</td>
<td>8/2/0</td>
</tr>
<tr>
<td>BS60</td>
<td>8/2/0</td>
<td>8/1/0</td>
<td>8/2/0</td>
</tr>
<tr>
<td>C29</td>
<td>8/2/0</td>
<td>8/3/1</td>
<td>8/2/0</td>
</tr>
<tr>
<td>Total</td>
<td>210/19/1</td>
<td>91/14/1</td>
<td>132/23/9</td>
</tr>
</tbody>
</table>

**NOTE.—** H2, H8, H25, H28, H33: animals from Halastra, Aegean Sea; BS29, BS45, BS60: animals from Sebastopol, Black Sea; C29: animal from Cefalonia island, Ionian Sea. cox3: cytochrome oxidase subunit 3, nad5: NAD dehydrogenase subunit 5, cox1: cytochrome oxidase subunit 1. Numbers: clones scored/different restriction patterns recovered/recombinant sequences recovered.

* Two clones, recovered from two independent attempts, contained the same recombinant RFLP pattern. Both clones were sequenced and produced the same recombinant sequence given in figure 2 as “cox1, H28-12R.”

end

Table 1. Different Sequences Recovered among Clones Screened for Recombinants.
a frequency of 2.5% of recombinant sequences in our sample. However, this figure can be misleading for reasons we explain below.

Parental and Recombinant Sequences and Crossover Points

We sequenced all clones that had restriction profiles different from the profiles of parental molecules, regardless of whether some of these clones produced the same RFLP profile. Also, we sequenced one clone with the parental F and one clone with the parental M profile from each individual. All clones with restriction profiles different from the parental ones turned out to be products of recombination (see below). We may assume that the actual number of recombinant sequences at the surveyed parts of the mtDNA genomes was higher and that a large number of them escaped detection. This could be because of the small number of examined clones or because a recombinant sequence may have had the same RFLP profile as a parental sequence (this will happen if the crossover point lies before the first or after the last restriction site, fig. 1). Such recombinants were likely to escape detection, even if present among our clones, because we sequenced only one clone with the maternal and paternal RFLP profile from each individual. In addition to these experimental limitations, a recombinant may not be detected because of its low representation in the mtDNA pool of the surveyed individuals.

Variable sites of recombinant and parental sequences are given in figure 2. The RDP package (Martin and Rybicki 2000) produced very strong support for recombination for all cases we describe. All seven programs supported recombination for \textit{cox1} and \textit{cox3} regions, and five programs supported recombination for \textit{nad5} (table 2). In all positive cases, the $P$ value for the null hypothesis of no recombination was between $10^{-3}$ and $10^{-23}$.

The \textit{cox3} recombinant (sequence H8-2R, fig. 2) is identical to the paternal sequence (H8-21M) for $>200$ bp and switches to the maternal sequence (H8-5F) somewhere between positions #209 and #216. The point of crossover cannot be identified with higher resolution because the part between these two sites is identical in the parental
sequences. Also, we cannot tell whether the other crossover point (the one that is not included in the sequenced part) is upstream of site #1 or downstream of site #813. Consequently, we cannot say anything about the length of the exchanged part, except that it should be.

208

bp (if the unidentified crossover point is upstream of the site of the first primer) or.

596 bp (if it is downstream of the site of the second primer; table 2). The situation is similar for the

nad5

recombinant (sequence C29-6R) where only one crossover point can be identified at site #170 or #171. At the

cox1

region, we recovered eight different recombinant sequences among nine nonparental clones from three individuals (the recombinant sequences from two clones were identical, table 1). We can deduce both crossover points for the recombinant sequences H8-13R and BS29-5R for which the exchanged part is 570–614 and 387–415 bp, respectively. For the remaining six

cox1

recombinant sequences (H28-2R, H28-9R, H28-12R, H28-15R, BS29-9R, and BS29-17R), only one crossover point lies within the amplified fragment, so the length of the exchanged part cannot be determined.

The exchanged part is identical in the recombinant and the donor sequence in one case (BS29-17R). Among the other recombinants, three had 1-nt, five had 2-nt, and one had 4-nt differences, amounting to 17 substitutions (12 synonymous and 5 nonsynonymous) or 1% in the total of 1,553 diagnostic sites. There was also one 2-nt deletion (gene

nad5

, sequence C29-6R, positions #814–815) that leads to premature termination of the reading frame (fig. 2). All differences occur at different nucleotide sites,
except for a G to A change at position #7 of *cox1* that was found in two recombinants from two different individuals (sequences H28-2R and BS29-9R) and a T to C change at position #19 of *cox1* that was found in four recombinants, three from the same individual (sequences H28-12R, H28-9R, and H28-15R) and one from a different individual (sequence BS28-9R). Among the 10 crossover points that we have identified for the 8 recombinants of the *cox1* region, two (the downstream points of recombinants H28-15R and H8-13R) lay between the same diagnostic sites (#605 and #621). However, the possibility that the two recombinants may represent reciprocal products of the same recombination event (that were subsequently found in two different individuals) can be excluded given that these recombinants differ in their upstream crossover points.

There are two ways to account for the differences between recombinant and parental sequences. One is to assume that the recombination events occurred in the animals we have examined (real-time recombination) and the differences resulted from mutations that occurred in these animals either before recombination (this would be the case if the copy that acted as the substrate sequence for recombination had suffered a mutation that differentiated it from the parental sequence), after recombination or even as a consequence of the event of recombination. The second way is to assume that the recombinants were produced sometime in the past and entered the individual as part of the mtDNA pool the

![FIG. 2. (Continued)](https://academic.oup.com/mbe/article-abstract/28/6/1847/1068474)
individual inherited from its parents. In this case, the nucleotide differences between the parental and the recombinant sequence would be due to mutations that occurred in the intervening time between the event of recombination and the time the recombinant entered the individual in which it was observed (historical recombination).

One way to distinguish between these two hypotheses is to compare the F and the M parts of a recombinant sequence with the corresponding parts of the parental sequences of the individual in which the recombinant was observed, as well as with homologous parts of sequences from other individuals. If the recombinant sequence is more closely related to the individual’s parental sequence than to an external sequence, it would argue for the hypothesis that the recombinant originated in the scored individual. We have, therefore, produced phylogenetic trees that contained all available *M. galloprovincialis* sequences that had a similar length with the fragment that we surveyed for recombination in each of the three genes (see Materials and Methods). For the *cox3* gene, the set included eight F and two M external sequences. The corresponding numbers for the *nad5* gene were 10 and 2, and for the *cox1* gene 13 and 4. There were 20 trees in total, one for the F and one for the M part for each of the 10 recombinant sequences we have observed (trees not shown).

In 10 trees, the recombinant sequence formed a pair with the corresponding part of the parental sequence of the individual in which it was detected (with bootstrap values ≥ 90% in seven cases). In all 10 cases, the number of nucleotide differences between the recombinant and the parental sequence was smaller than between the recombinant and any external sequence. In seven trees, the recombinant and the parental sequence coclustered with one or more external sequence. In these cases, at least one external sequence was as close to the recombinant in terms of nucleotide substitutions as the fragment that we surveyed for recombination.

![Figure 2](https://学术.org/academic.oup.com/mbe/article-abstract/28/6/1847/1068474)
The recombinant and parental sequence codes are as in figure 2. Column 5: length (bp) of the exchanged fragment. Column 6: nucleotide difference of parental sequences (DPS). Columns 7–16: Sequence identity (SI) and sequence length (in parentheses) at crossover points. These segments are assumed to contain the MEPS required for a crossover to occur and are given under the assumption that no mismatch is allowed or that one mismatch is allowed. For two recombinants, both crossover points were within the sequence. For these sequences, columns 7 and 8 refer to the first crossover point (in a 5′→3′ direction) and columns 9 and 10 to the second.

Table 2. Information about Recombination Events.

<table>
<thead>
<tr>
<th>Recombinant Seq.</th>
<th>Parental Seq. 1</th>
<th>Parental Seq. 2</th>
<th>Gene</th>
<th>Length</th>
<th>DPS (%)</th>
<th>SI (No Mismatch)</th>
<th>SI (One Mismatch)</th>
<th>LS (no mismatch)</th>
<th>LS (One Mismatch)</th>
<th>Rejecting H0</th>
</tr>
</thead>
<tbody>
<tr>
<td>H8-2</td>
<td>H8-5</td>
<td>H8-21</td>
<td>cox3</td>
<td>&gt;208 or &gt;596</td>
<td>21.00</td>
<td>—</td>
<td>—</td>
<td>209–216 (8)</td>
<td>194–216 (23)</td>
<td>1,2,3,4,5,6,7</td>
</tr>
<tr>
<td>C29-6</td>
<td>C29-8</td>
<td>C29-3</td>
<td>nad5</td>
<td>&gt;169 or &gt;823</td>
<td>23.40</td>
<td>—</td>
<td>170–171 (2)</td>
<td>164–171 (8)</td>
<td>2,3,4,5,7</td>
<td></td>
</tr>
<tr>
<td>H28-2R</td>
<td>H28-1F</td>
<td>H28-7M</td>
<td>cox1</td>
<td>&gt;664 or &gt;165</td>
<td>16.83</td>
<td>—</td>
<td>665–672 (8)</td>
<td>653–672 (20)</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
<tr>
<td>H28-9R</td>
<td>H28-1F</td>
<td>H28-7M</td>
<td>cox1</td>
<td>&gt;442 or &gt;381</td>
<td>16.83</td>
<td>—</td>
<td>443–456 (14)</td>
<td>440–456 (17)</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
<tr>
<td>H28-12R</td>
<td>H28-1F</td>
<td>H28-7M</td>
<td>cox1</td>
<td>&gt;574 or &gt;246</td>
<td>16.83</td>
<td>—</td>
<td>575–591 (17)</td>
<td>566–591 (26)</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
<tr>
<td>H28-15R</td>
<td>H28-1F</td>
<td>H28-7M</td>
<td>cox1</td>
<td>&gt;604 or &gt;216</td>
<td>16.83</td>
<td>—</td>
<td>605–621 (17)</td>
<td>599–621 (23)</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
<tr>
<td>H8-13R</td>
<td>H8-11M</td>
<td>H8-5F</td>
<td>cox1</td>
<td>570 to 614</td>
<td>16.70</td>
<td>7–35 (29)</td>
<td>7–40 (34)</td>
<td>605–621 (17)</td>
<td>605–624 (20)</td>
<td>1,2,3,4,5,6,7</td>
</tr>
<tr>
<td>BS29-5R</td>
<td>BS29-1F</td>
<td>BS29-19M</td>
<td>cox1</td>
<td>387 to 415</td>
<td>16.46</td>
<td>23–39 (17)</td>
<td>23–45 (23)</td>
<td>426–438 (13)</td>
<td>425–238 (14)</td>
<td>1,2,3,4,5,6,7</td>
</tr>
<tr>
<td>BS29-9R</td>
<td>BS29-19M</td>
<td>BS29-1F</td>
<td>cox1</td>
<td>&gt;379 or &gt;450</td>
<td>16.46</td>
<td>—</td>
<td>380–387 (8)</td>
<td>377–387 (11)</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
<tr>
<td>BS29-17R</td>
<td>BS29-19M</td>
<td>BS29-1F</td>
<td>cox1</td>
<td>&gt;235 or &gt;583</td>
<td>16.46</td>
<td>—</td>
<td>239–254 (16)</td>
<td>236–254 (19)</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
<tr>
<td>AY363687</td>
<td>AY823623</td>
<td>AY497292</td>
<td>nad3</td>
<td>217</td>
<td>19.70</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1,2,3,4,5,6,7</td>
<td></td>
</tr>
</tbody>
</table>

Note.—The recombinant and parental sequence codes are as in figure 2. Column 5: length (bp) of the exchanged fragment. Column 6: nucleotide difference of parental sequences (DPS). Columns 7–16: Sequence identity (SI) and sequence length (in parentheses) at crossover points. These segments are assumed to contain the MEPS required for a crossover to occur and are given under the assumption that no mismatch is allowed or that one mismatch is allowed. For two recombinants, both crossover points were within the sequence. For these sequences, columns 7 and 8 refer to the first crossover point (in a 5′→3′ direction) and columns 9 and 10 to the second. Columns 9 and 10 give also the information for the single crossover for the other eight recombinants. The last column gives the tests in the RDP package (seven tests in all) that reject the null hypothesis (H0) of no recombination (RDP tests: 1. RDP, 2. GENECONV, 3. Bootscan, 4. Maxchi, 5. Chimaera, 6. SiScan, and 7. 3seq).

Our data provide information only on sequence difference between recombining sequences and nucleotide identity at the crossover points. The length of identity at the crossover points is the number of nucleotides between the diagnostic site just before the recombinant sequence switches from one parental sequence to the other and the next diagnostic site (e.g., for the recombinant sequence cox3 H8-2R [fig. 2], the crossover point lies between sites #208 and #217, and the length of identity runs from site #209 to #216 included, i.e., it is 8 bp long). If these stretches of identity are assumed to play the same role as the MEPS of nuclear recombination, then the maximum MEPS length for the recombination events we have detected should vary from 2 to 17 bp on the assumption that the MEPS identity must be perfect or from 8 to 26 if one-base mismatch is allowed (table 2). We note that MEPS is defined as the minimum number of identical nucleotides required for recombination to occur rather than the number of identical nucleotides one observes after recombination has occurred. The actual MEPS length could therefore be smaller than the length of observed identity.

Nucleotide difference between the parental sequences varied from 16.46% to 23.40% (table 2), so one could ask whether recombinants were recovered at a higher rate in less different regions. There are several reasons why there can be no satisfactory answer to this question. One is that our prescreening assay depended on the density of restriction sites along the surveyed region. This precludes a comparison of the nad5 region with the other two regions (fig. 1). Another inherent difficulty is that the rate with which recombinants are produced and the frequency with which recombinants occur in the mtDNA pool at the time of the survey need not be correlated. There is no mechanism for ensuring that upon cell division the mtDNA content of two daughter cells will be the same, like mitosis does for the nuclear DNA, with the result that the mtDNA content varies from tissue to tissue, including the possibility of fixation of two cell lines for two different mtDNA variants.
NUMT sequences are much higher (Arctander 1995). We note, however, that the number of recombinant sequences we recovered at the cox1 region, at which the difference between F and M genomes in the exchanged regions is on average 16.7% (table 2), is much larger than that at the cox3, at which the difference is 21% (8 recombinants in 132 clones of cox1, 1 recombinant in 212 clones of cox3, \( P = 0.0023 \) by Fisher’s exact test). The two regions have about the same length, and the density of restriction sites is higher in cox3, which would have facilitated the detection of recombinants in cox3 at the same degree, if not higher, as in cox1.

Discussion

This study was motivated by our previous report of frequent mtDNA recombination in M. galloprovincialis males that were heteroplasmic for two mitochondrial genomes that differed by about 3% (Ladoukakis and Zouros 2001a). These were exceptional males that have inherited the standard F genome from their mother and a “masculinized,” that is, an F-like, genome from their father. In common males of this species, the two genomes, the standard F and M genomes, differ by about 20%. We report here 11 cases of homologous recombination between parts of these genomes. The nucleotide difference in these parts varied from 16.5% to 23.5% (table 2).

Recombinant mtDNA Sequences Are Not Artefacts

Our method for detection of recombinant sequences is theoretically open to two types of experimental error. The first stems from the fact that the source DNA for our amplifications contained both nuclear and mitochondrial DNA. This leaves open the possibility that the recombinants we have detected are nuclear mitochondrial-like sequences (NUMTs) (Lopez et al. 1994; Bensasson et al. 2000). This possibility appears highly unlikely in our case. NUMTs are nuclear sequences, and as such, they ought to appear in equal frequency in female and male mussels. But no recombinant-like sequences of the type reported here were yet recovered from females (Dalziel and Stewart 2002; Ladoukakis et al. 2002). For the same reason, NUMTs should appear in equal frequency in extracts from different tissues, either male or female. Yet recombinants are recovered almost exclusively from male gonads, where the two genomes coexist (Garrido-Ramos et al. 1998; Ladoukakis and Zouros 2001a).

This conclusion is also consistent with the fact that we have observed only one stop codon (due to a 2-bp deletion in the nad5 recombinant, see above) in a total length of 9,104 bp (862 for cox1, 1,8900 for cox3, and 1,042 for nad5) recombinant sequences (or 3,034 codons). In organisms in which NUMTs were found, the sequence divergence between NUMTs and mtDNA and the number of stop codons in NUMT sequences are much higher (Arctander 1995).

The second, more likely, source of error comes from the possibility of DNA polymerase “jumping” from one parental template to the other during the PCR reaction (Paabo et al. 1990). The products of this jumping will be indistinguishable from real recombinants. We tested this possibility in our first report of recombination between the F and the C mtDNA genomes in two ways (Ladoukakis and Zouros 2001a). First, we mixed cloned homologous parts of the F and the C genome and searched for recombinants on the rationale that polymerase jumping would produce recombinants under these in vitro conditions. We found none. Second, we returned to the same animals in which we first recovered a recombinant and repeated the amplification. In all cases, we recovered the same recombinants.

Origin and Fate of Recombinants

The main point of this study is that recombination may occur among highly divergent mtDNA sequences. A secondary but important issue that emerges from this observation is whether the recombinant sequences we have observed were the products of real-time recombination events, that is, they were produced in the individuals they were detected or are parts of an mtDNA polymorphism that entered the individuals via the egg or the sperm. The fact that all recombinant sequences we have recovered are unique and were found in the same individual with the parental sequences from which they could be produced by recombination argues in favor of the first hypothesis. This, however, requires an explanation of the 1% difference at the exchanged parts between parental and recombinant sequences. PCR errors or mutations that occurred, before or after recombination, in the specific molecules that were cloned and sequenced would be a possibility. This possibility is strongly supported by the observation that only in 2 of 20 comparisons did the recombinant sequence show a higher similarity with a sequence other than the individual’s parental sequence— and this by just 1 nt. There is also good independent evidence against the hypothesis that the recombinant genomes entered the individuals through one or the other gamete along with the parental genome. First, extensive surveys of mtDNA variation within and among populations of various species of Mytilus (Ladoukakis et al. 2002; Burzynski et al. 2006; Filipowicz et al. 2008; Zbawicka et al. 2010) produced no sign of any type of intra- or interspecies variation that could be attributed to mtDNA recombination. Second, there is good evidence that the Mytilus sperm contains only one type of M mtDNA (Venetis et al. 2006); thus, the possibility that a male could inherit more than one type of mtDNA from its male parent appears unlikely. This is less certain for the female parent. Extensive egg heteroplasmy was reported in M. galloprovincialis (Obata et al. 2006), but no such heteroplasmy could be detected in the venerid Venerupis philippinarum (Ghiselli et al. 2011).
Recombination and Sequence Divergence

The main factor that impedes homologous recombination between nuclear sequences of elevated degree of divergence is the mismatch repair mechanism (MMR). The anti–recombinational efficiency of the MMR system varies among organisms. Known cases of recombination between sequences of high divergence in Arabidopsis (Opperman et al. 2004) and in mammalian cells with a defective MMR system (Smith et al. 2007) do not exceed the level of 10% differentiation. In prokaryotes, recombination may occur among genomes with gene rearrangements or between organisms that belong to different taxa. As a result, recombination in prokaryotes, even though rarer than in eukaryotes, may involve sequences with higher levels of divergence (Zawadzki et al. 1995). In Bacillus, very little recombination was observed (6.9 × 10^{-6} sexual isolation reduced by a factor of 354.8) when the sequence divergence was 14.5% even though the mismatch repair system was mutated (Majewski and Cohen 1999), but in Streptococcus, there are reported cases of recombination between sequences that differed by as much as 18% (Majewski et al. 2000). In λ virus, where genome mosaics are common, cases of recombination between sequences differing by 22% are known (Martinsohn et al. 2008). This high degree of divergence is comparable to the 23% we report here.

Having established that recombination could occur among mtDNA sequences that differ by >20%, the next question is whether the frequency of recombination may be affected by the degree of nucleotide difference. The variation of the mtDNA pool of an organism is under a strong and constant stochastic or selective flux (Bergstrom and Pritchard 1998; Meiklejohn et al. 2007), which presents an inherent difficulty in approaching this question. Our data suggest that recombination may occur more frequently at the cox1 region where the nucleotide difference between the F and the M genome is 16.7% than in the cox3 region where the difference is 21% (divergence corrected for multiple hits is 0.1959 and 0.2608, respectively). A better approach would be to use cases in which the difference of the two parental genomes varies widely, as with as low as 3%, as in the case of male mussels with a masculinized paternal genome (Venetis et al. 2007), to as high as 40%, as is the case of F and M genomes of M. californianus (Ort and Pogson 2007).

Is There MMR Anti–Recombinational Activity in Mitochondria?

The role of the MMR system in mitochondria is much less understood. In yeast and mammals, there is mismatch repairing of mtDNA (Larsen et al. 2005), but the mammalian mitochondrial MMR system lacks strand recognition (Mason and Lightowlers 2003; Mason et al. 2003). Two core proteins of the MMR system, Msh1 and Mlh1, have been found in yeast mitochondria (Larsen et al. 2005), where deletion of the MSH1 gene has been linked to rearrangements in the mitochondrial genome (Reenan and Kolodner 1992). Evidence of an MMR mediated anti–recombinational action also exists for the mtDNA of Arabidopsis, where it has been shown that Msh1 suppresses recombination between repeated mtDNA sequences and increases the stability of the mitochondrial genome (Shedge et al. 2007; Arrieta-Montiel et al. 2009). In contrast, the presence of MMR proteins in mammalian mitochondria remains controversial (Chen et al. 2001).

Assuming that the MMR system is present in the mitochondrion, one may ask whether its anti–recombinational activity may be different in the mitochondrion than in the nucleus. The proteins that comprise the nuclear and mitochondrial MMR system are apparently encoded by the same nuclear genes, yet it is known that the mismatch repair mechanism is not as effective for the mtDNA as for the nuclear DNA in mammalian cells (Mason and Lightowlers 2003; Mason et al. 2003). This leaves open the possibility that the MMR system may have a reduced anti–recombinational activity for the mtDNA and that the etiology for this might be the same as for the reduced mismatch repair activity. Moreover, the mismatch repair and the anti–recombination function of the MMR system may involve different sets of proteins (Elez et al. 2007), so it is possible that one function is affected, whereas the other is maintained intact or is only modestly modified.

It is more difficult to see why selection would allow the reduction of the anti–recombinational activity of the MMR system in the mitochondrion while maintaining this activity at high levels in the nucleus. One hypothesis that may provide an answer to this question is that the strict uniparental inheritance of animal mtDNA and the ensuing homoplasy has reduced selection pressure against homologous recombination for the simple reason that recombination between identical genomes would not lead to genetic imbalance or interfere with gene-order organization. As noted, such disturbances are assumed to be the reason recombination has been restricted to nearly identical (homologous) sequences through the recruitment of the MMR system. One argument against this hypothesis might be that our observations were made in a species with DUI. In these species, maternal and paternal sequences are highly diverged, and the origin of DUI may be as old as 400 my (Hoeh et al. 1996; Liu et al. 1996). Under these circumstances, one may argue, the anti–recombinational strength of MMR ought to
be restored in species with DUI. The argument is muted if one takes into consideration that, despite their high divergence, the two genomes have the same gene arrangement and all genes are fully functional in both (Mizi et al. 2005; Breton et al. 2006; Zbawicka et al. 2007). In addition, we argue that the recombination events we report here are restricted to somatic tissues of males, in which case the recombinant sequences cannot be transmitted to the next generation. It is therefore reasonable to suggest that pressure for regaining the anti–recombination strength has never acted on the mitochondrial MMR of DUI species.

There remains the question of whether the information on mtDNA recombination that we gained from the atypical mitochondrial system of Mytilus can be extended to other species, outside those having the DUI system. This question must remain open, until the issue can be addressed in other animals, perhaps through rare cases of haploplasmic individuals. We only note that a similar concern on the very occurrence of mtDNA recombination in other animals, following our first report of such recombination in mussels (Ladoukakis and Zouros 2001a), proved unfounded, as evidenced by the appearance of subsequent reports of both real-time (Hoarau et al. 2002; Ganetbein et al. 2005; Sato et al. 2005) and historical (Pignaneau et al. 2004; Tsoukis et al. 2005) recombination. Also, the discovery of heteroplasmic with pathological consequences in a human subject (Schwartz and Vissing 2002) was soon followed by a report of mtDNA recombination in the same subject (Kraytsberg et al. 2004).

When viewed as a cause of mtDNA variation in a phylogenetic context, animal mtDNA recombination might not be considered important. This is because strict maternal inheritance leads to mtDNA homoplasy and recombination between identical genomes, no matter how common, does not generate new mtDNA variants. The situation might be different if one considered mtDNA mutations and deletions that are produced during the life of the organism. The importance of mitochondrial fusion for normal cellular function is well established for at least some types of mammalian cells, and one reason for this might be that fusion allows the cell to keep deleterious mtDNA mutations and deletions at low levels (for a most recent reference, see Chen et al. 2010). Recombination among mtDNA genomes of fused mitochondria might play an important role in this process (Neiman and Taylor 2009).

Acknowledgments

This work was supported by the Marine Genomics Europe Network of Excellence (to E.Z) and by the University of Crete (to E.D.L). We thank Dr. F. Cohan, an anonymous reviewer, and associate editor Dr. Thomas for constructive comments. We also thank Dr. F. Cohan for suggesting the phylogenetic test for the origin of recombinants.

References


the Bivalvia (Mollusca): searching for the origin and mitoge
nomic correlates of doubly uniparental inheritance of mtDNA.
BMC Evol Biol. 10:50.
Filipowicz M, Burzynski A, Smietanka B, Wenne R. 2008. Re-
Garrido-Ramos MA, Stewart DT, Sutherland B, Zouros E. 1998. The distribution of male-transmitted and female-transmitted mito-
ochondrial DNA tyles in somatic tissues of blue mussels: implications for the operation of doubly uniparental inheritance of mitochondrial DNA. Genome 41:818–824.
Ladoukakis ED, Zouros E. 2001a. Direct evidence for homologous recombination in mussel (Mytilus gallopriovinicialis) mitochond-
Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG. 2000. Barriers to genetic exchange between bacterial species: Strepto-
Mizi A, Zouros E, Moschonas N, Rodakis GC. 2005. The complete maternal and paternal mitochondrial genomes of the Mitter-
null