Molecular Population Genetics of the Male and Female Mitochondrial DNA Molecules of the California Sea Mussel, *Mytilus californianus*

Brian S. Ort¹ and Grant H. Pogson²

Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California 95064

Manuscript received March 6, 2007

Accepted for publication August 19, 2007

ABSTRACT

The presence of two gender-associated mitochondrial genomes in marine mussels provides a unique opportunity to investigate the dynamics of mtDNA evolution without complications inherent in interspecific comparisons. Here, we assess the relative importance of selection, mutation, and differential constraint in shaping the patterns of polymorphism within and divergence between the male (M) and female (F) mitochondrial genomes of the California sea mussel, *Mytilus californianus*. Partial sequences were obtained from homologous regions of four genes (*nad2*, *cox1*, *atp6*, and *nad5*) totaling 2307 bp in length. The M and F mtDNA molecules of *M. californianus* exhibited extensive levels of nucleotide polymorphism and were more highly diverged than observed in other mytilids (overall Tamura–Nei distances >40%). Consistent with previous studies, the M molecule had significantly higher levels of silent and replacement polymorphism relative to F. Both genomes possessed large numbers of singleton and low-frequency mutations that gave rise to significantly negative Tajima's *D* values. Mutation-rate scalars estimated for silent and replacement mutations were elevated in the M genome but were not sufficient to account for its higher level of polymorphism. McDonald–Kreitman tests were highly significant at all loci due to excess numbers of fixed replacement mutations between molecules. Strong purifying selection was evident in both genomes in keeping the majority of replacement mutations at low population frequencies but appeared to be slightly relaxed in M. Our results suggest that a reduction in selective constraint acting on the M genome remains the best explanation for its greater levels of polymorphism and faster rate of evolution.

ONE of the primary reasons that animal mitochondrial DNA (mtDNA) has gained prominence in evolutionary and population genetics studies has been the assumption that it evolves in a strictly neutral manner (BROWN *et al.* 1979; AVISE *et al.* 1987). Recently, however, this assumption has been questioned as a growing number of reports have documented departures from neutral patterns of mtDNA polymorphism within and divergence between species. Overall, the strongest evidence against the strictly neutral model has come from studies suggesting that the majority of amino acid replacement mutations in the mtDNA genome experience weak levels of purifying selection (*e.g.*, NACHMAN *et al.* 1996; NACHMAN 1998; RAND and KANN 1998), an observation consistent with the nearly neutral theory of molecular evolution (reviewed by RAND 2001; BALLARD and WHITLOCK 2004). Further evidence has come from studies documenting effects of different mtDNA haplotypes on individual fitness (*e.g.*, CLARK and LYCKEGAARD 1988; JAMES and BALLARD 2003), cytonuclear interac-

tions (*e.g.*, SCHMIDT *et al.* 2001; WILLETT and BURTON 2004), geographic patterns of mtDNA variation (*e.g.*, RUIZ-PESINI *et al.* 2004; GRANT *et al.* 2006), and a recent report by BAZIN *et al.* (2006) that, on the basis of the absence of an expected relationship between effective population size and mtDNA variability, implicated a role for selective sweeps in reducing polymorphism to a level below that expected at mutation-drift equilibrium (notably in invertebrates).

Much of our understanding of broad-scale patterns of mtDNA evolution has come from comparisons undertaken among different species. However, several issues complicate evaluating the relative importance of selection and random drift in the evolution of the mtDNA molecule from interspecific data. First, different species are likely to experience different demographic histories that can exert major effects on the patterns and levels of mtDNA polymorphism. Disentangling the relative effects of selection and demography on a single genetic locus (*i.e.*, mtDNA) has proven to be difficult and perhaps can be overcome only by the analysis of multi-locus data (*e.g.*, WILLIAMSON *et al.* 2005; LI and STEPHAN 2006). Second, if mtDNA evolution does conform to the nearly neutral theory, then differences in species' effective population sizes can be expected to play important roles in determining levels of polymorphism

¹Present address: Romberg Tiburon Center for Environmental Studies, San Francisco State University, 3152 Paradise Dr., Tiburon, CA 94920.

²Corresponding author: Department of Ecology and Evolutionary Biology, Earth and Marine Sciences Bldg., University of California, Santa Cruz, CA 95064. E-mail: pogson@biology.ucsc.edu

and the rates of both neutral and adaptive evolution. Here, too, evaluating the importance of demographic events (such as population bottlenecks) can be accomplished only from a multilocus perspective that is typically lacking in mtDNA studies.

The coexistence of two gender-associated mtDNA lineages in marine mussels provides an excellent system for studying how selection affects the evolution of mitochondrial genomes while avoiding many of the complications implicit in interspecific comparisons. Mussels belonging to the genus *Mytilus*, as well as some species of unionid mussels (LIU *et al.* 1996) and venerid clams (PASSAMONTI and SCALI 2001), have an unusual pattern of sex-specific mtDNA inheritance known as "doubly uniparental inheritance" or DUI (SKIBINSKI *et al.* 1994; ZOUROS *et al.* 1994). Females are homoplasmic for an "F" mtDNA molecule that is inherited from mother to daughter in the usual metazoan pattern. Males, however, are heteroplasmic for an F molecule in their somatic tissues and a male ("M") mtDNA that predominates in the gonad and is transmitted from father to son. Because the F and M mtDNA molecules reside in the same species, both genomes have experienced similar demographic histories and, given a 1:1 sex ratio in mytilids (FISHER and SKIBINSKI 1990), have similar effective population sizes (assuming negligible sperm limitation and/or competition). Furthermore, the evolution of the male and female mtDNA lineages in marine mussels occurs in the same nuclear genetic background lacking sex chromosomes, thus avoiding complications arising from lineage-specific cytonuclear interactions or mutational biases.

Previous studies on the F and M mtDNA molecules in marine mussels have established that the male molecule possesses higher levels of polymorphism and evolves more rapidly than the female molecule. Two main hypotheses have been proposed to explain these patterns. First, STEWART *et al.* (1996) proposed that the male molecule might experience a relaxation of selective constraint because the only tissue in which it experiences direct exposure to selection is the male gonad. Second, the M genome might experience a higher mutation rate on the basis of a greater number of cell divisions that occur during spermiogenesis (seven divisions) compared to oogenesis (four divisions) (RAWSON and HILBISH 1995; STEWART *et al.* 1995). These two hypotheses are clearly not mutually exclusive and evaluating their relative importance has proven to be difficult. The majority of studies on the DUI system in mytilids have focused on single-gene regions, thus precluding comparisons among loci experiencing different levels of selective constraint within and between molecules. Furthermore, the evolutionary histories of M and F molecules in the blue mussel complex (*i.e.*, *Mytilus edulis*, *M. trossulus*, and *M. galloprovincialis*) have been complicated by "masculinization" events (in which an F molecule invades the male lineage) and by historical

and contemporary introgression that appears to vary both among taxa and geographic regions (see QUESADA *et al.* 1995, 1998a; RAWSON *et al.* 1996; SAAVEDRA *et al.* 1996; RAWSON and HILBISH 1998; RIGINOS *et al.* 2004).

In this study, we evaluate the relative importance of differential constraint, mutation, and diversifying selection in shaping the patterns of polymorphism within and divergence between the male and female mitochondrial genomes of the California sea mussel, *Mytilus californianus*, which, unlike other mytilids, does not hybridize with any extant species in the northeast Pacific region. We collected large samples of DNA sequences from four homologous regions of the M and F mtDNA genomes to determine if the mutation rate of the male molecule is consistently elevated at both silent and replacement sites across different loci and to more fully assess differences in the degree of selective constraint acting within and between the two molecules. Our results suggest that a relaxation of selective constraint on the M genome can account for differences observed between patterns of polymorphism within and divergence between the M and F mtDNA molecules in *M. californianus* without needing to invoke a higher intrinsic mutation rate as an ancillary factor.

MATERIALS AND METHODS

Samples: Approximately 150 adult mussels (3–10 cm shell height) were collected in the winter and spring of 1999 and 2000 from four northeast Pacific localities: Strawberry Hill, Oregon (124° 4' 40.39" W, 44° 32' 8.88" N); Terrace Point, California (122° 5' 17.08" W, 36° 57' 38.23" N); Arroyo Hondo, California (119° 57' 41.22" W, 34° 26' 28.87" N); and Punta Baja, Baja California (115° 48' 29.54" W, 29° 57' 4.34" N). Animals were transported back to the laboratory on ice where tissues were dissected and preserved for DNA extraction.

DNA extractions: Individuals were sexed by visual inspection of the gonad under a dissecting microscope. Total DNA was extracted from ethanol-preserved gill tissues as described in POGSON *et al.* (1995) and used to PCR amplify F mtDNA sequences. The M molecule was amplified from DNA isolated from an enriched mitochondrial fraction prepared from ripe male gonad tissue. Approximately 1 g of male gonad was washed in high TE (100 mM Tris-HCl, 10 mM EDTA, pH 8.0) and homogenized in a 15-ml Wheaton glass tissue grinder in 10 ml ice-cold TEK buffer (50 mM Tris-HCl, pH 7.5, 40 mM EDTA, 1.5% KCl). The homogenate was centrifuged at $1475 \times g$ in a Sorvall RC58 centrifuge for 20 min at 4° to remove nuclear and cellular debris. The supernatant was transferred to a clean tube and spun again at $1475 \times g$ for 15 min at 4°. The resulting supernatant was transferred to a new tube and centrifuged at $14,500 \times g$ for 15 min at 4° to pellet mitochondria. The resulting pellet was resuspended in 610 µl salt lysis buffer (10 mM Tris-HCl, 400 mM NaCl, 2 mM EDTA, pH 8.3) and total DNA was extracted as described above.

PCR amplifications and DNA sequencing: PCR primers specific for the *M. californianus* M and F molecules were initially designed from the cytochrome oxidase subunit I (*coxI*) and NADH dehydrogenase subunit V (*nad5*) sequences of BEAGLEY *et al.* (1997). These were paired with a forward primer positioned near the beginning of the *M. edulis* NADH

dehydrogenase subunit II (*nad2*) sequence of HOFFMAN *et al.* (1992) that amplified both genomes. *M. californianus* M- and F-specific primers for the *nad2*, *cox1*, adenosine triphosphatase subunit 6 (*atp6*), and *nad5* genes of *M. californianus* were then designed from samples of 12–15 sequences obtained from the Terrace Point population. Primer sequences and reaction conditions for the eight genes are provided in the APPENDIX. PCR reactions contained 20 mM Tris-HCl (pH 8.8 at 25°), 10 mM KCl, 10 mM (NH₄)₂SO₄, 2.5 mM MgSO₄, 0.1% Triton X-100, 100 ng/μl bovine serum albumin, 200 μM each dNTP, 0.25 μM forward and reverse primers, 0.6 units of *Taq* 2000 DNA polymerase (Stratagene, La Jolla, CA), 0.6 units *Taq* Extender PCR additive (Stratagene), and 150 ng template DNA. Reactions were performed in 10-μl sealed glass capillary tubes in an Idaho Technology A1605 Air Thermo-Cycler. After an initial denaturation step of 45 sec at 94°, the tubes were exposed to 35 cycles of denaturation at 94° for 1 sec, primer annealing at 52°–54° for 1 sec, primer extension at 72° for 50–60 sec, and a final hold at 72° for 1 min (see APPENDIX for details).

PCR products were excised from 1% agarose gels stained with ethidium bromide and purified using ZymoSpin columns following the manufacturer's instructions (Zymo Research). For all genes, complete sequences off both strands were obtained using an Applied Biosystems (ABI; Foster City, CA) model 373 Automated Sequencer or an ABI model 3100 Genetic Analyzer. Sequences were edited with Sequence Navigator, and compiled into consensus sequences using Auto-Assembler (both programs from ABI). All nucleotide sequences have been deposited in GenBank (accession nos. EU028349–EU029064).

Data analyses: Standard measures of nucleotide polymorphism, tests of neutrality, and estimates of codon bias were obtained from the DnaSP (version 4.00) program of ROZAS *et al.* (2003). We also used the neutrality index of RAND and KANN (1996) to assess the magnitude and direction of departures from neutral expectations in the ratios of polymorphism and divergence. Tests of population differentiation (AMOVAs) were performed using the ARLEQUIN (version 2.000) program of SCHNEIDER *et al.* (2000). AMOVAs were performed using all of the data and on data sets that contained only silent or only replacement mutations. The resulting Φ_{ST} values were tested for significance by 10,000 random permutations of haplotypes among populations. Tests for recombination within and between the F and M genomes were performed using the four-gamete test of HUDSON and KAPLAN (1985) and the LDhat program of McVEAN *et al.* (2002), a coalescent-based method that is robust to unequal substitution rates across sites and multiple hits in genomes experiencing high substitution rates. The MEGA (v. 3.0) program of KUMAR *et al.* (2004) was used to estimate the net nucleotide and amino acid divergence between the four genes.

Phylogenetic reconstructions were performed with MRBAYES v.3.1 (RONQUIST and HUELSENBECK 2003) using a GTR + I + Γ model identified by MODELTEST v.3.8 (POSADA and CRANDALL 1998) on the combined data partitioned by gene. Five independent runs of four Markov chains were run for 2 million generations using the default temperature parameter (0.2) and sampling trees every 1000 steps after an initial burn-in of 200,000 generations. Convergence was assessed using the TRACER v.1.3 program (RAMBAUT and DRUMMOND 2003, 2004, 2005) and by comparing the output from three independent runs.

To evaluate the hypothesis that the M molecule experiences a relaxation of selective constraint, we compared the magnitude of Tajima's *D* between genes using all sites, as well as for silent and replacement sites separately (TAJIMA 1989). We further evaluated the differential constraint hypothesis using

GRANTHAM'S (1974) distances to quantify the physicochemical difference between amino acid changes that were either polymorphic within F or M or fixed between molecules. Each amino acid substitution was given a numeric score that quantifies the magnitude of change on the basis of atomic composition, polarity, and charge. Amino acid changes were categorized as conservative, moderately conservative, moderately radical, or radical, corresponding to GRANTHAM'S (1974) scores of 0–50, 51–100, 101–150, or >150, respectively (LI *et al.* 1985). The proportions of fixed and polymorphic amino acid replacement mutations falling into each category were compared between the M and F lineages.

To test whether the M molecule has an intrinsically higher rate of mutation than the F molecule, we estimated mutation-rate scalars for silent and replacement mutations by coalescent simulations implemented by the "Isolation with Migration" (IM program, version 4.0) model of NIELSEN and WAKELEY (2001). For these simulations, mutation rates were estimated for each locus by setting the time of separation parameter (*t*) and two migration parameters (*m*₁ and *m*₂) equal to zero and constraining $\theta_1 = \theta_2 = \theta_A$. Under these conditions, the IM model mimics a single panmictic population, allowing estimation of only θ_A and the mutation-rate scalars. For all four genes, rate scalars for the F and M molecules were estimated for silent and replacement mutations separately. To compare the relative differences between M and F, we included nucleotide sequence data (105 sequences, 615 bp in length) from an anonymous nuclear gene (*Mcal25*) sequenced from the same four populations to serve as a "control" for each IM run. Since the F and M molecules both fail the four-gamete test (due to back mutation; see below), we used the Hasegawa-Kishino-Yano model for all loci. We ran six parallel chains with a geometric heating scheme (*g*₁ = 0.8, *g*₂ = 0.9) to improve mixing. After a burn-in of 100,000 steps, the chains were run for at least 1.5×10^6 steps, sampling every 1000 steps, until we achieved MCMC effective sample sizes (ESSs) of at least 10,000. Each analysis was repeated at least three times to assess the convergence of the estimates. Mutation-rate scalars were also determined separately for first, second, and third positions using the concatenated data from all four loci and nuclear controls for each run. For these simulations, we ran six chains with geometric heating (*g*₁ = 0.80, *g*₂ = 0.90) with burn-ins of 100,000 steps. The chains were run for at least 1.5×10^6 steps and achieved ESSs of at least 5000. To determine the probability that the mutation-rate estimate for M exceeds that for F, we modified the IM program to record an indicator variable *I*(*j*) that took a value of 1 if the hypothesis is true (*i.e.*, $\mu_M > \mu_F$) and a value of zero if the hypothesis is false (*i.e.*, $\mu_F > \mu_M$) at each sampling step *j* of the MCMC data set. An estimate of the probability that $\mu_M > \mu_F$ is the average of the indicator variable over *s* samples: $\Pr(\mu_M > \mu_F) = 1/s \sum I(j)$. Posterior odds ratios are computed as $P/(1 - P)$, where $P > 0.5$.

RESULTS

Partial sequences were obtained from four genes (*nad2*, *cox1*, *atp6*, and *nad5*) from the M and F mtDNA molecules of *M. californianus*. Sample sizes for each gene ranged from 20 to 23 sequences from each of four populations producing total sample sizes of 87–94 sequences/locus (Table 1). Combining the data from all four genes yielded 2307 bp of aligned sequence from 87 F and 85 M molecules. No insertions or deletions of amino acids were observed between the two genomes. M sequences were successfully amplified from all individuals

visually identified as males, confirming the absence of masculinized F molecules in our sample.

Population structure: AMOVAs were used to test for population differentiation among four populations spanning ~2000 km of Pacific coastline. The final column of Table 1 lists Φ_{ST} values for each locus individually and for the combined data. No significant differences were observed among populations for either F or M; Φ_{ST} values were negative for the four F genes and always <0.01 for M. Similar results were obtained for analyses considering only silent or replacement mutations (not shown). Sequentially pruning mutations with frequencies <0.10 from the data (*i.e.*, singletons, doubletons, etc.) also failed to reveal any significant population structure at deeper levels in the gene genealogies (not shown). F and M sequences from all populations were thus combined for subsequent analyses.

Recombination: Both the M and F molecules failed the four-gamete test of HUDSON and KAPLAN (1985), suggesting the presence of recombination (minimum numbers of recombination events in the combined data were 47 and 18, respectively). However, permutation tests implemented by the LDhat program of McVEAN *et al.* (2002) did not detect significant levels of recombination in either genome. Since the approach of McVEAN *et al.* (2002) uses a finite-sites model of nucleotide substitution, this discrepancy is likely due to recurrent mutation. The magnitude of homoplasy is likely higher in the M molecule in which 39 sites (1.7%) were segregating for three or four mutations compared to only 9 (0.4%) in F. No recombination events were observed between the F and M molecules across the region studied.

DNA polymorphism: The levels of nucleotide polymorphism detected in the M and F mtDNA molecules are summarized in Table 1. As found for other *Mytilus* species, the M molecule exhibited significantly higher levels of polymorphism than F. Overall, the M sequences had 66.2% more segregating sites, 75.2% more total mutations, and two to four times the levels of nucleotide diversity compared to F. Despite this difference in polymorphism, haplotype diversity in the combined data was nearly 100% in both genomes. Among males, only one haplotype was seen twice (in two individuals from Oregon) while in females only two haplotypes were found in multiple copies. Gene trees constructed for the combined F and M sequences were similar in showing this extreme sequence diversity but the male genealogy exhibited longer tips and deeper internal branches (Figure 1, a and b).

In both molecules, the *nad2* locus exhibited the highest level of polymorphism. The rank ordering of the three other loci differed among molecules, with *cox1* and *atp6* being the least-variable genes in males and females, respectively. Both genomes were AT rich and exhibited similar base compositions and patterns of codon usage (Table 1). Transitions strongly outnumbered

TABLE 1
Nucleotide polymorphism, base composition, and tests of neutrality

Locus	N	Length (bp)	No. of segregating sites	No. of haplotypes	Haplotype diversity	Effective no. of codons	Base composition				Transitions	Transversions	π^a	θ /site	Tajima's <i>D</i>	Φ_{ST}
							C	T	A	G						
Female																
<i>mad2</i>	94	726	102	75	0.981	48.7	12.9	35.5	27.5	24.1	13	97	0.0062	0.0288	-2.61***	-0.00493
<i>cox1</i>	92	444	43	49	0.861	50.6	14.9	36.0	26.6	22.5	3	42	0.0042	0.0199	-2.51***	-0.00174
<i>atp6</i>	88	501	39	37	0.770	38.4	11.0	41.0	25.0	23.0	1	38	0.0029	0.0154	-2.57***	-0.00267
<i>mad5</i>	91	636	83	57	0.885	46.2	12.8	38.8	25.1	23.3	13	74	0.0039	0.0272	-2.83***	-0.00231
Combined																
	87	2307	260	82	0.997	45.8	12.8	37.7	26.1	23.3	30	244	0.0045	0.0233	-2.76***	-0.00293
Male																
<i>mad2</i>	85	726	152	81	0.999	43.9	11.1	36.5	27.3	25.2	29	139	0.0137	0.0462	-2.38**	0.00534
<i>cox1</i>	87	444	71	69	0.987	50.4	15.1	35.6	27.5	21.9	6	69	0.0101	0.0335	-2.30**	0.00330
<i>atp6</i>	92	501	95	79	0.990	40.6	12.0	37.9	26.8	23.3	21	83	0.0121	0.0408	-2.33**	0.00315
<i>mad5</i>	87	636	116	81	0.998	47.0	14.0	35.8	29.8	20.4	22	105	0.0112	0.0400	-2.42**	-0.00447
Combined																
	85	2307	432	84	0.999	48.5	12.9	36.4	27.9	22.8	77	396	0.0122	0.0409	-2.42**	0.00101

*** $P < 0.01$, **** $P < 0.001$.

^a Nucleotide diversity.

FIGURE 1.—Bayesian consensus trees for the (a) F mtDNA sequences, (b) M mtDNA sequences, and (c) the combined F and M data.

transversions in both molecules and the bias favoring the former was consistently higher in the F genome. Tajima's *D* statistics were significantly negative for all genes in both molecules, indicating a significant excess of low-frequency mutations. However, Tajima's *D* values were consistently more negative in the F sequences, indicating a more pronounced skew of the site frequency spectra consistent with stronger purifying selection.

The increased diversity of the male molecule was attributable to elevated levels of both silent and replacement polymorphism (Figure 2). The magnitude of silent-site polymorphism exhibited less variation among loci in M compared to F. A similar pattern was seen for replacement polymorphism although the *cox1* gene was not variable in F. With the exception of the *atp6* locus in M, the majority of silent and replacement mutations in both molecules occurred as singletons (Table 2). The impact of purifying selection was clearly stronger in the female genome; the highest frequency of a replacement mutation in the sample was 0.043 (a V → F change in *nad2*). In contrast, seven replacement mutations were observed in the M molecule with frequencies >0.10 but only three exceeded 0.20. Both silent and replacement mutations had significantly lower mean population frequencies in the F molecule, giving rise, on average, to more negative values of Tajima's *D*. However, median allele frequencies were very similar in F and M (0.0110

FIGURE 2.—Comparison of silent and replacement nucleotide diversities in the M and F mtDNA molecules.

and 0.0118, respectively) and, surprisingly, did not differ between silent and replacement categories.

Divergence: Table 3 presents several estimates of net nucleotide and amino acid distances between the F and M lineages and a tree of the combined male and female data is shown in Figure 1c. The M and F lineages of *M. californianus* were highly differentiated; for the combined data, any randomly sampled F and M molecule differed, on average, by 530 mutations. Uncorrected *p*-distances at the nucleotide level exceeded 30% at all loci except *cox1*. Similarly high levels of divergence were observed at the amino acid level with overall Tamura-Nei distances exceeding 40%. More realistic models that allow for multiple hits and rate heterogeneity (among lineages) produced considerably higher estimates but did not affect the net rankings. Genetic distances estimated between genes tended to parallel patterns of polymorphism observed within genes; the *nad2* locus was the most differentiated and *cox1* the most highly constrained. Allowing for unequal substitution rates among sites increased the distances by nearly 50% (not shown).

Estimating a divergence time between the F and M mtDNA molecules of *M. californianus* is difficult because of high sequence divergence and the fact the homologous regions have not been studied in other mytilids. However, data for regions of the *nad4* and *cox3* genes of the F genome are available in GenBank for *M. trossulus*, *M. galloprovincialis*, and *M. edulis* and we have obtained homologous sequences for *M. californianus* (B. S. ORT and G. H. POGSON, unpublished data). Assuming a divergence time of 3.5 million years between *M. trossulus* and *M. edulis*/*M. galloprovincialis* (WARES and CUNNINGHAM 2001), we estimate a second position substitution rate of 3.53×10^{-9} for *nad4* and a first position rate of 6.34×10^{-9} for *cox3*. These provide rough estimates for the divergence between the *M. edulis* species complex and the *M. californianus* lineage at 7.64 and 7.61 million years ago (MYA) for *nad4* and *cox3*, respectively, and between the F and M genomes of *M. californianus* at 23.4 and 24.1 MYA for *nad4* and *cox3*, respectively. Given the many problems inherent in estimating divergence times

TABLE 2
Comparison of silent and replacement mutations

Locus	<i>N</i>	Silent mutations				Replacement mutations			
		Singletons	Total	Mean frequency	Tajima's <i>D</i>	Singletons	Total	Mean frequency	Tajima's <i>D</i>
Female									
<i>nad2</i>	94	47	91	0.0218	−2.56***	14	19	0.0151	−2.45**
<i>cox1</i>	92	23	45	0.0208	−2.51***	0	0	0	NA
<i>atp6</i>	88	22	32	0.0199	−2.50**	6	7	0.0146	−2.04*
<i>nad5</i>	91	53	72	0.0148	−2.79***	12	15	0.0125	−2.45**
	87	142	229	0.0193	−2.73***	33	41	0.0141	−2.72***
Male									
<i>nad2</i>	85	78	148	0.0301	−2.48**	11	20	0.0654	−1.42
<i>cox1</i>	87	36	63	0.0370	−2.22**	6	12	0.0201	−2.14**
<i>atp6</i>	92	36	75	0.0420	−2.15*	21	29	0.0184	−2.52***
<i>nad5</i>	87	43	94	0.0358	−2.30**	24	33	0.0218	−2.50**
	85	193	379	0.0350	−2.39**	62	94	0.0288	−2.45**

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

(e.g., Ho *et al.* 2005), these estimates must be viewed as provisional.

McDONALD and KREITMAN (1991) tests: All McDonald–Kreitman (MK) tests produced significant results (Table 4). At all loci, departures from neutral expectations were caused by excess numbers of fixed replacement mutations resulting in values of the neutrality index < 1 . With the exception of *cox1*, $> 50\%$ of the mutations that had fixed between the M and F genomes were amino acid replacements. This stands in marked contrast to the polymorphism data where replacement mutations were far less common than silent mutations. Combining data across all genes, the ratio of replacement to silent mutations (R:S ratio) fixed between molecules (55.0%) was roughly three times the R:S ratio observed for the polymorphism data (18.6%).

To further test for the importance of relaxed constraint in contributing to divergence between molecules, we compared the extent of replacement polymorphism in two classes of amino acid sites. The first were “constrained” sites that had not experienced an amino acid substitution between M and F. The second were “di-

verged” sites that had experienced at least one amino acid substitution between genomes. We then determined whether an amino acid site belonging to either class was monomorphic or polymorphic. Here, a diverged site was classified as polymorphic only if the amino acid mutation differed from the one(s) that had fixed between molecules. Table 5 shows that all four loci exhibited significantly higher polymorphism at diverged sites than at constrained sites. This result is consistent with the hypothesis that diverged amino acid codons experience, on average, reduced levels of purifying selection.

Estimation of mutation-rate scalars: Figure 3 presents mutation-rate scalars for F and M silent and replacement sites estimated by the IM program (NIELSEN and WAKELEY 2001; HEY and NIELSEN 2004). Convergence of mutation-rate estimates was assessed by replicating each simulation at least three times while including data from the same anonymous nuclear locus (*Mca125*) to serve as a standard (G. H. POGSON, unpublished data). The relative differences between M and F rate scalars were similar among replicates and those shown in Figure 3 represent typical results. At the *cox1*, *atp6*, and *nad5* loci, the male molecule exhibited 30–40% higher silent mutation rates than the F molecule. However, at the *nad2* locus, the silent mutation-rate scalar for M exceeded F by 69.0% and the probability that $\mu_M > \mu_F$ was estimated as 0.990. The posterior odds ratio that the male *nad2* silent mutation rate exceeds the female rate was 102:1. The M genome also exhibited higher replacement mutation rates than F. At the *nad2* and *atp6* loci, the male sequences showed a 23.0% increase in replacement mutation rate relative to females. At the *nad5* locus, the male replacement mutation rate exceeded the female rate by 47.0%. However, the probability that $\mu_M > \mu_F$ for *nad5* was not significant (0.883).

TABLE 3

Net nucleotide and amino acid divergence between genomes

Locus	Nucleotide		Amino acid	
	<i>p</i> -Distance	Tamura–Nei	<i>p</i> -Distance	JTT matrix
<i>nad2</i>	0.384	0.583	0.435	0.666
<i>cox1</i>	0.198	0.239	0.114	0.133
<i>atp6</i>	0.331	0.467	0.296	0.411
<i>nad5</i>	0.345	0.497	0.341	0.484
Combined	0.326	0.455	0.317	0.441

JTT, Jones–Taylor–Thornton matrix.

TABLE 4
McDonald and Kreitman tests

Locus	Mutation	Fixed	Polymorphic	G-test	Neutrality index
<i>nad2</i>	Replacement	131	32	118.30***	0.097
	Silent	82	206		
<i>cox1</i>	Replacement	17	12	8.17**	0.301
	Silent	38	89		
<i>atp6</i>	Replacement	66	34	14.64***	0.368
	Silent	65	91		
<i>nad5</i>	Replacement	93	32	59.70***	0.156
	Silent	63	139		
Combined	Replacement	306	117	175.43***	0.186
	Silent	250	513		

** $P < 0.01$, *** $P < 0.001$.

Comparison of amino acid substitutions: We made two predictions concerning the nature of amino acid substitutions (*i.e.*, Grantham distances) in the M and F genomes and their population frequencies. First, a relaxation of selective constraint on the M molecule should allow more nonconservative and radical amino acid mutations to reach higher frequencies in the male lineage. As predicted, a higher proportion of mutations in the M genome involved moderately radical or radical changes (*i.e.*, Grantham scores >100) compared to the F genome (13.8% *vs.* 4.9%; Figure 4). In the M molecule, the majority of mutations with frequencies >0.10 were conservative (*i.e.*, scores <50), confirming the action of strong purifying selection. The only exception was a S $\rightarrow$ L mutation at position 25 at the *atp6* locus with a Grantham score of 145. Second, we predicted inverse relationships between allele frequencies and Grantham distances of replacement mutations in both genomes (reflecting variation in the intensity of purifying selection). Although the M molecule exhibited the predicted negative relationship, the correlation was not significant in either genome (Figure 4).

No significant differences were observed between lineages in the numbers of amino acid polymorphisms classified as conservative, moderately conservative, or moderately radical (MR) and radical (R) combined (MR + R; $\chi^2 = 2.51$, d.f. = 2, $P = 0.285$; Table 6). However, there was a marginally significant difference between the numbers of fixed *vs.* polymorphic amino acid replacement mutations in the four Grantham categories ($\chi^2 = 8.10$, d.f. = 3, $P = 0.044$; Table 7). This was attributable to the larger proportion of moderately radical mutations that had fixed (18.9%) compared to those polymorphic in the sample (7.4%).

DISCUSSION

Despite a number of previous studies on the gender-associated mtDNA molecules in *Mytilus* spp., the relative roles of selection, mutation, and differential constraint

in determining the patterns of polymorphism and divergence between the genomes remain unclear. Our study has attempted to circumvent limitations of earlier reports by comparing four homologous gene regions of the F and M molecules (totaling 2307 bp) in large population samples from *M. californianus* that, unlike other mytilids, does not hybridize with any extant species. The two genomes of *M. californianus* were highly divergent (originating at least 20 MYA) and each possessed extensive levels of nucleotide polymorphism (haplotype diversities approached 100% in both lineages). As found for other *Mytilus* spp., the M molecule exhibited significantly higher levels of polymorphism (both silent and replacement) than the F molecule. Overall, nucleotide diversity in M was two to four times greater than in F, which is comparable to other mytilids (*e.g.*, SKIBINSKI *et al.* 1994; RAWSON and HILBISH 1995; STEWART *et al.* 1995; QUESADA *et al.* 1998b), unionid mussels (LIU *et al.* 1996), and venerid clams (PASSAMONTI *et al.* 2003). The unprecedented numbers of segregating mutations (274 and 473 in the combined F and M data, respectively) provide us with a unique ability to evaluate

TABLE 5
Comparison of polymorphism at diverged or constrained amino acid sites

Locus	Status	Classification		G-test
		Diverged	Constrained	
<i>nad2</i>	Polymorphic	22	8	12.26***
	Monomorphic	84	128	
<i>cox1</i>	Polymorphic	6	5	13.05***
	Monomorphic	12	125	
<i>atp6</i>	Polymorphic	15	11	10.41**
	Monomorphic	35	106	
<i>nad5</i>	Polymorphic	22	10	18.67***
	Monomorphic	51	129	
Combined	Polymorphic	65	34	54.39***
	Monomorphic	182	488	

** $P < 0.01$, *** $P < 0.001$.

FIGURE 3.—Estimates of mutation-rate scalars for (a) silent mutations and (b) replacement mutations in the F (dotted lines) and M genomes (solid lines) relative to an anonymous nuclear gene (*Mcal25*; dashed lines). Mutation rates were estimated by coalescent simulations using the IM program of NIELSEN and WAKELEY (2001) assuming a single panmictic population (*i.e.*, constraining t , m_1 , and m_2 to equal zero and $\theta_1 = \theta_2 = \theta_A$).

the relative importance of drift and selection in shaping patterns of polymorphism within and divergence between genomes.

Selective constraint: There are several features of our data that suggest that the F molecule experiences a higher degree of selective constraint than the M molecule but that the latter is still subject to strong purifying selection. First, the site frequency spectrum in both genomes was highly skewed toward significant excesses of low-frequency mutations that generated highly significant and negative Tajima's D values (Tables 1 and 2). Departures from neutral expectations were consistently stronger in F, suggesting that replacement mutations in the female lineage experience higher levels of purifying selection than those in the male lineage. Although a recent increase in effective population size may have contributed to the excess of rare polymorphism, this cannot account for differences in Tajima's D values because both genomes have experienced the same demographic history. Second, the majority of replacement polymorphisms in both lineages involved conservative or moderately conservative changes (*i.e.*, Grantham scores <100 ; Figure 4). Although about three times more moderately radical and radical replacement mutations were detected in the M molecule, only one (in *atp6*)

achieved a moderate population frequency ($P = 0.16$). The six remaining replacement mutations in the M lineage with frequencies >0.10 involved conservative changes (Grantham scores <32 ; Figure 4). In contrast, not a single replacement mutation in the F lineage achieved a population frequency >0.05 . These observations suggest that the vast majority of replacement mutations in both genomes are deleterious and that the magnitude of purifying selection is only slightly reduced in M.

Mutation: Following the initial suggestions by STEWART *et al.* (1995) and RAWSON and HILBISH (1995), both QUESADA *et al.* (1998) and SKIBINSKI *et al.* (1999) have argued that the *Mytilus* M genome might possess a higher rate of mutation than F, which contributes to its higher polymorphism and faster rate of divergence. The coalescent simulations undertaken in this study provide support for this hypothesis; mutation-rate scalars estimated for the four male genes exceeded females by 30–70% and 20–50% for synonymous and nonsynonymous mutations, respectively (Figure 3). However, if effective population sizes of the two molecules are similar, these differences appear far below that required to explain the two to four times higher level of polymorphism in M (assuming a strictly neutral

FIGURE 4.—Relationships between the allele frequencies and Grantham scores of amino acid replacement mutations in the (a) M and (b) F genomes.

TABLE 6

Comparison of Grantham scores of amino acid polymorphisms in F and M molecules

Classification	F molecule	M molecule	Total
Conservative	21	47	68
Moderately conservative	18	34	52
Moderately radical and radical	2	13	15
Total	41	94	135

$$\chi^2 = 2.51, \text{ d.f.} = 2, P = 0.285.$$

TABLE 7

Comparison of Grantham scores of fixed or polymorphic replacement mutations

Classification	Fixed	Polymorphic	Total
Conservative	78	68	146
Moderately conservative	58	52	110
Moderately radical	33	10	43
Radical	6	5	11
Total	175	135	310

$$\chi^2 = 8.10, \text{ d.f.} = 3, P = 0.044.$$

model). Our simulations also detected significant heterogeneity in silent, but not replacement, mutation rates among loci, suggesting that factors in addition to mutation are determining levels of polymorphism. It is important to note, however, that our approach to estimating mutation rates was phylogenetic, which is known to underestimate intrinsic mutation rates by one to two orders of magnitude (DENVER *et al.* 2000; HOWELL *et al.* 2003; HAAG-LIAUTARD *et al.* 2007). If the M genome experiences reduced levels of purifying selection, both replacement and silent mutations will achieve higher population frequencies providing the appearance of increased mutation rates (see Figure 2). This may not be due to intrinsic differences in mutation rate *per se* but to the diminished ability of selection to act on mildly deleterious mutations. We believe that our data fit this scenario and that a relaxation of selective constraint on M alone is sufficient to explain its elevated level of polymorphism and faster rate of evolution documented in earlier studies without the need to infer a higher mutation rate.

Tests of neutrality: Significant MK tests have been reported in previous studies comparing F and M sequences in *Mytilus* spp. (STEWART *et al.* 1996; QUESADA *et al.* 1998b, 1999) and in the manila clam, *Tapes philippinarum* (PASSAMONTI *et al.* 2003). However, the causes of departures from neutral expectations have varied among studies. For example, in North American populations of *M. edulis* and *M. trossulus*, STEWART *et al.* (1996) observed an excess number of fixed replacement mutations at the *cox3* gene. At the same locus, QUESADA *et al.* (1998b) detected a significant excess of replacement polymorphisms in Atlantic and Mediterranean populations of *M. galloprovincialis* but Atlantic *M. edulis* conformed closely to neutral expectations. RIGINOS *et al.* (2004) failed to detect significant MK tests at the *cox1* and *cox3* genes in comparisons between Atlantic and European *M. edulis* and *M. trossulus*. These discrepancies appear, in part, to be caused by the complex history of mtDNA introgression between *Mytilus* spp. that appears to differ between European and Atlantic regions (see RIGINOS *et al.* 2004).

In this study, we observed a highly significant excess of fixed replacement mutations at all four loci (Table 4),

similar to that reported by PASSAMONTI *et al.* (2003) who studied seven genes in *T. philippinarum*. However, unlike PASSAMONTI *et al.*'s (2003) study, more mutations in *M. californianus* were polymorphic within lineages than fixed between lineages and, in the latter group, replacement fixations outnumbered silent fixations. The large number of fixed amino acid replacements between lineages appears paradoxical in light of the strong signature of purifying selection in both genomes. If the vast majority of replacement mutations are mildly deleterious and maintained at low population frequencies by negative selection, how can so many have fixed between the two molecules?

Several factors may have contributed to this paradox. First, the high sequence divergence between lineages may have led us to underestimate the numbers of fixed silent differences. In the combined data, ~30% of amino acid positions (236 of 769) have experienced at least one replacement substitution between the two molecules ("diverged codons"). At 70% of diverged codons, there have been more than two nucleotide substitutions with the vast majority (88.5%) occurring at third positions. At these diverged amino acid sites, additional synonymous substitutions will go undetected in our sample, but silent polymorphisms will be counted. This will result in an underestimation of the numbers of fixed silent mutations and create an apparent excess of fixed replacements. A similar counting problem will occur at amino acid sites that are identical in the two molecules ("constrained codons") where 46% of synonymous sites have experienced at least one substitution between lineages. Although replacement substitutions at diverged codons will also be underestimated, the bias will not be as extreme as at synonymous sites due to the greater degree of saturation at the latter.

Second, fluctuations in effective population size may have contributed to the significant MK tests (EYRE-WALKER 2002). Population bottlenecks associated with Pleistocene glacial cycles may have accelerated the fixation of mildly deleterious replacement mutations that became effectively neutral in smaller populations. Furthermore, if mussel populations have experienced recent expansions in size, the frequencies of replacement mutations could be well below their historical means due to the increased efficacy of purifying selection. Under either scenario, the numbers of fixed amino acid replacements will be greater than expected from the contemporary polymorphism data. The highly significant negative Tajima's *D* values suggest recent population growth, but data from additional nuclear loci will be needed to test this possibility.

Finally, it is not possible for us to exclude some role played by adaptive evolution in the two lineages of *M. californianus*. Recently, BAZIN *et al.* (2006) have argued that the absence of a general relationship between mtDNA diversity and population size among animal groups might be explained by selective sweeps and suggest that

such events may be more common in invertebrates than vertebrates. The extensive levels of nucleotide polymorphism displayed by the M and F mitogenomes of *M. californianus* are clearly incompatible with recent selective sweeps. However, given the antiquity of the two lineages, it is not possible to exclude the possibility of historical sweeps. If a relaxation of selective constraint on the M molecule consistently maintains higher levels of polymorphism than F, then sweeps in the male lineage are expected to fix more mutations by genetic hitchhiking (both silent and replacement) compared to similar events occurring in the female lineage. In the four gene regions that we studied, any two randomly sampled M molecules differ, on average, by 4.9 replacements and 22.8 silent mutations (compared to 1.1 replacement and 9.2 silent mutations for the F molecule). If polymorphism in these regions is representative of other protein-coding loci, then (in the absence of recombination) a selective sweep in M would be expected to result in the mean fixation of ~24 amino acid substitutions and 111 silent changes (compared to 5 nonsynonymous and 45 synonymous mutations in F). If sweeps occur in both lineages with similar probabilities, the M molecule will exhibit a faster rate of evolution, as documented in a number of earlier studies (e.g., RAWSON and HILBISH 1995; HOEH *et al.* 1996).

Comparison with other mytilids: The F and M mtDNA molecules of *M. californianus* differ in a number of important ways from those previously characterized from species in the *M. edulis* complex. First, neither genome exhibited any detectable population structure over the 2000 km of Pacific coast sampled in our study. This is consistent with allozyme studies on *M. californianus* (LEVINTON and KOEHN 1976; LEVINTON and SUCHANEK 1978; ENGEL 2004) and other broadcast spawning marine invertebrates with similar geographic distributions in the northeast Pacific (see BURTON 1998), but differs from several reports documenting significantly greater population structuring for the M than for the F molecule (QUESADA *et al.* 1998a; SKIBINSKI *et al.* 1999; RIGINOS *et al.* 2004). Second, we failed to detect any masculinization events in which an F molecule has successfully invaded the M lineage. Although our study did not test for masculinization directly, we were able to successfully amplify genes from an M molecule in every individual identified as a male. This indicates that masculinized F molecules, if present, occur at population frequencies <1%. Our inability to detect such “role reversals” stands in marked contrast to masculinization frequencies of 33% in European *M. trossulus* (WENNE and SKIBINSKI 1995) and up to 40% in hybrids between American *M. trossulus* and *M. edulis* (ZOUROS *et al.* 1994), but is similar to that reported for unionid mussels by HOEH *et al.* (2002).

Third, unlike a number of recent studies (e.g., LADOUKAKIS and ZOUROS 2001; BURZYNSKI *et al.* 2003, 2006; RAWSON 2005; BRETON *et al.* 2006), no recombi-

nation was detected within or between the M and F molecules. Although both genomes of *M. californianus* failed the four-gamete test (suggesting recombination within both F and M), the finite-sites model implemented by the method of McVEAN *et al.* (2002) failed to detect significant recombination, suggesting that its apparent signal was caused by homoplasy. Finally, the magnitude of divergence between the F and M genomes of *M. californianus* was markedly higher than observed for other *Mytilus* species but lower than that exhibited among unionid mussels (HOEH *et al.* 2002). We estimate that the two mtDNA lineages of *M. californianus* last shared a common ancestor at least 20 MYA, which is three to four times older than the common ancestor of the standard M and F molecules in the *M. edulis* complex (~5.5 MYA) (RAWSON and HILBISH 1995).

The extreme divergence between the M and F mtDNA molecules of *M. californianus* and the absence of both masculinization events and population structuring could all result from the fact that this taxon, unlike others in the *M. edulis* complex, does not hybridize with any other extant species. Hybridization events are known to disrupt the transmission of DUI and result in the appearance of both males lacking an M molecule and in M-positive females (RAWSON *et al.* 1996; WOOD *et al.* 2003). Successful backcrosses involving F₁ hybrids have the potential to cause the introgression of mtDNA lineages among mussel species that has been previously documented on both historical (QUESADA *et al.* 1998a; RAWSON and HILBISH 1998) and contemporary time scales (SMIETANKA *et al.* 2004; see, however, SAAVEDRA *et al.* 1996). The absence of hybridization between *M. californianus* and other mytilids in the northeast Pacific could account for the high fidelity of DUI transmission and the large divergence observed between the F and M lineages. It is also possible that some threshold of divergence between the M and F lineages of *M. californianus* that reduces the success of masculinization events has been exceeded.

Broader implications: In agreement with a growing number of studies, our results suggest that the majority of amino acid replacement mutations that occur in the mtDNA molecule are mildly deleterious and subject to purifying selection. The strength of purifying selection in our study appears to be particularly strong in maintaining the vast majority of replacement mutations at low population frequencies but is clearly relaxed on the M genome. Despite their extensive sequence divergence, the M and F mtDNA molecules of *M. californianus* were remarkably similar in nucleotide composition and patterns of codon usage (Table 1), suggesting that both lineages have diversified largely through the sustained action of mutation pressure and random genetic drift.

Although DUI provides unique opportunities for studying sexual conflict and the relative importance of neutral/nearly neutral *vs.* adaptive evolution in the

mtDNA molecule, the resolution of these issues may require the functional characterization of mitochondria possessing the M or F genomes. For example, a relaxation of selective constraint on the M molecule should result in a more rapid accumulation of deleterious mutations (*i.e.*, Muller's ratchet) that should compromise its metabolic performance compared to F. At the same time, however, the uniparental transmission of the M molecule from father to son might facilitate the evolution of metabolic traits that benefit sperm performance. At present, there is no evidence of adaptations in M affecting sperm swimming performance (*e.g.*, EVERETT *et al.* 2004), and the frequent masculinization events observed in the *M. edulis* complex contradict that expected if the M molecule had evolved to enhance male fitness. Future studies quantifying the metabolic performance of mitochondria isolated from *Mytilus* spp. possessing gender-associated mtDNA molecules at varying levels of divergence might distinguish between these alternatives.

Conclusions: Purifying selection appears to play a dominant role in determining the levels and patterns of polymorphism in both the M and F mtDNA molecules of *M. californianus*. Both genomes experience strong levels of purifying selection but amino acid replacements appear to be slightly more deleterious in the F lineage. A relaxation of selective constraint on the M molecule appears capable of explaining its higher population frequencies of synonymous and nonsynonymous mutations without needing to invoke a higher intrinsic mutation rate. Both the extensive sequence divergence between the F and M genomes of *M. californianus* and the high fidelity of DUI in this species may be indirect consequences of its inability to hybridize with congeners in the northeast Pacific region. The relative importance of drift *vs.* selective sweeps in causing the excess numbers of fixed replacement difference observed between lineages remains unclear. We recommend that future studies combine investigations of polymorphism and divergence with metabolic characterizations of mitochondrial performance among mussel species with gender-associated mtDNA molecules exhibiting a range of sequence divergence.

We thank Matt Edwards and Jonna Engel for their assistance in collecting mussels, Rob Young for advice in running the coalescent simulations, and D. M. Rand and two anonymous reviewers for providing helpful comments for improving the manuscript. Funding for the study was provided by the Myers Oceanographic Trust, the Western Malacological Society, the Partnership for the Interdisciplinary Study of Coastal Oceans, and the National Science Foundation (OCE-0350443 and DEB-0412976).

LITERATURE CITED

- AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB *et al.*, 1987 Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* **18**: 489–522.
- BALLARD, J. W. O., and M. C. WHITLOCK, 2004 The incomplete natural history of mitochondria. *Mol. Ecol.* **13**: 729–744.
- BAZIN, E., S. GLEMIN and N. GALTIER, 2006 Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**: 570–572.
- BEAGLEY, C. T., K. A. TAYLOR and D. R. WOSTENHOLME, 1997 Gender-associated diverse mitochondrial DNA molecules of the mussel, *Mytilus californianus*. *Curr. Genet.* **31**: 318–324.
- BRETTON, S., G. BURGER, D. T. STEWART and P. U. BLIER, 2006 Comparative analysis of gender-associated complete mitochondrial genomes in marine mussels (*Mytilus* spp.). *Genetics* **172**: 1107–1119.
- BROWN, W. M., M. GEORGE and A. C. WILSON, 1979 Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **76**: 1967–1971.
- BURTON, R. S., 1998 Intraspecific phylogeography across the Point Conception biogeographic boundary. *Evolution* **52**: 734–745.
- BURZYNSKI, A., M. ZBAWICKA, D. O. F. 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.
- BURZYNSKI, A., M. ZBAWICKA, D. O. F. SKIBINSKI and R. WENNE, 2006 Doubly uniparental inheritance is associated with high polymorphism for rearranged and recombinant control region haplotypes in Baltic *Mytilus trossulus*. *Genetics* **174**: 1081–1094.
- CLARK, A. G., and E. M. LYCKEGAARD, 1988 Natural selection with nuclear and cytoplasmic transmission. III. Joint analysis of segregation and mtDNA in *Drosophila melanogaster*. *Genetics* **118**: 471–481.
- DENVER, D. R., K. MORRIS, M. LYNCH, L. L. VASSILIEVA and W. K. THOMAS, 2000 High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. *Science* **289**: 2342–2344.
- ENGEL, J. D., 2004 Population genetic structure of the California sea mussel: influence of the Pleistocene, biogeography, and micro-evolutionary processes. Ph.D. Thesis, University of California, Santa Cruz, CA.
- EVERETT, E. M., P. J. WILLIAMS, G. GIBSON and D. T. STEWART, 2004 Mitochondrial DNA polymorphisms and sperm motility in *Mytilus edulis* (Bivalvia; Mytilidae). *J. Exp. Zool.* **301**: 906–910.
- EYRE-WALKER, A., 2002 Changing effective population size and the McDonald-Kreitman test. *Genetics* **162**: 2017–2024.
- FISHER, C., and D. O. F. SKIBINSKI, 1990 Sex-biased mitochondrial heteroplasmy in the marine mussel *Mytilus*. *Proc. R. Soc. Lond. B Biol. Sci.* **242**: 149–156.
- GRANT, W. S., I. B. SPIES and M. F. CANINO, 2006 Biogeographic evidence for selection on mitochondrial DNA in the north Pacific walleye pollock *Theragra chalcogramma*. *J. Hered.* **97**: 571–580.
- GRANTHAM, R., 1974 Amino acid difference formula to help explain protein evolution. *Science* **185**: 862–864.
- HAAG-LIAUTARD, C., M. DORRIS, X. MASIDE, S. MACASKILL, D. L. HALLIGAN *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* **445**: 82–85.
- HEY, J. and R. NIELSEN, 2004 Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**: 747–760.
- HO, S. Y. W., M. J. PHILLIPS, A. COOPER and A. J. DRUMMOND, 2005 Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* **22**: 1561–1568.
- HOEH, W. R., D. T. STEWART, B. W. SUTHERLAND and E. ZOUIROS, 1996 Multiple origins of gender-associated mitochondrial DNA lineages in bivalves (Mollusca: Bivalvia). *Evolution* **50**: 2276–2286.
- HOEH, W. R., D. T. STEWART and S. I. GUTTMAN, 2002 High fidelity of mitochondrial genome transmission under the doubly uniparental mode of inheritance in freshwater mussels (Bivalvia: Unionidae). *Evolution* **56**: 2252–2261.
- HOFFMAN, R. J., J. L. BOORE and W. M. BROWN, 1992 A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. *Genetics* **131**: 397–412.
- HOWELL, N., C. B. SMEJKAL, D. A. MACKAY, P. F. CHINNERY, D. M. TURNBULL *et al.*, 2003 The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *Am. J. Hum. Genet.* **75**: 659–670.

- HUDSON, R. R., and N. L. KAPLAN, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**: 147–164.
- JAMES, A. C., and J. W. O. BALLARD, 2003 Mitochondrial genotype affects fitness in *Drosophila simulans*. *Genetics* **164**: 187–194.
- KUMAR, S., K. TAMURA and M. NEI, 2004 MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* **5**: 150–163.
- LADOUKAKIS, E. D., and E. ZOUROS, 2001 Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. *Mol. Biol. Evol.* **18**: 1168–1175.
- LEVINTON, J. S., and R. K. KOEHN, 1976 Population genetics of mussels, pp. 357–384 in *Marine Mussels: Their Ecology and Physiology*, edited by B. L. BAYNE. Cambridge University Press, Cambridge, UK.
- LEVINTON, J. S., and T. H. SUCHANEK, 1978 Geographic variation, niche breadth and genetic differentiation at different geographic scales in the mussels *Mytilus californianus* and *M. edulis*. *Mar. Biol.* **49**: 363–375.
- LI, H., and W. STEPHAN, 2006 Inferring the demographic history and rate of adaptive substitution in *Drosophila*. *PLoS Genet.* **2**: 1580–1589.
- LI, W.-H., C.-I. WU and C.-C. LUO, 1985 A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**: 150–174.
- LIU, H. P., J. B. MITTON and S. K. WU, 1996 Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the freshwater mussel, *Anodonta grandis grandis*. *Evolution* **50**: 952–957.
- MCDONALD, J. H., and M. KREITMAN, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.
- MCVEAN, G., P. AWADALLA and P. FEARNHEAD, 2002 A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics* **160**: 1231–1241.
- NACHMAN, M. W., 1998 Deleterious mutations in animal mitochondrial DNA. *Genetica* **103**: 61–69.
- NACHMAN, M. W., W. M. BROWN, M. STONEKING and C. F. AQUADRO, 1996 Nonneutral mitochondrial DNA variation in humans and chimpanzees. *Genetics* **142**: 953–963.
- NIELSEN, R., and J. WAKELEY, 2001 Distinguishing isolation from migration: a Markov chain Monte Carlo approach. *Genetics* **158**: 885–896.
- PASSAMONTI, M., and V. SCALI, 2001 Gender-associated mitochondrial DNA heteroplasmy in the venerid clam *Tapes philippinarum* (Mollusca: Bivalvia). *Curr. Genet.* **39**: 117–124.
- PASSAMONTI, M., J. L. BOORE and V. SCALI, 2003 Molecular evolution and recombination in gender-associated mitochondrial DNAs of the Manila clam *Tapes philippinarum*. *Genetics* **164**: 603–611.
- POGSON, G., H., K. A. MESA and R. G. BOUTILIER, 1995 Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. *Genetics* **139**: 375–385.
- POSADA, D., and K. CRANDALL, 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.
- QUESADA, H., R. WENNE and D. O. F. SKIBINSKI, 1995 Differential introgression of mitochondrial DNA across species boundaries within the marine mussel genus *Mytilus*. *Proc. R. Soc. Lond. B Biol. Sci.* **262**: 51–56.
- QUESADA, H., C. GALLAGHER, D. A. G. SKIBINSKI and D. O. F. SKIBINSKI, 1998a Patterns of polymorphism and gene flow of gender-associated mitochondrial DNA lineages in European mussel populations. *Mol. Ecol.* **7**: 1041–1051.
- QUESADA, H., M. WARREN and D. O. F. SKIBINSKI, 1998b Nonneutral evolution and differential mutation rate of gender-associated mitochondrial DNA lineages in the marine mussel *Mytilus*. *Genetics* **149**: 1511–1526.
- QUESADA, H., R. WENNE and D. O. F. SKIBINSKI, 1999 Interspecies transfer of female mitochondrial DNA is coupled with role-reversals and departure from neutrality in the mussel *Mytilus trossulus*. *Mol. Biol. Evol.* **16**: 655–665.
- RAMBAUT, A., and A. DRUMMOND, 2005 TRACER, version 1.3. (<http://evolve.zoo.ox.ac.uk>).
- RAND, D. M., 2001 The units of selection on mitochondrial DNA. *Annu. Rev. Ecol. Syst.* **32**: 415–448.
- RAND, D. M., and L. M. KANN, 1996 Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol. Biol. Evol.* **13**: 735–748.
- RAND, D. M., and L. M. KANN, 1998 Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. *Genetica* **103**: 393–407.
- RAWSON, P. D., 2005 Nonhomologous recombination between the large unassigned region of the male and female mitochondrial genomes in the mussel, *Mytilus trossulus*. *J. Mol. Evol.* **61**: 717–732.
- RAWSON, P. D., and T. J. HILBISH, 1995 Evolutionary relationships among the male and female mitochondrial DNA lineages in the *Mytilus edulis* species complex. *Mol. Biol. Evol.* **12**: 893–901.
- RAWSON, P. D., and T. J. HILBISH, 1998 Asymmetric introgression of mitochondrial DNA among European populations of blue mussels (*Mytilus spp.*). *Evolution* **52**: 100–108.
- RAWSON, P. D., C. L. SECOR and T. J. HILBISH, 1996 The effects of natural hybridization on the regulation of doubly uniparental mtDNA inheritance in blue mussels (*Mytilus spp.*). *Genetics* **144**: 241–248.
- RIGINOS, C., M. J. HICKERSON, C. M. HENZER and C. W. CUNNINGHAM, 2004 Differential patterns of male and female mtDNA exchange across the Atlantic Ocean in the blue mussel, *Mytilus edulis*. *Evolution* **58**: 2438–2451.
- RONQUIST, F., and J. P. HUELSENBECK, 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.
- ROZAS, J., J. C. SANCHEZ-DELBARRIO, X. MESSEGUER and R. ROZAS, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- RUIZ-PESINI, E., D. MISHMAR, M. BRANDON, V. PROCACCIO and D. C. WALLACE, 2004 Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* **303**: 223–226.
- SAAVEDRA, C., D. T. STEWART, R. R. STANWOOD and E. ZOUROS, 1996 Species-specific segregation of gender-associated mitochondrial DNA types in an area where two mussel species (*Mytilus edulis* and *M. trossulus*) hybridize. *Genetics* **143**: 1359–1367.
- SCHMIDT, T. R., W. WU, M. GOODMAN and L. I. GROSSMAN, 2001 Evolution of nuclear- and mitochondrial-encoded subunit interaction in cytochrome c oxidase. *Mol. Biol. Evol.* **18**: 563–569.
- SCHNEIDER, S., D. ROESSLI and L. EXCOFFIER, 2000 *Arlequin (Version 2.000): A Software for Population Genetic Analyses*. Genetics and Biometry Laboratory, University of Geneva, Geneva.
- SKIBINSKI, D. O. F., C. GALLAGHER and C. M. BEYNON, 1994 Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. *Genetics* **138**: 801–809.
- SKIBINSKI, D. O. F., C. GALLAGHER and H. QUESADA, 1999 On the roles of selection, mutation and drift in the evolution of mitochondrial DNA diversity in British *Mytilus edulis* (Mytilidae; Mollusca) populations. *Biol. J. Linn. Soc. Lond.* **68**: 195–213.
- SMIETANKA, B., M. ZBAWICKA, M. WOLOWICZ and R. WENNE, 2004 Mitochondrial DNA lineages in the European populations of mussels (*Mytilus spp.*). *Mar. Biol.* **146**: 79–92.
- STEWART, D. T., C. SAAVEDRA, R. R. STANWOOD, A. O. BALL and E. ZOUROS, 1995 Male and female mitochondrial DNA lineages in the blue mussel (*Mytilus edulis*) species group. *Mol. Biol. Evol.* **12**: 735–747.
- STEWART, D. T., E. R. KENCHINGTON, R. K. SINGH and E. ZOUROS, 1996 Degree of selective constraint as an explanation of the different rates of evolution of gender-specific mitochondrial DNA lineages in the mussel *Mytilus*. *Genetics* **143**: 1349–1357.
- TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- WARES, J. P., and C. W. CUNNINGHAM, 2001 Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* **55**: 2455–2469.
- WENNE, R., and D. O. F. SKIBINSKI, 1995 Mitochondrial DNA heteroplasmy in European populations of the mussel *Mytilus trossulus*. *Mar. Biol.* **122**: 619–624.
- WILLETT, C. S., and R. S. BURTON, 2004 Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. *Mol. Biol. Evol.* **21**: 443–453.
- WILLIAMSON, S. H., R. HERNANDEZ, A. FLEDEL-ALON, L. ZHU, R. NIELSEN *et al.*, 2005 Simultaneous inference of selection and population

- growth from patterns of variation in the human genome. Proc. Natl. Acad. Sci. USA **102**: 7882–7887.
- WOOD, A. R., G. TURNER, D. O. F. SKIBINSKI and A. R. BEAUMONT, 2003 Disruption of doubly uniparental inheritance of mitochondrial DNA in hybrid mussels (*Mytilus edulis* × *M. galloprovincialis*). Heredity **91**: 354–360.
- ZOUROS, E., A. O. BALL, C. SAAVEDRA and K. R. FREEMAN, 1994 An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. Proc. Natl. Acad. Sci. USA **91**: 7463–7467.
- Communicating editor: D. M. RAND

APPENDIX

PCR primers and amplification conditions

Gene	Primer name and sequence (5'–3')	Annealing temperature	Extension time
F molecule			
<i>nad2</i>	F: ATGGTAAGTTTAGTAATAAGGCC R: GAAGTTACTAACATTCATATAGC	52°	50 sec
<i>cox1</i>	F: GGGATTAACTGGGGTGATGTTG R: AAGCCCTTTYCCATTAAGGCG	52°	60 sec
<i>atp6</i>	F: GGTATCTATGCCGTTGTGGG R: ACTCCTCCRCCTATYAAGAAGCC	52°	50 sec
<i>nad5</i>	F: GTGAGGCTAGGATTTGYTTGGC R: CACTGCGCTTGAACCAGCTAAT	52°	60 sec
M molecule			
<i>nad2</i>	F: CTAATAATGAGAGTGGTGAGGCCC R: CAGAAAAGTTTACCGCCAARACC	54°	50 sec
<i>cox1</i>	F: CTGGGATTATACTATCTAGTGCGTCC R: TAAYACCAAGCAACCCAGATG	52°	60 sec
<i>atp6</i>	F: GATGCGCACAGGTATAATCTTGTG R: AACTTCCGCTCCTCTAAACACC	52°	50 sec
<i>nad5</i>	F: CTCTATGTATGGCTGTTTGTGAGGC R: CCTGCCACCAGCAAAGTTAAACT	52°	60 sec