Mytilus mitochondrial DNA haplotypes in southeastern Australia

Donald J. Colgan*, Peter Middelfart

Research Branch, The Australian Museum, 6 College Street, Sydney 2010, Australia

ABSTRACT: Mitochondrial DNA (mtDNA) analyses suggest that Southern Hemisphere mussels of the Mytilus edulis species complex include an endemic lineage and invasive genetic elements derived from the Northern Hemisphere species M. galloprovincialis and M. edulis. Haplotypes from these species have been reported in substantial frequencies from South America, South Africa, Western Australia and New Zealand, but only at very low frequencies in Tasmania and the coast near Melbourne in Victoria, the only regions studied in southeastern Australia. Here we investigate cytochrome c oxidase sequences from a wider region in southeastern Australia. Substantial frequencies of Northern Hemisphere Mytilus haplotypes of the female form of the mtDNA were observed, particularly in southern New South Wales. These haplotypes represented a broad range of divergent lineages among the Northern Hemisphere clades. All 3 haplotypes of the male form of mtDNA belonged to the lineage previously identified as the ‘European’ clade of M. edulis.

KEY WORDS: Mytilus · Invasive species · Introgression · Mitochondrial DNA

INTRODUCTION


Interpreting these patterns is complicated by the unusual method of mitochondrial inheritance in Mytilus. Males have 2 forms of mitochondrial genomes, 1 passed through the male line (M-type) and 1 inherited from the mother (F-type) (Skibinski et al. 1994, Hoeh et al. 1997). Females have only the F-type, but recombination between the types does occur, and the pre-existing M-type can be supplanted by the F-type to form a new M-type in a process called ‘masculinisation’ (Hoeh et al. 1997).

Hilbish et al. (2000) found 3 distinct sets of haplotypes in 16S rRNA F-type sequences from female Southern Hemisphere Mytilus spp. One consisted of haplotypes that were closely related to the Northern Hemisphere species characteristic of M. galloprovincialis. A second, found only in the Southern Hemisphere, comprised haplotypes that were sufficiently distinct from the first to suggest that they had been separated during the Pleistocene. The third group consisted of haplotypes similar to those found in M. edulis, some of which also occur, apparently as a result of introgression, in M. galloprovincialis. Gérard et al. (2008), using cytochrome c oxidase I (COI) sequences, identified 2 clades in Southern Hemisphere mussels. The first was denoted the northern clade (N), and is composed entirely of haplotypes found in either or both M. edulis and M. galloprovincialis. The second was denoted the southern clade (S) and was endemic to the Southern Hemisphere. It was differentiated into 3 reciprocally monophyletic sub-clades: S1 from South

*Email: don.colgan@austmus.gov.au © Inter-Research 2011 · www.int-res.com
America and Kerguelen, S2 from New Zealand and S3 from Tasmania. COI analyses suggested that the S clade, comprising all 3 sub-clades, is the sister group of the N clade comprising all the haplotypes from *M. edulis* and *M. galloprovincialis* (Gérard et al. 2008).

Although haplotypes of the N clade have been observed in many parts of the Southern Hemisphere, including South America, South Africa, Western Australia and New Zealand (Gérard et al. 2008), they were not observed in the Tasmanian populations examined by Hilbish et al. (2000) or Gérard et al. (2008). Recently, while this manuscript was in review, Westfall & Gardner (2010) found Northern Hemisphere 16S rRNA haplotypes at a frequency of about 15% in southeastern Australia (principally Tasmania and southern Victoria). The survey of COI sequences in the present study was conducted, in light of the earlier results, to collect data from a broader region in order to obtain a better estimate of the contribution of Northern Hemisphere *Mytilus* spp. mtDNA to mussel populations in southeastern Australia. Individuals of both sexes were included in the survey, and type-specific primers were developed to confirm the identity of M-type sequences and to generate F-type sequences from males.

Nuclear genetic contributions from Northern Hemisphere *Mytilus* spp. to Tasmanian populations have been suggested by allozyme data (McDonald et al. 1991) and genomic DNA size polymorphisms, particularly in the intron 1 of the *mac-I* locus (Borsa et al. 2007). Borsa et al. (2007) interpret their data as indicating locus-specific levels of introgression of Northern Hemisphere genetic material into Southern Hemisphere populations. They noted that it was not possible to determine from their data (Borsa et al. 2007) whether the putative *M. edulis* mac-1 variants in the Southern Hemisphere were derived from an invasion of the species itself or from introgression into *M. galloprovincialis* prior to the invasion of that species. We sequenced mac-1 intron 1 for the specimens for which we have mitochondrial DNA sequences and which were potentially homozygous, as judged by the presence of single bands in PCR amplifications, in an attempt to confirm the introgression of Northern Hemisphere alleles in southeastern Australian *Mytilus* spp.

**MATERIALS AND METHODS**

**Maintenance and experimentation.** After collection (Table 1) specimens were generally stored in 95% ethanol or frozen at −80°C. DNA was extracted from mantle tissue using DNeasy Spin columns (Qiagen).

<table>
<thead>
<tr>
<th>Locality</th>
<th>Site description</th>
<th>Registration number</th>
<th>Sample sizes for F-type COI</th>
<th>Accession numbers</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tura Head</td>
<td>E-SE-facing, steeply sloping, very exposed platform</td>
<td>AMS C.469469</td>
<td>8</td>
<td>HQ864835 (M) to HQ864847 to HQ864854</td>
<td>36°51.341’ S</td>
<td>149°56.992’ E</td>
</tr>
<tr>
<td>Haycock Headland</td>
<td>S-facing jagged rocks with high splash zone</td>
<td>AMS C.469468</td>
<td>4</td>
<td>HQ864834 (M) to HQ864844 to HQ864862 (A)</td>
<td>36°57.025’ S</td>
<td>149°56.461’ E</td>
</tr>
<tr>
<td>Bittangabee Bay</td>
<td>Rock block formation facing 50° E-NE, more sheltered</td>
<td>AMS C.469466</td>
<td>1</td>
<td>HQ864855 to HQ864840</td>
<td>37°12.922’ S</td>
<td>150°01.164’ E</td>
</tr>
<tr>
<td>Green Cape</td>
<td>S-SW-facing high energy rock formation</td>
<td>AMS C.469467</td>
<td>4</td>
<td>HQ864833 (M) to HQ864841 to HQ864843</td>
<td>37°15.758’ S</td>
<td>150°03.051’ E</td>
</tr>
<tr>
<td>Snug</td>
<td>Rocks at The Esplanade, sheltered</td>
<td>AMS C.469470</td>
<td>4</td>
<td>HQ864836 to HQ864839 to HQ891001</td>
<td>43°3’ 58” S</td>
<td>147°15’ 53” E</td>
</tr>
<tr>
<td>Triabunna</td>
<td>Aquaculture sample provided by Spring Bay Seafoods</td>
<td>AMS C.469465</td>
<td>1</td>
<td>HQ864856 to HQ864858 (A)</td>
<td>42°31’ S</td>
<td>147°55’ E</td>
</tr>
</tbody>
</table>
The specimens from Triabunna were stored in RNAlater (Applied Biosystems/Ambion) at room temperature for 18 mo. During this period, the shells dissolved and crystals, presumably of calcium carbonate, formed in the solution. DNA was extracted from the remaining tissue, but only 1 specimen could be successfully amplified for the genes studied here.

PCR amplifications of mitochondrial COI were performed at an annealing temperature of 43°C using the primers constructed by Folmer et al. (1994), following the methods of Colgan & da Costa (2009). PCR products were purified using ExoSAP-IT (USB) and sequenced by Macrogen. Both sequence directions were determined, using the individual primers from the original reaction.

M- and F-type specific primers were designed from the alignments compiled for the present study. These were then used separately in combination with Primer 1490 by Folmer et al. (1994), to amplify type-specific products. Cycling parameters were the same as those used for COI above. The type-specific primers were: (male) MCOIR: AGR GCT GCA GTG ATC CTR ATT and (female) FCOIR: CAATAC GGC AGTAAC TCT RAT C.

Amplification of mac-1 intron 1 using the primers by Daguin & Borsa (1999) was performed, in general, as it was for COI, except that the annealing temperature was increased to 48°C.

Analytical methods. Chromatograms were checked, and sequences were assembled using Sequencher v. 4.9 (Gene Codes). Alignments were made using CLUSTALX (Thompson et al. 1997), using the default parameters, and were inspected in BioEdit (Hall 1999).

We compiled a large range of Mytilus spp. sequences from GenBank to add to the data collected in the present study (Table 2, see Figs. 1 & 2 for accession numbers for the sequences used in the analyses of F- and M-type sequences, respectively). The data were initially analysed as a single set, but were subsequently separated into male and female data sets for detailed analyses. The ends of the male alignment were excluded from these, as many GenBank accessions were shorter than the product of the primers by Folmer et al. (1994).

Maximum-parsimony analyses for the COI data were conducted with PAUP v. 4.0 (Swofford 2003) using heuristic searches of 1000 replicates with random stepwise addition, keeping no more than 200 trees of length ≥400 steps in each replicate. Bootstrap searches were conducted with 500 pseudoreplicates, with 20 random addition replicates in each, and keeping no more than 200 trees of length >400. Bayesian analyses were performed with MRBAYES 3.1.2 (Huelsenbeck & Ronquist 2001), with parameters estimated separately for each codon position. The ‘ratepr’ was set to ‘variable’, the ‘invgamma’ option was invoked, and NST was set to 6 (for the GTR model). The analysis was run for 10 million generations with a tree sampled after every 1000 generations. Ten percent of the sampled trees were discarded as ‘burn-in’ using the ‘sumt’ command. All non-discarded trees from both runs by MRBAYES were included in the calculation of posterior probabilities.

Table 2. Mytilus spp. Data sources for additional sequences deposited in GenBank. The columns designate whether the sequence was M- or F-type, the GenBank accession number or the lowest in the range of numbers, the highest accession number and the source. Note that not all accessions in the ranges were included in the analyses.

<table>
<thead>
<tr>
<th>Sequence type</th>
<th>Accession number or lowest in range</th>
<th>Highest accession number in range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-type</td>
<td>U68775</td>
<td></td>
<td>Hoeh et al. (1997)</td>
</tr>
<tr>
<td>F-type</td>
<td>AY130013</td>
<td>AY130067</td>
<td>Riginos et al. (2004)</td>
</tr>
<tr>
<td>M-type</td>
<td>AY101397</td>
<td>AY101420</td>
<td>Riginos et al. (2004)</td>
</tr>
<tr>
<td>M-type</td>
<td>AY713946</td>
<td>AY713960</td>
<td>Riginos et al. (2004)</td>
</tr>
<tr>
<td>M-type</td>
<td>AY363687</td>
<td></td>
<td>Mizi et al. (2005)</td>
</tr>
<tr>
<td>F-type</td>
<td>AY497292</td>
<td></td>
<td>Mizi et al. (2005)</td>
</tr>
<tr>
<td>M-type</td>
<td>DQ198225</td>
<td></td>
<td>Burzyński et al. (2006)</td>
</tr>
<tr>
<td>F-type</td>
<td>DQ198231</td>
<td></td>
<td>Burzyński et al. (2006)</td>
</tr>
<tr>
<td>F-type</td>
<td>AM905146</td>
<td>AM905225</td>
<td>Gérard et al. (2008)</td>
</tr>
<tr>
<td>M-type, except</td>
<td>EU018170 and EU018184 (F-type)</td>
<td>EU018148</td>
<td>Jha et al. (2008)</td>
</tr>
<tr>
<td>EU018170 and EU018184 (F-type)</td>
<td>EU018204</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-type, except</td>
<td>GU570473</td>
<td>GU570522</td>
<td>Lesser et al. (2010)</td>
</tr>
<tr>
<td>GU570514 (F-type)</td>
<td>DQ864379</td>
<td>DQ864421</td>
<td>P. N. Lewis &amp; S. N. Jarman (unpubl.)</td>
</tr>
<tr>
<td>F-type</td>
<td>EU332488</td>
<td>EU332510</td>
<td>S. Breton et al. (unpubl.)</td>
</tr>
<tr>
<td>M-type except</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU332506 (F-type)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RESULTS

The sequences of 3 individuals generated by the Folmer et al. (1994) primers were most similar to the M-type sequences in GenBank. These were the only 3 individuals in which a product was found when the MCOIR primer was used in PCR amplification. For all 3 individuals, the M-type-specific sequences were identical to the original sequence for that specimen. When the FCOIR primer was used, single-banded amplification products were found for all of the mussels examined here. The F-type sequences found for the individuals originally returning only M-type sequences were added to the F-type data set.

The F-type alignment comprised 638 base positions, including 488 that were invariant, 27 that were variable, but not parsimony informative, and 123 that were parsimony informative. The aligned M-type data set comprised 549 bases, with 368 bases that were invariant, 77 that were variable, but not parsimony informative, and 104 that were parsimony informative. No insertions or deletions were inferred in either alignment.

There were 34,800 maximum-parsimony trees in the analyses of the F-type alignment, each of length 247, with a consistency index (CI) of 0.676. In the Bayesian analysis, the average standard deviation of the ‘splits frequency’ for the 2 independent Monte Carlo runs performed by MRBAYES was 0.012, and the ‘potential scale reduction factor’ (PSRF) for all parameters approached 1.00, indicating that the runs were converging. The averages of the likelihoods of the trees remaining after ‘burn-in’ were −2307.184 and −2306.128.

Of the 22 sequences collected here, 15 belonged to the N clade. These were found in Tasmania (1 individual) and in southern New South Wales (NSW; 14 individuals) (Fig. 1). The other 7 sequences, comprising 3 from southern NSW and 4 from Tasmania, all belonged to the S3 sub-clade of Gérard et al. (2008). In total, 1 of 5 Tasmanian individuals had an N clade haplotype as did 14 of 17 sequences from NSW. A χ² test of homogeneity for these frequencies gave a value of 6.92 (p < 0.01).

Southeastern Australian haplotypes were represented in all of the major lineages in the N clade, including 2 distinct groups whose sequences have previously been referred to as Mytilus galloprovincialis (Fig. 1). These were well-supported monophyletic groups in parsimony and Bayesian analyses, although they were not shown to be sister groups. Only 1 of the Australian individuals from the N clade had a sequence that was not also found in other countries. Three haplotypes were found in multiple populations in southeastern Australia. One of these was observed in specimens from both Snug (Tasmania) and Green Cape (NSW).

There were 15,200 maximum-parsimony trees in analyses of the M-type alignment, each 390 steps long, with a CI of 0.59. In the Bayesian analysis, the average standard deviation of the ‘splits frequency’ was 0.012, and the PSRF approached 1.00 for all parameters. The averages of the likelihoods of the trees remaining after ‘burn-in’ were −2766.30 and −2765.928.

All male Mytilus spp. haplotypes sequenced in the present study belonged to the ‘European’ lineage of M. edulis identified by Riginos et al. (2004). This lineage is also found at low frequency in North America (Fig. 2). Notably, 1 of the F-type sequences from these individuals apparently belonged to the S3 sub-clade (Green Cape 10), and the other 2 to divergent sub-clades of the N clade.

Six specimens were homozygous for mac-1 intron 1. These represented 4 haplotypes, 1 of which was observed in 3 individuals, and 3, in 1 individual each. Two of these differed at only 1 base in an alignment of 256 bases. The maximum number of nucleotide differences between any Australian haplotypes was 4, but there were at least 17 differences between any of these and the Northern Hemisphere mac-1b allele. An insertion of 27 bases that was observed in all 6 Australian sequences was not present in any reported Northern Hemisphere sequences (Ohresser et al. 1997). One of the Australian sequences also had the large insertion that distinguishes the mac-1b allele from other Northern Hemisphere sequences. The COI haplotypes of 3 of the sequenced individuals belonged to the N clade, and 2 to the S3 sub-clade (Triabunna 03 was not scored for COI).

DISCUSSION

Three main results were observed. First, Northern Hemisphere haplotypes are present at high frequencies in F-type Mytilus spp. mtDNA in at least 1 region of southeastern Australia (southern NSW). Second, all 3 M-type sequences collected here represented the ‘European’ M. edulis clade identified by Riginos et al. (2004). Third, a wide diversity of F-type lineages from the Northern Hemisphere is present in southeastern Australian populations.

Substantial frequencies of both male and female Northern Hemisphere Mytilus spp. mitochondrial haplotypes were observed in southern NSW. This contrasts with previous studies of Tasmanian populations, in which they were observed at low frequency (Westdall & Gardner 2010) or not at all (Hilbish et al. 2000, Gérard et al. 2008). The inter-regional differentiation in southeastern Australia may be due to a number of factors. There may be environmental differences between the regions that are associated with differential
Fig. 1. *Mytilus* spp. An example of the maximum parsimony tree for F-type *Mytilus* spp. haplotypes. The tree is drawn to scale and rooted on the *M. trossulus* female haplotypes. GenBank sequences are specified by accession numbers (references listed in Table 2). Sequences determined for the present study are indicated by an asterisk. Specimens in the Northern Hemisphere haplotype clade are denoted as Mg (for *M. galloprovincialis*) or Me (for *M. edulis*) if identifications were made by the corresponding authors in Table 2. Nodes seen in the strict consensus of the maximum-parsimony trees are indicated by filled circles. Bootstrap support percentages >50 and Bayesian posterior probabilities >0.50 are written near branches or to the right of bars linking the relevant specimens. Locality abbreviations used in specimen identifications are as follows — Afr: South Africa; Aus: Australia; Can: Canada; Ch: Chile; Fin: Finland; Fr: France; Gr: Greece; It: Italy; Ker: Kerguelen Islands; NSW: New South Wales; NZ: New Zealand; Pol: Poland; Tas: Tasmania; USA: (northeastern) United States of America; WA: Western Australia. The inset shows the location of the sample sites in southeastern Australia.
One possibility is that there are differences in the habitats of the regions that differentially affect the persistence of native and introduced Mytilus spp. mtDNA types. Wave exposure might be one such difference. But the occurrence of haplotypes from the S3 sub-clade and from a wide range of sub-clades of the N clade at Tura Head suggests that is not the complete explanation. The high frequency of Northern Hemisphere haplotypes in southern NSW may be associated with shipping traffic to ports in the area during their history as whaling stations in the 19th century, some aspects of which are discussed by Pearson (1983).

Both M- and F-type Mytilus edulis mitochondrial genes are found introgressed in Mytilus galloprovincialis in Europe (Rawson & Hilbish 1998). Consequently, it is difficult to know whether the southeastern Australian haplotypes that are closely related to specimens identified as M. edulis in the literature or GenBank are derived from introgression in M. galloprovincialis or whether introduction of M. edulis occurred from a distinct location. The latter possibility would require corroboration by finding a suite of M. edulis-characteristic alleles in an individual mussel. This does not appear likely given the absence of co-inheritance in the M- and F-type sequences of individuals and the finding of Northern Hemisphere mitochondrial haplotypes in individuals that are apparently homozygous for Southern Hemisphere alleles of the mac-1 intron 1 polymorphism.

The failure to discover M-type haplotypes in Mytilus galloprovincialis and the Southern Hemisphere endemic lineage(s) in Australia may be a sampling artefact. It is also possible that any endemic male mtDNA in the region has been replaced by competition from invasive haplotypes or that there has been a recent masculinisation event, in which a pre-existing Southern Hemisphere M-type has been replaced by the current F-type mtDNA (Hoeh et al. 1997).
The dispersion of southeastern Australian specimens throughout the Northern Hemisphere clade (Figs. 1 & 2) shows that much of the variation in the *Mytilus edulis* species complex is represented in the region. Since there are few geographically localised clades among the F-type haplotypes worldwide (Fig. 1), it cannot be proven whether there were many introductions of *Mytilus* spp. from separate populations or a low number of introductions from genetically highly variable sources, although both possibilities seem likely. It appears unlikely that there is an endemic *M. edulis*-type taxon in the Southern Hemisphere, although this possibility cannot be completely dismissed. Gardner (2004) found no historic evidence of shells of this species in New Zealand. One of the Australian *M. edulis* F-type sequences is quite basal (Fig. 2), but the others are dispersed throughout the topology, with no groups of haplotypes forming monophyletic lineages.

**Acknowledgements.** We thank the NSW Environmental Trust for supporting this project through Grant 2006/RD/0071. Collections for this project were made under Permits F86/2163(A) (New South Wales Fisheries Research permit), 7118 (Tasmania DPIW permit) and RP972 (Victoria General Research permit). We thank Spring Bay Seafoods for providing us with specimens (New South Wales Fisheries Research permit) and RP972 (Victoria General Research permit).

**LITERATURE CITED**


Swofford DL (2003) PAUP: phylogenetic analysis using parsimony; Version 4.0. Laboratory of Molecular Systematics, Smithsonian Institution, Washington, DC


Submitted: September 8, 2010; Accepted: January 27, 2011

Proofs received from author(s): March 14, 2011

Editorial responsibility: Francesco Patti, Ischia, Italy