Limited genetic variation and structure in softshell clams (*Mya arenaria*) across their native and introduced range

Running title: Genetic structure of softshell clams

C. A. Strasser*

Biology Department, Woods Hole Oceanographic Institution, Woods Hole MA 02543

P. H. Barber

Biology Department, Boston University, Boston MA 02215

Mailing Address for corresponding author:

MS 34, Woods Hole Oceanographic Institution

Woods Hole MA 02543

Phone: 508-289-2358. Fax: 508-457-2134. Email: cstrasser@whoi.edu.

Keywords: *Mya arenaria*, bivalve, COI, northwest Atlantic, genetic structure

ABSTRACT

To offset declines in commercial landings of the softshell clam, *Mya arenaria*, resource managers are engaged in extensive stocking of seed clams throughout its range in the northwest Atlantic. Because a mixture of native and introduced stocks can disrupt locally adapted genotypes, we investigated genetic structure in *M. arenaria* populations across its current distribution to test for patterns of regional differentiation. We sequenced mitochondrial cytochrome oxidase I (COI) for a total of 212 individuals from 12 sites in the northwest Atlantic (NW Atlantic), as well as two introduced sites, the northeast Pacific (NE Pacific) and the North Sea and Europe (NS Europe). Populations exhibited extremely low genetic variation, with one haplotype dominating (65-100%) at all sites sampled. Despite being introduced in the last 150-400 years, both NE Pacific and NS Europe populations had higher diversity measures than those in the NW Atlantic and both contained private haplotypes at frequencies of 10% to 27% consistent with their geographic isolation. While significant genetic structure ($F_{ST} = 0.159$, $p<0.001$) was observed between NW Atlantic and NS Europe, there was no evidence for genetic structure across the pronounced environmental clines of the NW Atlantic. Reduced genetic diversity in mtDNA combined with previous studies reporting reduced genetic diversity in nuclear markers strongly suggests a recent population expansion in the NW Atlantic, a pattern that may result from the retreat of ice sheets during Pleistocene glacial periods. Lack of genetic diversity and regional genetic differentiation suggests that present management strategies for the commercially important softshell clam are unlikely to have a significant impact on the regional distribution of genetic variation, although the possibility of disrupting locally adapted stocks cannot be excluded.

INTRODUCTION

Benthic marine habitats of the Northwest Atlantic Ocean (NW Atlantic) are structured into distinct biogeographic provinces (Engle and Summers 1999). These biogeographic divisions are a function of environmental gradients resulting from the synergy of the coastal geography of Eastern North America with the Gulf Stream and Labrador Currents, combined with latitudinal gradients in temperature and salinity (Hutchins 1947). The most commonly recognized biogeographic divisions are the Nova Scotian and Virginian Provinces, with Cape Cod serving as the boundary between the two (Hall 1964; Hutchins 1947). Superimposed on these divisions is a history of Pleistocene glaciations that extirpated many benthic marine species from northern latitudes and formed Cape Cod (Upham 1879a; Upham 1879b), reshaping regional patterns of biological and genetic diversity (Hewitt 1996; reviewed in Wares 2002; Wares and Cunningham 2001).

The presence of distinct biogeographic provinces in the NW Atlantic has significant implications for management of fish and invertebrates in this region because species spanning multiple provinces of the NW Atlantic can have populations adapted to local environmental conditions. For example, the Atlantic Silverside *Menidia menidia* exhibits heritable local variation in growth rate and vertebral number, resulting in a latitudinal phenotypic cline across the NW Atlantic (Billerbeck *et al.* 1997; Present and Conover 1992; Yamahira *et al.* 2006). On a smaller scale, the mussel *Mytilus edulis* exhibits a sharp cline in the leucine aminopeptidase (LAP) allele across salinity gradients in Long Island Sound (Gardner and Kathiravetpillai 1997; Gardner and Palmer 1998). The presence of regional genetic structure, particularly if it is locally adaptive, needs to

be accounted for in fisheries management so that genetic diversity is conserved and locally adaptive gene complexes are not disrupted through indiscriminate stocking (Hansen 2002).

Mya arenaria is a commercially important bivalve with a contemporary distribution that includes 1) the northwest Atlantic ranging from Nova Scotia to Virginia, 2) the North Sea and European waters, including the Black, Baltic, Wadden, White, and Mediterranean Seas, and 3) northeast Pacific from San Francisco to Alaska (Strasser 1999). *M. arenaria* has a complex history of extensive global distributions, with several extinctions and re-colonization events (reviewed in Strasser 1999). The species originated in the Pacific Ocean during the Miocene, and then extended its range to the Atlantic and European waters in the early Pliocene. Extinction of Pacific and European populations in the early Pleistocene left the only surviving populations in the NW Atlantic until recent history (MacNeal 1965). *M. arenaria* re-invaded European waters in the 17th century after being brought from the NW Atlantic by Vikings (Petersen et al. 1992). In the late 19th century *M. arenaria* was reintroduced into the Pacific, first accidentally and then as a potential commercial fishery (Carlton 1979; Powers et al. 2006). The natural and introduced distribution of *M. arenaria* results partly from the species' ability to withstand wide salinity and temperature ranges, and its capability of inhabiting different sediment types from fine mud to coarse sand (Abraham and Dillon 1986; Hidu and Newell 1989; Newell and Hidu 1982).

The last two decades have seen appreciable declines in softshell clam landings in New England (Anonymous 2007; Brousseau 2005). This decline has been attributed to habitat degradation or loss, overfishing, contamination, and predation by invasive species

(Brousseau 2005). Managers and state agencies have enacted various management strategies to combat these declines, including using protective nets to reduce predation on newly recruited clams, and seeding flats using hatchery-reared juveniles (H. Lind pers. comm. 2006, Marcotti and Leavitt 1997).

While stocking of fish and shellfish is a long-standing practice, research is increasingly showing that the genetic impacts of stocking cannot be ignored. Stocking should seek to maintain levels of genetic diversity (Waples and Do 1994); although multiple individuals are spawned to produce seed clams, it is unknown whether the genetic diversity represented among these individuals is reduced in comparison to naturally occurring cohorts, where entire adult populations spawn simultaneously (Brousseau 1978). In addition, given that brood stock is not always taken from the flat into which seed clams are stocked, locally appropriate genotypes could be introduced into inappropriate areas. For example, *Mya arenaria* exhibit local variation in resistance to paralytic shellfish toxins (Connell et al. 2007). Seeding flats using brood stock from other clam populations may result in either reduction of the locally dominant alleles due to success of the introduced seed clams, or significant loss of seed clams due to a lack of a genetic background appropriate to the local environment. Similar declines in local fitness have been documented in salmonids (Hansen 2002). He examined the relative genetic contributions of domesticated and wild trout in several populations and found that for at least one population, the observed contribution of domesticated fish (6%) was well below what was expected assuming equal survival and reproductive performance of wild and domesticated fish.

Previous genetic studies on softshell clams have found limited genetic diversity despite the wide geographic ranges represented among studies. Morgan *et al.* (1978) used allozymes to study *Mya arenaria* genetic variation in the NW Atlantic, and found low polymorphism and low heterozygosity per individual for both populations examined. Caporale *et al.* (1997) found similar low variability in three regions of the NW Atlantic (12 locations total) using the internal transcribed spacer ribosomal DNA region (nDNA), and concluded that although the three regions were not genetically distinct, the data from the study were insufficient to indicate a panmictic population. More recently, Lasota *et al.* (2004) used allozymes to study seven locations in the northeast Atlantic and two in the North Sea. They also found low genetic variability and a lack of genetic differentiation, and concluded that *M. arenaria* is a successful invader despite a high degree of genetic homogeneity. They suggested the patterns observed were evidence of rapid population expansion, allele neutrality, and high gene flow. Nuclear DNA (nDNA), however, is known to evolve slower than mitochondrial DNA (mtDNA), and allozyme studies may mask underlying sequence variation. The results seen in these studies therefore might be because of the markers chosen by the investigators.

In this study, we examine population genetic variability of *Mya arenaria* across its natural range in the NW Atlantic and portions of its introduced range in the northeast Pacific and European waters using the highly variable mitochondrial cytochrome oxidase I (COI) gene that commonly resolves phylogeographic structure in marine invertebrates (Barber *et al.* 2006; Wares 2002) including bivalves (King et al. 1999; May et al. 2006). First, we examine how populations may be geographically structured across the NW Atlantic to determine whether the distinct environments and biogeographic provinces

partition softshell clams into genetically distinct regional stocks. Second, we compare NW Atlantic to populations in the NE Pacific and NS Europe. *M. arenaria* was introduced recently to both of these locations; we can examine the geographic origins of these populations and the effects of recent introduction on genetic diversity. The results of this study have implications for management of softshell clams in New England, in addition to insights gained about historical extinction and colonization events of *M. arenaria* with reference to biogeographic boundaries and glaciation.

METHODS

Sampling and sequencing

Juvenile and adult *Mya arenaria* (N = 212) were collected between 2001 and 2006 from 12 locations: one northeast Pacific (NE Pacific) site (n = 20), ten NW Atlantic sites (n = 177), and one North Sea, Europe (NS Europe) site (n = 15) (Table 1). Most *M. arenaria* were frozen after collection to prevent DNA degradation, and then transferred to 70-95% ethanol for at least 24 hours prior to DNA extraction to improve the success of DNA extractions. Some individuals were preserved directly in ethanol without freezing. For clams <1cm total length, we used the entire clam for DNA extraction. For larger clams we extracted DNA from small fibers of adductor muscle tissue. All DNA extractions were performed with a 10% Chelex® (BioRad) solution following Walsh *et al.* (1991). A 661 bp fragment of the mitochondrial cytochrome oxidase subunit-I gene (COI) was amplified via polymerase chain reaction (PCR) using the primers HCO-2198 and LCO-1490 (Folmer *et al.* 1994). PCR occurred in 25 µl reactions with 2.5 µl of 10x buffer, 2 µl $MgCl_2$ (25 mM), 2.5 µl DNTPs (8 mM), 1.25 µl of each 10 mM primer, 1 µl

of template, and 0.625 units of Amplitaq™ (Applied Biosystems, Inc., Foster City, CA). Hot-start thermocycling parameters were as follows: initial denaturation 94°C (3 min); followed by 38 cycles of 94°C (30 s), 50°C (30s), 72°C (45s); then a final extension of 72°C (10 min).

PCR products were visualized on 1% agarose PAC 1% sodium hydroxide and boric acid gels. They were then prepared for sequencing by digestion in 0.5 units of Shrimp Alkaline Phosphotase and 5 units of exonuclease per 5 μl of PCR product, incubated at 37°C for 30 min followed by 80°C for 15 min. Sequencing reactions were performed for both forward and reverse strands using BigDye (Applied Biosystems Inc., Foster City, CA) terminator chemistry, and visualized on an ABI 377 (Applied Biosystems, Inc., Foster City, CA) following isopropanol precipitation according to manufacturer instructions. Complementary strands for each sample were proofread and aligned in SEQUENCHER v4.0 (GeneCode, Ann Arbor, MI), and translations confirmed using MACCLADE v4.05 (Maddison and Maddison 2002).

Genetic analyses for all sites

To explore regional distribution of genetic diversity in *Mya arenaria,* we calculated haplotype diversity (h), nucleotide diversity (π), and the population parameter theta (θ_s, where $\theta = 2N\mu$, estimated from the number of segregating sites, Watterson 1975) for all populations using ARLEQUIN v3.1 (Excoffier et al. 2005). To explore patterns of phylogeographic structure we constructed a minimum-spanning tree using the MINSPNET algorithm as employed in ARLEQUIN. Frequency of haplotypes was then plotted against geography for NW Atlantic populations.

To further explore geographic genetic structure, we investigated genetic partitions in AMOVA as implemented in ARLEQUIN. Values of Φ_{ST} were calculated with statistical significance determined by 20,000 random permutations. Analyses were run both unstructured (one region), and partitioned into two regions, (scenario A = NW Atlantic + NE Pacific and NS Europe; scenario B = NW Atlantic + NS Europe and NE Pacific) or three regions (NW Atlantic, NE Pacific, NS Europe). Patterns of genetic structure were similarly estimated within the NW Atlantic by excluding NE Pacific and NS Europe populations. NW Atlantic analyses were assuming one region or three regions (north of Cape Cod, Cape Cod, and South of Cape Cod). Pairwise F_{ST} values were also calculated among all populations with 20,000 permutations used to establish significance. To adjust significance levels for multiple comparisons, the Benjamini and Yekutieli (2001) modification of the Bonferroni correction was employed as standard Bonferroni corrections have been shown to be excessively conservative (Narum 2006).

Genetic analyses of NW Atlantic only

To determine whether our sampling in the NW Atlantic was sufficient to collect most of the haplotypes present, we constructed a rarefaction curve. We used data only from populations in the NW Atlantic and used equations appropriate for population sample sizes much smaller than the total sample size (Heck et al. 1975). In addition, we estimated the expected number of haplotypes in the NW Atlantic using two indices, the Chao1 index (Chao 1984) and the abundance-based coverage estimator (ACE) (Chao and Lee 1992).

We tested for neutrality by calculating Fu's F_S statistics (Fu 1997), which establishes whether non-neutrality might be due to population growth and range

expansion. To further explore the possibility of recent demographic or spatial population expansion, we used mismatch distributions, which compares the expected and observed number of differences between pairs of haplotypes under a null model of population expansion (Ray *et al.* 2003; Rogers and Harpending 1992). Finally, we used Bayesian Markov Chain Monte Carlo analysis of molecular sequences to produce a Bayesian skyline plot using BEAST v1.4 and TRACER v1.4 (Drummond and Rambaut 2006), which plots population size over time and estimates the approximate time since population expansion (Drummond et al. 2002; Drummond et al. 2005). We used MODELTEST (Posada and Crandall 1998) implemented in PAUP* ver.4.0b10 (Swofford 1998) to find the most appropriate model for BEAST (Hasegawa, Kishino, and Yano Model). We produced the skyline plot based on five groups using a strict molecular clock, which assumes a global clock rate with no variation among lineages within a tree. We ran the program using default priors for Bayesian skyline analysis for 50 million generations, and repeated the program run four times to increase effective sample size and assure that results were converging. Results reported in mutational units were converted to years for the skyline plot by assuming a molluscan-specific COI divergence rate of either 1 % per million years (% Myr^{-1}) for all COI sites or 5 % Myr^{-1} for third positions alone (Marko 2002). As all variation was located third position sites (see Results), using these values provides a five-fold range to account for error associated with molecular clock estimations.

RESULTS

Analyses for all sites

A total of 661 bp of COI was collected from 212 individuals, yielding only 27 unique haplotypes that differed by one or two nucleotide substitutions, all in the third codon position and silent with one exception, haplotype I (Table 2), where serine was substituted for proline. There was one dominant haplotype (A) found at all of the locations sampled, ranging in frequency from 0.53 to 1.00 for individual populations, with an overall frequency of 0.79. Of the remaining 26 haplotypes, only five were found more than once in a single population, ranging in frequency from 0.10 to 0.27 (haplotypes B-F). Two private haplotypes, defined as haplotypes that occur more than one time in only one site (Slatkin 1985), were found in the NS Europe (Haplotype E) and the NE Pacific (Haplotype G).

Haplotype diversity (h) in the NW Atlantic ranged from 0.178 to 0.583 (Table 1). Comparable levels of haplotype diversity occurred in NS Europe (h = 0.648) and NE Pacific (h = 0.574) populations. Nucleotide diversity (π) was low for all NW Atlantic populations, ranging from 0.0003 to 0.0014 (Table 1), while π = 0.0010 in NS Europe and π = 0.0012 in NE Pacific. Theta ranged from 1.85 to 0.549 in the NW Atlantic and was 0.615 and 1.13 in NS Europe and NE Pacific, respectively (Table 1). There were no clear geographic patterns in genetic diversity measures.

Consistent with the low nucleotide diversity, the minimum spanning tree of *Mya arenaria* COI haplotypes revealed a star-shaped phylogeny (Figure 1). The dominant haplotype (A) was located at the center of the star with 21 of 26 remaining haplotypes differing from haplotype A by a single nucleotide substitution. Five haplotypes differed by 2 mutational steps (haplotypes J, O, T, U, AA). No geographic structure is evident in the minimum spanning tree topology and NE Pacific and NS Europe haplotypes are

scattered throughout the tree. Plotting the frequency of the 6 non-singleton haplotypes revealed no clear phylogeographic patterns in the NW Atlantic except for the lack of genetic diversity in Nova Scotia (Figure 2).

Results from AMOVA found that the majority of variability was within populations, regardless of any regional partitions imposed on the locations sampled (Table 3). Examining all data, assuming no *a priori* regional structure, AMOVA analyses indicate the presence of subtle genetic structure ($\Phi_{ST} = 0.027$, $p < 0.005$) with 3% of the variation between populations and 97% of the variation within populations. Imposing regional partitions comparing North American (NW Atlantic + NE Pacific) and European (NS Europe) populations produced $\Phi_{ST} = 0.16$ ($p < 0.005$) with 16% of variation among regions, 0.010% among populations within regions, and 84.1% of the variation within populations. Comparing North Atlantic (NW Atlantic + NS Europe) and Pacific (NE Pacific) populations, $\Phi_{ST} = 0.34$ ($p < 0.005$) with 1.4% of variation among regions, 2.4% among populations within regions, and 96% of the variation within populations. Imposing three regional partitions (NW Atlantic, NE Pacific, and NS Europe), $\Phi_{ST} = 0.090$ ($p < 0.005$) with 91% of the variation within populations, 9.6% among regions, and no variation among populations within regions. Results are summarized in Table 3.

Within the NW Atlantic, AMOVA analyses revealed no significant genetic structure with 99.9% of all genetic variation contained within populations (Table 3). Similarly, when locations were grouped into regions north of Cape Cod, Cape Cod, and south of Cape Cod (Table 1), $\Phi_{ST} = 0.0010$ (n.s., $p = 0.43$) with 99.9% of all genetic variation contained within populations and no significant variation among regions or

among populations within regions. After correction for multiple comparisons, significant pairwise F_{ST} values were observed between NS Europe and 4 populations from the NW Atlantic (F_{ST} = 0.12 to 0.22; $p < 0.05$) (Table 4). Within the NW Atlantic there were no significant pairwise F_{ST} values.

Both NS Europe and NE Pacific had one haplotype each that was not shared with the NW Atlantic. A rarefaction curve of NW Atlantic haplotypes constructed using an equation appropriate for our system (Heck *et al.* 1975) did not asymptote, suggesting this may result from inadequate sampling. Chao1 and ACE indices both predicted that more than 100 haplotypes were present in the NW Atlantic (125 and 102, respectively); this further suggests that our sampling was not sufficient to characterize all of the diversity present.

Fu's F_S statistic was significantly large and negative for 9 of the 12 populations (Table 1) suggesting non-equilibrium dynamics. Mismatch analysis revealed no significant deviation from the null model of population expansion, and the raggedness index confirmed a left-shifted unimodal distribution characteristic of population expansion (Table 1, Figure 5). Further support for a range expansion comes from the Bayesian skyline plot, indicating that *Mya arenaria* populations in the NW Atlantic were much smaller in recent history (Figure 4). The plot indicates that a pronounced 100-fold demographic expansion event took place in NW Atlantic populations of *M. arenaria* approximately 75,000 or 15,000 years ago. These values correspond to 0.00035 mutational units and a mutation rate of 0.005 to 0.025 mutations Myr^{-1}, based on the clock calibrations of Marko (2002).

DISCUSSION

Patterns in the Northwest Atlantic

Genetic analysis of *Mya arenaria* populations across the Northwest Atlantic revealed an absence of genetic structure. This result stands in contrast to previous studies of other marine species that show pronounced phylogeographic structure in the NW Atlantic (see Wares 2002 for a review), particularly among populations along the northern and southern coastline of the NW Atlantic (e.g. Brown *et al.* 2001; Dahlgren *et al.* 2000; Smith *et al.* 1998; Waldman *et al.* 1996). The lack of genetic diversity and limited genetic structure reported here echoes previous genetic studies on softshell clams (Caporale *et al.* 1997; Lasota *et al.* 2004; Morgan *et al.* 1978). The concordant results among these multiple studies provide strong evidence for lack of genetic boundaries in *M. arenaria.*

One potential explanation for the observed pattern is that *Mya arenaria* is characterized by high levels dispersal and gene flow, as suggested by Lasota *et al.* (2004). This species has a planktonic larval phase that can last up to three weeks in the water column, during which time the larva feeds on algae and is transported by currents (Abraham and Dillon 1986). Transport via strong currents along the NW Atlantic could promote high dispersal and gene flow among NW Atlantic populations, with further mixing augmented by human-mediated transport (although this view must be tempered by the observation of low levels of genetic diversity). Given that

$$N_e = \frac{N}{1 \quad F_{ST}}$$

where N_e is effective population size and N is actual population size (Wright 1943), as gene flow increases and F_{ST} approaches zero, then N_e should approach N. Commercial

landings of this species totaled 7.9 million pounds in 1984 (Abraham and Dillon 1986), suggesting N exceeds 10 million individuals (assuming an average of 4 clams lb^{-1}). However, if gene flow is high, $N_e \approx N$, and N is extremely large, then genetic drift should be extremely low, preserving high levels of genetic diversity. This expectation is contradicted by the minimum spanning tree and the minimal genetic diversity measures. Thus, while high dispersal may contribute to genetic homogeneity in *M. arenaria*, other processes must also be acting to reduce diversity within this species.

A second process suggested by Lasota et al. (2004) that could contribute both to limited genetic structure and low genetic diversity across the NW Atlantic is a recent population expansion event. Evidence for a demographic expansion comes from the star-like phylogeny, low genetic diversity measures, the significantly large and negative values of Fu's F_S, and the mismatch distribution and raggedness index values that do not differ from a model of population expansion. Further evidence of demographic expansion comes from the Bayesian skyline plot produced using NW Atlantic data (Figure 4), suggesting a rapid 100 fold increase in population size. Although the lack of a species-specific clock and associated error requires cautious interpretation of age estimates, even increasing or decreasing the assumed rates by an order of magnitude puts this expansion squarely in the Pleistocene.

Based on the geology of the NW Atlantic, a recent range expansion is required to achieve contemporary distributions of *Mya arenaria*. During Plio-Pleistocene glacial cycles, glaciers covered most of the NW Atlantic coastline with the southern limit of glaciation near Cape Cod, Massachusetts (Cronin 1988; Shackleton *et al.* 1984). Glacial build-up would displace *M. arenaria* populations from much of its contemporary range,

and the lowering of sea levels by as much as 130 m (Porter 1989) would likely reduce available estuarine habitat as the shallow continental shelf of the NW Atlantic became exposed. After glaciers subsided, individuals from southern and potentially northern refugia would spread into previously unavailable glaciated habitats (Wares 2002; Wares and Cunningham 2001). Reduced genetic variation, such as that observed in *M. arenaria,* is common in reinvaded habitats following Pleistocene glacial periods (see Hewitt 2000 for review), as is reduced genetic variation in formerly glaciated regions of the NW Atlantic compared to glacier-free regions of Europe (Wares 2002).

An alternative explanation for the signal of demographic expansion is recovery following a selective sweep, where selectively advantageous haplotypes go to fixation (e.g. Berry et al. 1991). Recovery from a recent selective sweep could also yield a star-like phylogeny and lower genetic diversity that would inflate genetic similarity and gene flow estimates among populations, resulting in genetic patterns very similar to those observed during a population expansion or recovery from a bottleneck (Tajima 1989). However, lowered genetic diversity in NW Atlantic populations, characteristic of departures from neutral expectations in stable populations, are reported in this study as well as studies focusing on allozymes (Lasota *et al.* 2004; Morgan *et al.* 1978) and nuclear sequence data (Caporale *et al.* 1997). For selective sweeps to occur in multiple unlinked mtDNA and nuclear markers seems unlikely, particularly given that the signal of demographic expansion occurs during the Pleistocene. Furthermore, similarly reduced patterns of genetic diversity are also seen in co-distributed NW Atlantic populations of the ocean quahog, *Arctica islandica* (Dahlgren et al. 2000), *Mercenaria mercenaria* (Baker *et al.* 2008), suggesting that a common physical process is at work (Avise 2000).

Thus, while a selective sweep or purifying selection cannot be totally excluded, postglacial expansion is likely a more parsimonious explanation for the observed patterns, especially given the timing of demographic expansion suggested by the skyline plot (Figure 4).

Patterns across the Northwest Atlantic, European waters and the northeast Pacific

While populations in the NW Atlantic had minimal genetic structure, the strongest signal of regional genetic structure comes from comparing NW Atlantic populations to NS Europe. AMOVA results with NS Europe, NE Pacific, and NW Atlantic defined as separate regions resulted in a significant Φ_{ST} of 0.0903 (Table 3). Furthermore, of 11 pair-wise comparisons, a total of 4 pair-wise F_{ST} values among NS Europe and the NW Atlantic were significant (Table 2). This result indicates that despite being introduced from NW Atlantic populations, there are significant genetic differences among these regions.

Pacific and European populations also contained private haplotypes (Figure 1). The presence of unique haplotypes found multiple times in a single population suggests genetic isolation (Hartl and Clark 1997). Given the geographic separation of the NW Atlantic, Pacific and European waters, observation of genetic isolation should be expected. This result is surprising, however, given that both NE Pacific and NS Europe populations are thought to have been introduced from the NW Atlantic within the last 150 and 400 years, respectively. This seems a particularly short amount of time for local variation to evolve *in situ* and increase in frequency sufficiently to be detected by sampling 15-20 individuals. In contrast, no private haplotypes were detected in sampling of 177 individuals from the entire range of *Mya arenaria* in the NW Atlantic.

We constructed a rarefaction curve plotting number of haplotypes versus number of samples. Although the slope shallows, it did not asymptote over the range of number of individuals sampled (Figure 3) indicating that sampling 177 NW Atlantic individuals was insufficient to detect the rare haplotypes that founded NE Pacific and NS Europe populations. Paradoxically, if they are in very low frequencies in the NW Atlantic, it seems unlikely that they would be introduced to the Pacific and European waters. One interpretation of this result is that these private haplotypes may represent ancestral polymorphism from relic populations that survived the Pliocene extinction events in the Pacific and European waters. If, however, these were relic haplotypes, genetic divergence in excess of one mutational step would be expected, as a single 3rd position substitution in 661 bases over a minimum period of 2 million years would yield a substitution rate of 0.076% per million years, nearly two orders of magnitude lower than the 5% per million years reported by Marko (2002) for another bivalve. Thus, the minimal divergence of these haplotypes suggests that these are indeed introductions of rare NW Atlantic haplotypes.

Management implications

One of the current management strategies for NW Atlantic softshell clam populations is to increase local abundances by seeding flats with hatchery-reared juvenile clams. As has been demonstrated in fish, this approach has the potential to decrease or alter genetic variability by introducing non-native genotypes that may affect the fitness of both introduced and native stocks (Hansen 2002).

The low genetic diversity and minimal genetic structure observed in COI combined with previous results showing limited genetic diversity in *Mya arenaria* using

nuclear sequences (Caporale et al. 1997) and allozymes (Lasota *et al.* 2004; Morgan *et al.* 1978) suggests that brood stock origins may not be critical to maintaining current levels of genetic diversity and patterns of genetic structure across the NW Atlantic. Results of this study suggest that brood stocks should be quite similar regardless of their locality, and their resulting juvenile seed clams are likely interchangeable across geography.

Although we did not detect genetic structure using the mitochondrial COI gene, there may yet be other genes that might show variability within the NW Atlantic. Given that multiple previous genetic studies of *Mya arenaria* showed low genetic variability, however, the odds of distinguishing locally adapted stocks using neutral genetic variation is remote. Local adaptation has been noted in *M. arenaria* for toxin resistance (Connell *et al.* 2007), and there may very well be important regional genetic differences among clam stocks in non-neutral genes, even though such differences may be difficult or impossible to detect using genetic methods.

Although genetic methods can be extremely informative when significant geographic subdivisions are detected, issues surrounding ancestral polymorphism and non-neutral processes make it nearly impossible to make strong inferences when no genetic differentiation is observed (Hedgecock *et al.* 2007). As such, the results of this study cannot be interpreted as the genetic equivalence of all *Mya arenaria.* Instead, our results indicate that because of the demographic history of this species, neutral genetic markers are likely to be uninformative in distinguishing regional stocks. Given that ecological methods have succeeded in demonstrating local adaptation (Connell *et al.* 2007) where genetics has failed, managing using precautionary principles suggests that seeding from local stocks should be preferred, if possible, particularly when

physiological or immunological differences have been demonstrated with non-genetic methods.

ACKNOWLEDGEMENTS

This work was supported by NSF grants OCE-0326734 and OCE-0215905 to L. Mullineaux and OCE- 0349177 (Biological Oceanography) to PHB. Initial stages of this work were conducted as part of course BI536 (Molecular Ecology and Evolution) at Boston University and is contribution 003 from this course. We are grateful to the many people who collected samples for this study. We also thank L. Mullineaux, E. Crandall, E. Jones, J. Drew, D. Adams, and R. Jennings for helpful advice and discussion. E. Crandall, L. Mullineaux, S. Mills, and S. Beaulieu provided useful comments on early drafts. We also thank the two anonymous reviewers for their contributions to improving the manuscript. The experiments in this study comply with the current laws of the United States of America.

REFERENCES

Abraham BJ, Dillon PL (1986) Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (mid-Atlantic): Softshell clam. USFWS Biological Report TR EL-82-4**:** 18pp.

Anonymous (2007) Massachusetts Commercial Shellfish Harvest; 1996-2000. Available at: www.mass.gov/dfwele/dmf/commercialfishing/5yr_shellfish.pdf. Cited 8 Nov 2007

Avise JC (2000) Phylogeography. Harvard University Press, Boston, MA

Baker P, Austin JD, Bowen BW, Baker SM (2008) Range-wide population structure and history of the northern quahog (*Merceneria merceneria*) inferred from mitochondrial DNA sequence data. ICES J Mar Sci 65**:** 155-163

Barber PH, Erdmann MV, Palumbi SR (2006) Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the coral triangle. Evolution 60**:** 1825-1839

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29**:** 1165-1188

Berry AJ, Ajioka JW, Kreitman M (1991) Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. Genetics 129**:** 1111-1117

Billerbeck JM, Orti G, Conover DO (1997) Latitudinal variation in vertebral number has a genetic basis in the Atlantic silverside, *Menidia menidia*. Can J Fish Aquat Sci 54**:** 1796-1801

Brousseau DJ (1978) Spawning cycle, fecundity, and recruitment in a population of soft-shell clam, *Mya arenaria*, from Cape Ann, Massachusetts. Fish Bull 76**:** 155-166

Brousseau DJ (2005) Effects of natural mortality and harvesting on inshore bivalve population trends. In: Buchsbaum R, Pederson J and Robinson WE, (eds) The Decline of Fisheries Resources in New England: Evaluating the Impact of Overfishing, Contamination, and Habitat Degradation. MIT Sea Grant 05-5, Cambridge, MA

Brown AF, Kann LM, Rand DM (2001) Gene flow versus local adaptation in the northern acorn barnacle, *Semibalanus balanoides*: insights from mitochondrial DNA variation. Evolution 55**:** 1972-1979

Caporale DA, Beal BF, Roxby R, VanBeneden RJ (1997) Population structure of *Mya arenaria* along the New England coastline. Mol Mar Biol Biotechnol 6**:** 33-39

Carlton JT (1979) History, biogeography and ecology of the introduced marine and estuarine invertebrates of the Pacific coast of North America. Dissertation, University of California Davis

Chao A (1984) Non-parametric estimation of the number of classes in a population. Scand J Statist 11**:** 265-270

Chao A, Lee SM (1992) Estimating the number of classes via sample coverage. J Am Stat Assoc 87

Connell LB, MacQuarrie SP, Twarog BM, Iszard M, Bricelj VM (2007) Population differences in nerve resistance to paralytic shellfish toxins in softshell clam, *Mya arenaria*, associated with sodium channel mutations. Mar Biol 150**:** 1227-1236

Cronin TM (1988) Evolution of marine climates of the U.S. Atlantic coast during the past four million years. Phil Trans R Soc Lond, Ser B 318**:** 661-678

Dahlgren TG, Weinberg JR, Halanych KM (2000) Phylogeography of the ocean quahog (*Arctica islandica*): influences of paleoclimate on genetic diversity and species range. Mar Biol 137**:** 487-495

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161**:** 1307-1320

Drummond AJ, Rambaut A (2006) BEAST v.1.4. available from http://beast.bio.ed.ac.uk/

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22**:** 1185-1192

Engle VD, Summers JK (1999) Latitudinal gradients in benthic community composition in western Atlantic estuaries. J Biogeogr 26**:** 1007-1023

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform 1**:** 47-50

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit 1 from diverse metazoan invertebrates. Mol Mar Biol Biotech 3**:** 294-299

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147**:** 915-925

Gardner JPA, Kathiravetpillai G (1997) Biochemical genetic variation at a leucine aminopeptidase (LAP) locus in blue (*Mytilus galloprovincialis*) and greenshell

(*Perna canaliculus*) mussel populations along a salinity gradient. Mar Biol 128**:** 619-625

Gardner JPA, Palmer NL (1998) Size-dependent, spatial and temporal genetic variation at a leucine aminopeptidase (LAP) locus among blue mussel (*Mytilus galloprovincialis*) populations along a salinity gradient. Mar Biol 132**:** 275-281

Hall CA, Jr. (1964) Shallow-water marine climates and molluscan provinces. Ecology 45**:** 226-234

Hansen MM (2002) Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: an approach using microsatellite DNA analysis of historical and contemporary samples. Mol Ecol 11**:** 1003-1015

Hartl DL, Clark AG (1997) Population Substructure. In: Principles of Population Genetics, p.111-162. Sinauer, Sunderland MA

Heck KL, Jr., van Belle G, Simberloff D (1975) Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. Ecology 56**:** 1459-1461

Hedgecock D, Barber PH, Edmands S (2007) Genetic approaches to measuring connectivity. Oceanog 20**:** 70-79

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405**:** 907-913

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58**:** 247-276

Hidu H, Newell CR (1989) Culture and ecology of the soft-shelled clam *Mya arenaria*. In: Manzi JJ and Castagna M, (eds) Clam Mariculture in North America. Elsevier, Amsterdam

Hutchins LW (1947) The bases for temperature zonation in geographical distribution. Ecol Monogr 17**:** 325-335

King TL, Eackles MS, Gjetvaj B, Hoeh WR (1999) Intraspecific phylogeography of *Lasmigona subviridis* (Bivalvia: Unionidae): conservation implications of range discontinuity. Mol Ecol 8**:** S65-S78

Lasota R, Hummel H, Wolowicz M (2004) Genetic diversity of European populations of the invasive soft-shell clam *Mya arenaria* (Bivalvia). J Mar Biol Assoc UK 84**:** 1051-1056

MacNeal FS (1965) Evolution and distribution of the genus *Mya,* and tertiary migrations of mollusca. Prof Papers USGS 483-G**:** 1-51

Maddison DR, Maddison WP (2002) MacClade 4: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, MA

Marcotti T, Leavitt DF (1997) The Barnstable Harbor Shellfish Recruitment Enhancement Project (BHSREP): A Final Report. Woods Hole Oceanographic Institution: T-97-001**:** 47pp.

Marko PB (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. Mol Biol Evol 19**:** 2005-2021

May GE, Gelembiuk GW, Panov VE, Orlova MI, Lee CE (2006) Molecular ecology of zebra mussel invasions. Mol Ecol 15**:** 1021-1031

Morgan RP, Block SB, Ulanowicz NI, Buys C (1978) Genetic variation in the soft-shelled clam, *Mya arenaria*. Estuaries 1**:** 255-258

Narum SR (2006) Beyond Bonferroni: Less conservative analyses for conservation genetics. Conserv Genet 7**:** 783-787

Newell CR, Hidu H (1982) The effects of sediment type on growth rate and shell allometry in the soft shelled clam, *Mya arenaria* L. J Exp Mar Biol Ecol 65**:** 285-295

Petersen KS, Rasmussen KL, Heinemeler J, Rud N (1992) Clams before Columbus? Nature 359**:** 679

Porter SC (1989) Some geological implications of average Quaternary glacial conditions. Quat Res 32**:** 245-261

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14**:** 817-818

Powers SP, Bishop MA, Grabowski JH, Peterson CH (2006) Distribution of the invasive bivalve *Mya arenaria* L. on intertidal flats of south central Alaska. J Sea Res 55**:** 207-216

Present TMC, Conover DO (1992) Physiological basis of latitudinal frowth differences in *Menidia menidia*: variation in consumption or efficiency? Funct Ecol 6**:** 23-31

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol Biol Evol 20**:** 76-86

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9**:** 552-569

Shackleton NJ, Backman J, Zimmerman H, Kent DV, Hall MA *et al.* (1984) Oxygen isotope calibration of the onset of ice-rafting and history of glaciation in the North Atlantic region. Nature 307**:** 620-623

1 Slatkin M (1985) Rare alleles as indicators of gene flow. Evolution 39**:** 53-65
2 Smith MW, Chapman RW, Powers DA (1998) Mitochondrial DNA analysis of Atlantic
3 coast, Chesapeake Bay, and Delaware Bay populations of the teleost *Fundulus*
4 *heteroclitus* indicates temporally unstable distributions over geologic time. Mol
5 Mar Biol Biotechnol 7**:** 79-87
6 Strasser M (1999) *Mya arenaria*- an ancient invader of the North Sea coast. Helgol
7 Meeresunters 52**:** 309-324
8 Swofford DL (1998) Phylogenetic analysis using parsimony and other methods (PAUP*).
9 Available from: http://paup.csit.fsu.edu, Sinauer, Sunderland MA
10 Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA
11 polymorphism. Genetics 123**:** 585-595
12 Upham W (1879a) The Formation of Cape Cod. Am Nat 13**:** 489-502
13 Upham W (1879b) The Formation of Cape Cod (Continued). Am Nat 13**:** 552-565
14 Waldman JR, Nolan K, Hart J, Wirgin II (1996) Genetic differentiation of three key
15 anadromous fish populations of the Hudson River. Estuaries 19**:** 759-768
16 Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction
17 of DNA for PCR based typing from forensic material. Biotechniques 10**:** 506-513
18 Waples RS, Do C (1994) Genetic risk associated with supplementation of Pacific
19 salmonids: captive broodstock programs. Can J Fish Aquat Sci 51**:** 310-329
20 Wares JP (2002) Community genetics in the Northwestern Atlantic intertidal. Mol Ecol
21 11**:** 1131-1144
22 Wares JP, Cunningham CW (2001) Phylogeography and historical ecology of the North
23 Atlantic intertidal. Evolution 55**:** 2455-2469

Watterson G (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7**:** 256-276

Wright S (1943) Isolation by distance. Genetics 28**:** 114-138

Yamahira K, Lankford Jr. TE, Conover DO (2006) Intra- and interspecific latitudinal variation in vertebral number of *Menidia* spp. (Teleostei: Atherinopsidae). Copeia 2006**:** 431-436

FIGURE CAPTIONS

Figure 1. Unrooted minimum-spanning tree depicting the relationship of the 27 mitochondrial COI haplotypes from 212 individuals, collected from 12 sites in the NW Atlantic (white; n = 177), one site in the North Sea, Europe (black; n = 15), and one site in the northeast Pacific (gray; n = 20). Line distance between circles corresponds to the number of nucleotide differences (1 or 2). Each circle represents a unique haplotype, and the area of each circle corresponds to the number of individuals with that haplotype (the smallest circles are singletons). Letters correspond to Table 2.

Figure 2. Distribution of mitochondrial COI haplotypes for *Mya arenaria* in the NW Atlantic. Gray shades are unique haplotypes found in only one location; patterns are haplotypes shared among two or more locations. See Table 1 for site abbreviations and sample sizes.

Figure 3. Rarefaction curve constructed using data from *Mya arenaria* populations in the NW Atlantic. Number of haplotypes = 25; number of individuals sampled = 177.

Figure 4. Bayesian skyline plot derived from *Mya arenaria* NW Atlantic sequences. The solid line is the median estimate of population size, and the shaded region shows 95% highest posterior density limits (see Drummond *et al.* 2005). The dashed line indicates where in time the population expanded. Axis A is the years before present when a 1% per million years divergence rate is used; Axis B is the years before present when a 5% per million years divergence rate is used (see text for details). The gray arrows on each axis represent the approximate timing of the last glacial maximum. The y-axis is population size, with N = historical population size. ybp = years before present.

Figure 5. Frequency distributions of pairwise number of mutational differences between individuals for all NW Atlantic samples combined.

Table 1. Sampling and genetic information

Group	Population	Abrv	Latitude/Longitude	n	*h*	π	θ_S	F_S	P(*SSD*)	P(*Rag*)	Collected by
North of Cape	Antigonish, NS	NS	45° 37' N / 61° 59' W	20	0.000	0.000	-	-			D. Garbary
	St. John, NB	NB	45° 17' N / 66° 04' W	11	0.491 ± 0.175	0.0008 ± 0.0008	1.024 ± 0.676	-2.04 (0.008)	0.10	0.05	M.J. Maltais & H. Hunt
	Pembroke, ME	ME	44° 57' N / 67° 10' W	21	0.271 ± 0.124	0.0006 ± 0.0006	1.112 ± 0.641	-2.14 (0.006)	0.35	0.60	Gulf of Maine, Inc.
	Quincy , MA	QMA	42° 17' N / 71° 02' W	22	0.178 ± 0.106	0.0003 ± 0.0004	0.549 ± 0.408	-1.97 (0.008)	0.40	0.55	C. Strasser
Cape	Barnstable, MA	BMA	41° 42' N / 70° 20' W	19	0.205 ± 0.119	0.0003 ± 0.0005	0.572 ± 0.427	-1.80 (0.008)	0.35	0.80	C. Strasser
	Mashpee, MA	MMA	41° 36' N / 70° 27' W	20	0.447 ± 0.137	0.0009 ± 0.0008	1.691 ± 0.863	-3.95 (< 0.001)	0.50	0.60	C. Strasser
	Wareham, MA	WMA	41° 47' N / 70° 00' W	25	0.430 ± 0.124	0.0009 ± 0.0009	1.854 ± 0.888	-4.90 (< 0.001)	0.35	0.60	C. Strasser
South of Cape	Stony Brook, NY	NY	40° 54' N / 73° 07' W	15	0.476 ± 0.155	0.0008 ± 0.0008	1.230 ± 0.725	-3.23 (< 0.001)	0.50	0.80	L. Davies
	Miles River, MD	MMD	38° 47' N / 76° 08' W	15	0.371 ± 0.153	0.0006 ± 0.0007	0.923 ± 0.598	-2.37 (0.003)	0.70	0.65	M. Kramer & T. Hines
	Eastern Bay, MD	EMD	38° 51' N / 76° 15' W	9	0.583 ± 0.183	0.0014 ± 0.0012	1.470 ± 0.905	-1.28 (0.046)	0.90	0.99	M. Homer & C. Dungan
Outside NWA	Newport, OR	NEP	44° 36' N / 124° 03' W	20	0.574 ± 0.121	0.0010 ± 0.0009	1.127 ± 0.652	-4.28 (0.016)	0.95	0.90	J. Chapman & J. Chapman
	Sylt, Germany	NSE	54° 55' N / 8° 21' E	15	0.648 ± 0.088	0.0012 ± 0.0010	0.615 ± 0.462	0.365 (0.549)	0.15	0.20	S. Jacobsen

Abrv = abbreviation for site used in text; n = sample size, h = haplotype diversity (± SD); π = nucleotide diversity (± SD); θ_s, where $\theta = 2N\mu$ estimated from the number of segregating sites, F_S = Fu's F statistic and its associated p value. P(SSD) is probability of observing by chance a worse fit between the observed data and the mismatch distribution. P(Rag) is the probability of observing by chance a higher raggedness index than the observed index; non-significant raggedness indices indicate that the data fit a population expansion model.

Table 2. Haplotype distributions

	Location												
Haplotype	NS	NB	ME	QMA	BMA	MMA	WMA	NY	MMD	EMD	NEP	NSE	ALL
A	1.00 (20)	0.73 (8)	0.86 (18)	0.91 (20)	0.91 (17)	0.75 (15)	0.76 (19)	0.73 (11)	0.80 (12)	0.67 (6)	0.65 (13)	0.53 (8)	0.79 (167)
B		0.091 (1)					0.040 (1)	0.067 (1)	0.067 (1)		0.050 (1)	0.20 (3)	0.038 (8)
C			0.048 (1)							0.11 (1)	0.10 (2)		0.019 (4)
D		0.091 (1)				0.050 (1)					0.10 (2)		0.019 (4)
E												0.27 (4)	0.019 (4)
F						0.050 (1)	0.040 (1)	0.067 (1)					0.014 (3)
G											0.10 (2)		0.0094 (2)
H		0.091 (1)											0.0047 (1)
I			0.048 (1)										0.0047 (1)
J			0.048 (1)										0.0047 (1)
K				0.045 (1)									0.0047 (1)
L				0.045 (1)									0.0047 (1)
M					0.053 (1)								0.0047 (1)
N					0.053 (1)								0.0047 (1)
O						0.050 (1)							0.0047 (1)
P						0.050 (1)							0.0047 (1)
Q						0.050 (1)							0.0047 (1)
R							0.040 (1)						0.0047 (1)
S							0.040 (1)						0.0047 (1)
T							0.040 (1)						0.0047 (1)
U							0.040 (1)						0.0047 (1)
V								0.067 (1)					0.0047 (1)
W								0.067 (1)					0.0047 (1)
X									0.067 (1)				0.0047 (1)
Y									0.067 (1)				0.0047 (1)
Z										0.11 (1)			0.0047 (1)
AA										0.11 (1)			0.0047 (1)
Total individuals	20	11	21	22	19	20	25	15	15	9	20	15	212

Haplotype frequencies are given for each locality sampled, with the number of individuals per haplotype in parentheses. Site abbreviations are given in Table 1. ALL column is sum of all sites sampled.

Table 3. Results of AMOVAs

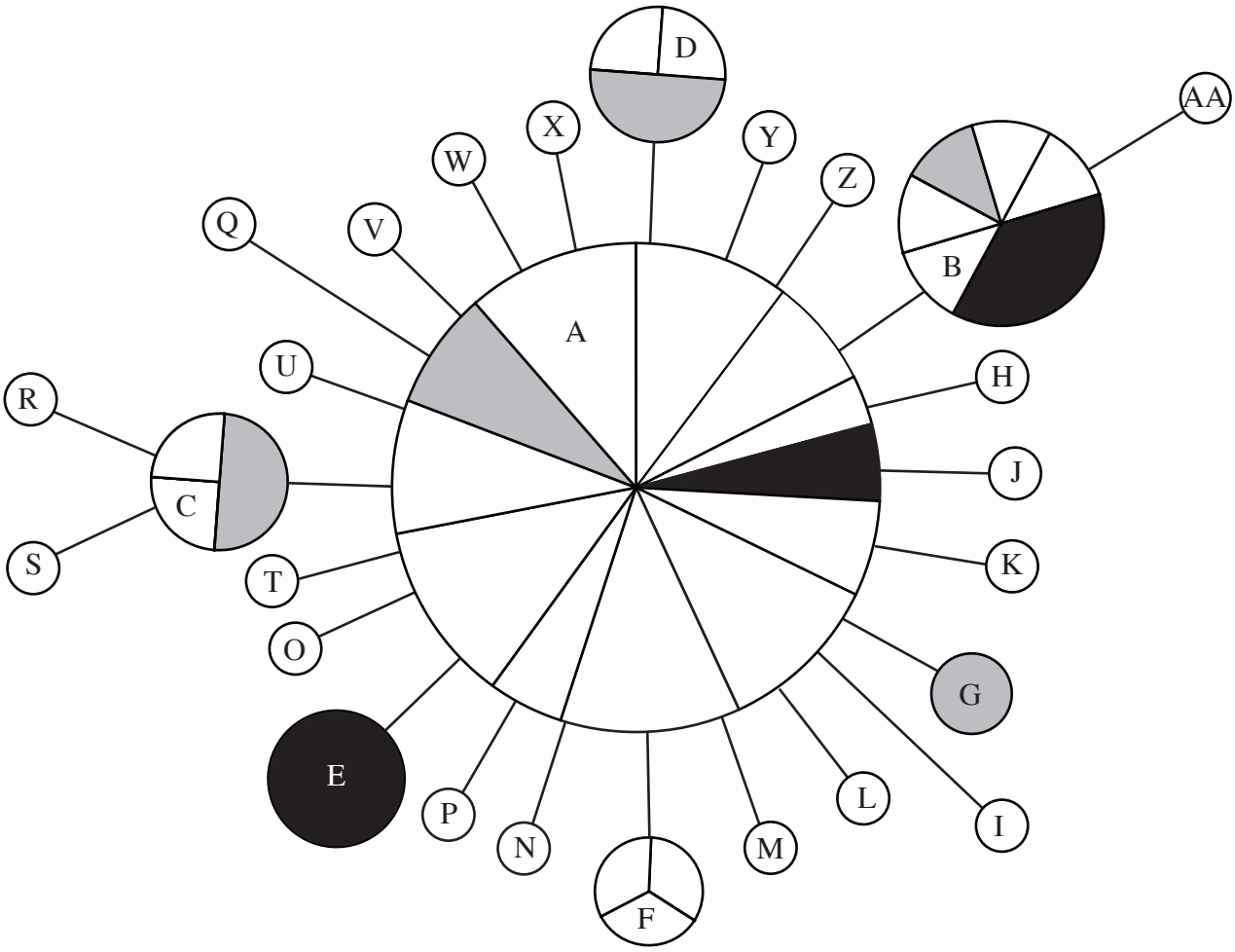
	Among regions					Among populations within regions					Within populations				
	df	Var.	% Var.	Φ_{CT}	p value	df	Var.	% Var.	Φ_{SC}	p value	df	Var.	% Var.	Φ_{ST}	p value
All Locations															
1 region						11	0.0062	2.7			200	0.23	97	0.027	0.0013
2 regions A	1	0.043	16	0.16	0.084	10	0.00003	0.01	0.0001	0.27	200	0.23	84	0.16	0.0011
2 regions B	1	0.0032	1.4	0.14	0.17	10	0.0056	2.4	0.024	0.23	200	0.23	96	0.34	0.0014
3 regions	2	0.024	9.6	0.960	0.045	9	-0.0014	-0.54	-0.0060	0.43	200	0.23	91	0.090	0.0011
NWA Only															
1 region						9	0.00008	0.04			167	0.20	100	0.00040	0.43
3 regions	2	0.00043	0.21	0.0021	0.19	7	-0.0002	-0.12	-0.0012	0.43	167	0.20	100	0.0010	0.43

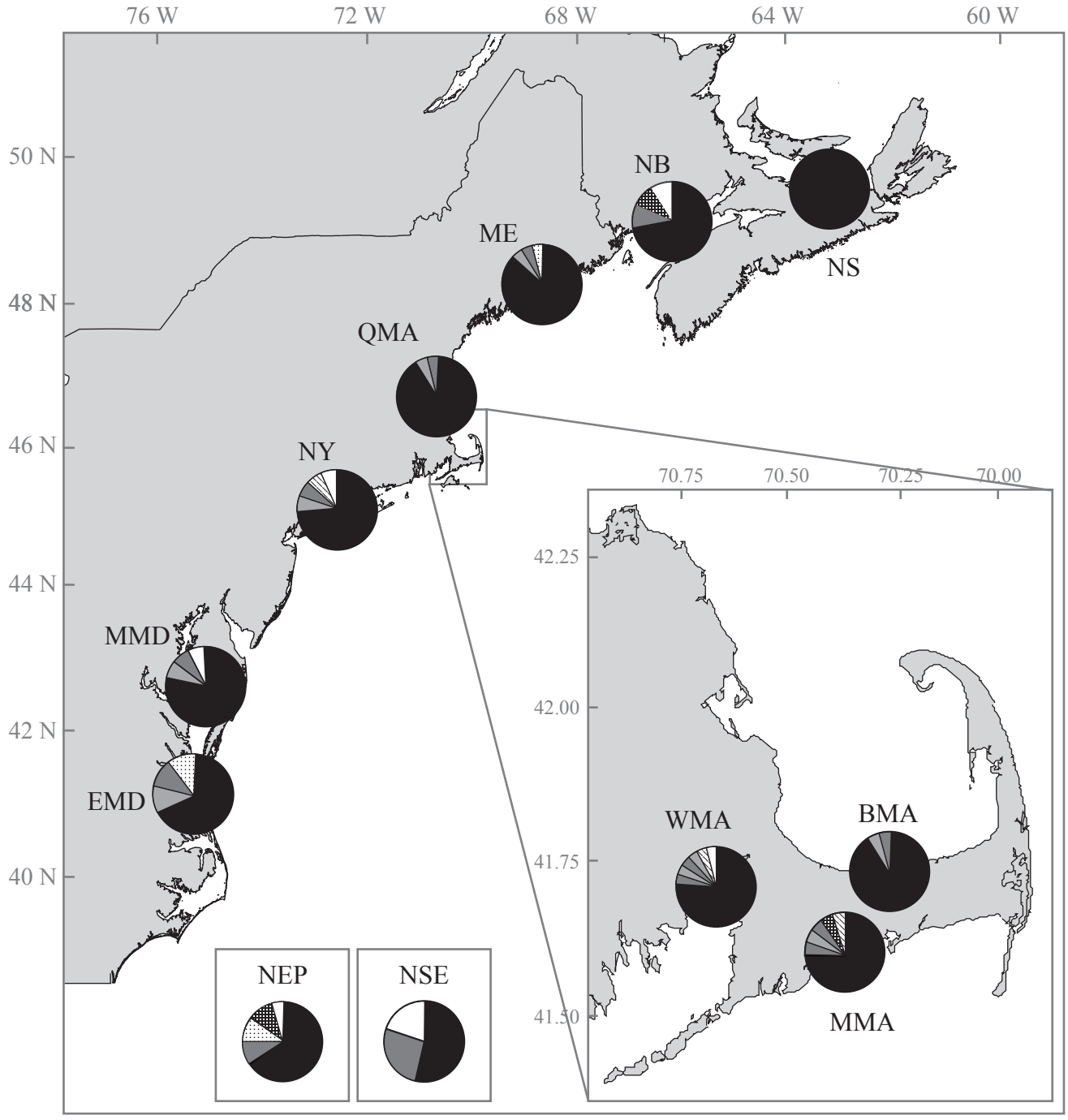
Results of AMOVAs testing for geographic structure across all locations and in the NWA only. Regions were as follows for *All Locations* analyses. 2 regions A = North American (NWA+NEP) and European (NSE) populations; 2 regions B = Atlantic (NWA+NSE) and Pacific (NEP); 3 regions = NWA, NEP, and NSE. For NWA only, 3 regions = North of Cape, Cape Cod, and South of Cape. p-values are associated with Φ values.

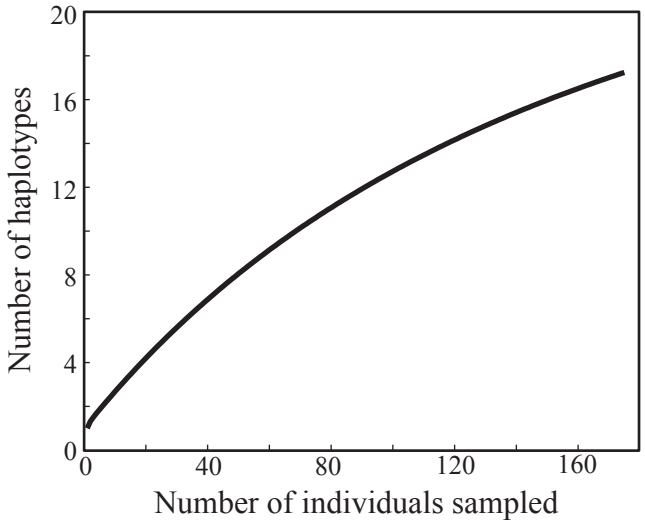
Table 4. Pairwise population comparisons, F_{ST}

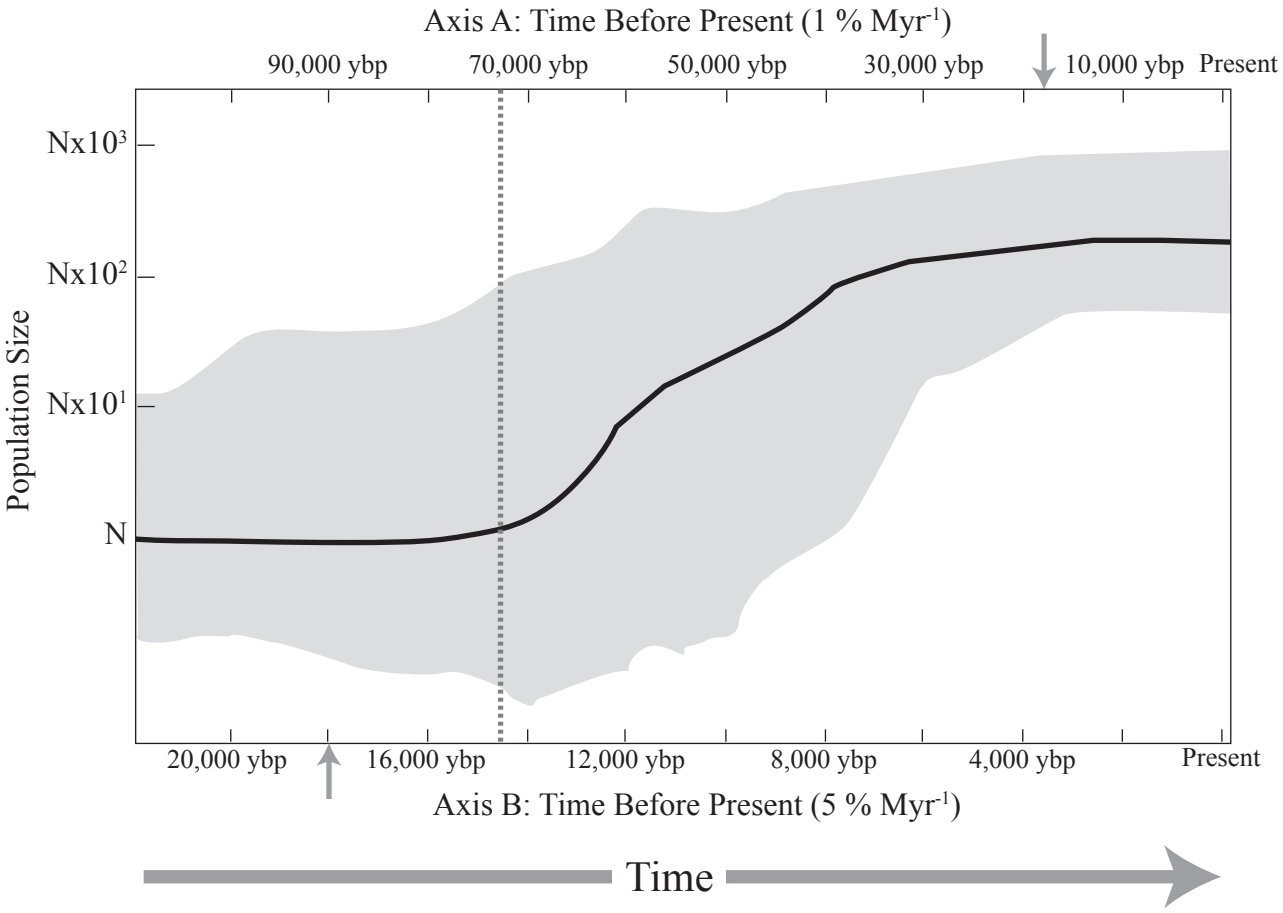
	NSE	NS	NB	ME	QMA	BMA	MMA	WMA	NY	MMD	EMD	NEP
NSE	-	**0.00066**	0.096	**0.0**	**0.00030**	*0.0013*	**0.00030**	*0.0093*	0.11	0.11	0.10	*0.0079*
NS	0.22*	-	0.036	1.0	1.0	0.23	1.0	0.50	0.026	0.62	0.023	0.11
NB	0.065	0.058*	-	0.45	0.22	0.41	0.86	0.67	0.85	0.77	0.69	0.74
ME	0.14*	-0.0024	0.0084	-	0.34	0.83	0.37	0.93	0.45	0.60	0.37	0.20
QMA	0.18*	-0.0045	0.028	0.00079	-	0.65	0.13	0.47	0.16	0.32	0.083	*0.0063*
BMA	0.170	0.003	0.0056	-0.0015	0.00054	-	0.99	0.47	0.23	0.40	0.12	0.22
MMA	0.12*	0.000	-0.018	0.00055	0.0026	-0.0078	-	0.56	0.88	0.81	0.39	0.42
WMA	0.099	0.000	-0.010	-0.010	0.0049	0.0013	-0.0014	-	0.83	0.75	0.80	0.48
NY	0.084	0.020*	-0.023	0.0034	0.012	0.0066	-0.013	-0.015	-	1.0	0.69	0.31
MMD	0.092	0.020	-0.023	0.000	0.0088	0.0047	-0.0033	-0.010	-0.019	-	0.64	0.30
EMD	0.052	0.097*	-0.024	0.012	0.060	0.047	0.013	-0.020	-0.010	-0.0060	-	0.60
NEP	0.100	0.045	-0.023	0.013	0.039	0.022	0.0091	-0.0011	0.013	0.012	-0.015	-

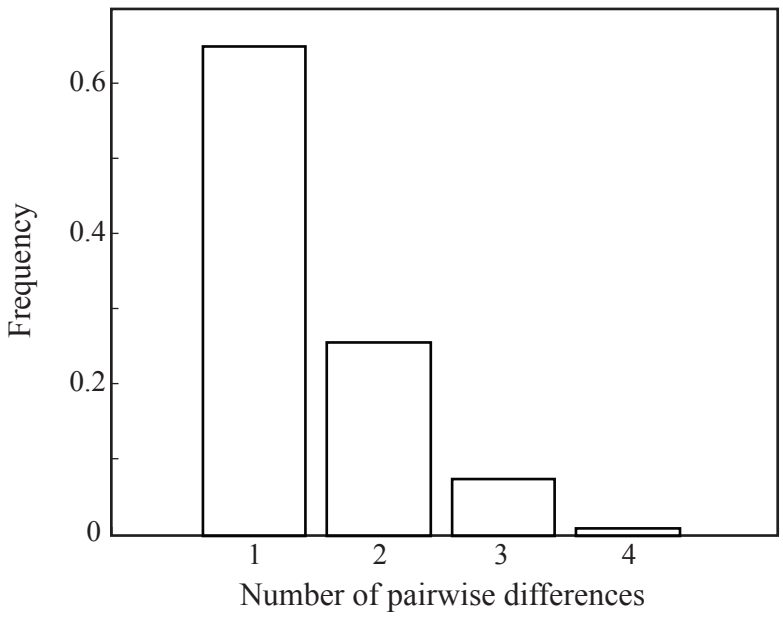
Below diagonal: population pairwise F_{ST} values. * indicates that the value is significant at the $p = 0.05$ level. Above diagonal: p value for F_{ST} values. Bold values remained significant after Bonferroni correction; italicized values became insignificant after correction.


D
AA
X
Y
W
Z
Q
V
B
A
U
H
R
J
C
K
S
T
O
G
E
P
N
L
M
I
F


76 W
72 W
68 W
64 W
60 W
50 N
48 N
46 N
44 N
42 N
40 N
NB
ME
NS
QMA
NY
MMD
EMD
NEP
NSE
70.75
70.50
70.25
70.00
42.25
42.00
41.75
41.50
WMA
BMA
MMA


20
16
12
8
4
0
Number of haplotypes
0
40
80
120
160
Number of individuals sampled


Axis A: Time Before Present (1 % Myr-1)
90,000 ybp
70,000 ybp
50,000 ybp
30,000 ybp
10,000 ybp
Present
Nx10³
Nx10²
Nx10¹
N
Population Size
20,000 ybp
16,000 ybp
12,000 ybp
8,000 ybp
4,000 ybp
Present
Axis B: Time Before Present (5 % Myr-1)
Time


0.6
0.4
0.2
0
Frequency
1
2
3
4
Number of pairwise differences