Aquat. Living Resour. 24, 35–43 (2011)

DOI: 10.1051/alr/2011107
www.alr-journal.org

Aquatic
Living
Resources

Changes in molecular genetic variation at AFLP loci associated with naturalization and domestication of the Pacific oyster (*Crassostrea gigas*)⋆

Mark D. Camara[a]

USDA, Agricultural Research Service, Shellfish Genetics, 2030 SE Marine Science Dr., Newport, Oregon 97365, USA

Received 9 September 2010; Accepted 14 March 2011

Abstract – The Pacific oyster (*Crassostrea gigas*) is an important commercial species in the US Pacific Northwest with a history of production initially based on wild-caught seed imported directly from the Miyagi region of northern Japan (1920s–1970s) followed by an extended period of seed collection from a small number of naturalized, self-recruiting populations in US estuaries (early 1970s-present) and more recently through large-scale hatchery production of seed oysters (mid 1970s-present). I studied the genetic level consequences of each of these major transformations of the oyster industry by examining the patterns of private alleles (bands unique to one sample), the number of polymorphic loci, expected heterozygosity, genetic distance between populations and genetic divergence between individuals within and among three native Pacific oyster populations in Japan (Hiroshima, Miyagi, Midori River), five naturalized populations in North America (Pipestem Inlet BC, Nootka Inlet BC, Dabob Bay WA, Willapa Bay, WA, and Tillamook Bay, OR), two in New Zealand (Chance Bay and Kaipara Harbor), and seven domesticated and selectively bred cohorts from an ongoing genetic improvement program on the West Coast of the United States using amplified fragment length polymorphism (AFLP) markers. All but one of the naturalized populations in both the USA and New Zealand are genetically more similar to native populations from the Ariake Sea (Midori River) than to the Miyagi region of their origin, but all domesticated stocks more closely resemble the wild Miyagi population. According to local oyster producers, the one exceptional naturalized population (Tillamook) is a very recent colonization derived from farmed oysters. Such consistency is unexpected under random genetic drift, and I speculate that both natural and artificial selection may have altered AFLP allele frequencies in this species in the course of naturalization and domestication.

Key words: Amplified fragment length polymorphism / Naturalization / Domestication selective breeding / Population genetics / Genetic diversity / Oyster

1 Introduction

Pacific oyster (*Crassostrea gigas*) aquaculture has become a world-wide industry and is particularly important for coastal communities in the Pacific Northwest region of the USA.The species is not native to this area; it was imported from Japan after populations of the only native oyster species (*Ostrea lurida*, the Olympia oyster) were decimated by a combination of over harvesting and environmental degradation caused by deforestation, mining, and industrial pollution in the early 20th century (Breese and Wick 1974). The first attempts in the early 1900s to grow imported *C. gigas* seed in the US Pacific Northwest used seed from the Hiroshima region of Japan and were unsuccessful, but seed from the Miyagi/Sendai region fared much better and eventually became the basis for the Pacific coast oyster industry (Breese and Wick 1974). From the late 1920s to the early 1970s, commercial production of *C. gigas* depended on annual large-scale importation and transplantation of wild-collected Miyagi oyster seed (Breese and Wick 1974; Clark and Langmo 1979). Even though this transplanted seed grew well in many West Coast locations, in sharp contrast to the situation in Europe (Troost 2010), *C. gigas* established self-recruiting populations in only a few estuaries where sufficiently high summer water temperatures support spawning and larval development and sufficiently retentive hydrodynamics prevent the flushing of externally fertilized larvae with a longer free-swimming period than the native species into the open ocean. At present, there are large, self-recruiting populations of *C. gigas* in Dabob Bay in Puget Sound, Willapa Bay on the coast of Washington state, and several sites in Pendrell Sound in British Columbia (Breese and Wick 1974). In 2000 or 2001, another colonization event occurred in Tillamook Bay, Oregon, but this population seems to have had little or no recruitment since then (Weigardt, pers. comm.). Beginning

⋆ Supporting information is only available in electronic form at www.alr-journal.org

[a] Corresponding author: Mark.Camara@ars.usda.gov

in the early 1970s, West Coast growers of Pacific oysters relied heavily upon wild-caught seed from these few naturalized populations and transplanted juveniles to a wide range of grow-out sites in Washington, Oregon, and California (Clark and Langmo 1979).

In the late mid-to-late 1970s, however, a combination of recruitment failures in naturalized populations and a tripling of the cost of Japanese oyster seed prompted an intensification of efforts to develop commercial-scale oyster hatchery technology (Clark and Langmo 1979) and remote setting technology (Jones and Jones 1983) to stabilize the Pacific oyster seed supply. Broodstock animals for early hatchery production were collected directly from naturalized populations, but as hatchery-produced seed became more common, broodstock were also collected from hatchery-produced plantings and were undoubtedly subjected to some level of informal selective breeding through both conscious decisions to spawn the best animals and through unconscious domestication selection on gamete production and larval development in novel hatchery environments where mortality is high. In addition, the extremely high fecundity of oysters allowed for the production of huge numbers of seed from small numbers of parents, facilitating genetic drift and inbreeding (Newkirk 1978). Large numbers of hatchery-produced animals have been reared alongside their progenitor populations and doubtlessly spawned, potentially influencing the genetic composition of the naturalized populations. Formal selective breeding in the 1980s to reduce stress-related mortality (Hershberger et al. 1984, and references therein) and the initiation in 1996 of several closed populations as part of the Molluscan Broodstock Program (MBP) – a family-based selective breeding program that provides improved broodstock to commercial hatcheries (Langdon et al. 2003) – may also have had genetic impacts. MBP's semi-domesticated populations have been propagated exclusively in a research-scale hatchery from a limited number of founders and subjected to artificial selection for enhanced yield for several generations.

This history of repeated, large-scale introduction from a single region in Japan, limited naturalization and potential adaptation to novel environmental conditions, informal domestication selection in commercial hatcheries, formal artificial selection in breeding programs, and commercial-scale grow out of reproductively competent hatchery-produced animals side-by-side with their wild cousins has potentially influenced the molecular genetic diversity of both naturalized populations and selectively bred Pacific oysters. Previous studies have, for example, shown that despite a long history of importation of presumably very genetically diverse groups of seed oysters from Japan, at least one naturalized population of Pacific oysters in the US (Dabob Bay, WA) has a very low effective population size (Hedgecock et al. 1992).

Because Pacific oysters have become such an economically important species in the US Pacific Northwest, the levels of genetic variation in both naturalized and domesticated populations are of considerable interest insofar as the availability of genetic variation is an important determinant of existing populations' capacity to adapt to changing environmental conditions such as ocean acidification through natural selection (Feely et al. 2004) and the potential for artificial selection to improve economically important traits. The objective of this study was to investigate the impacts of historical and ongoing naturalization, domestication, and artificial selection on the levels of genetic diversity in Pacific oyster germplasm by quantifying the genetic variation within and the extent of genetic differentiation between naturalized and domesticated populations and their wild progenitors using molecular genetic markers. Specifically, using amplified fragment length polymorphism (AFLP) markers I examine how naturalization and domestication have altered the molecular genetic variability within and promoted genetic differentiation between naturalized and domesticated populations relative to their wild progenitors.

2 Materials and methods

2.1 Biological materials

I obtained 50 or more tissue or DNA samples from each of three wild populations in Japan, five naturalized populations in the Pacific Northwest and two naturalized populations in New Zealand (Fig. 1) in addition to samples from seven selectively-bred cohorts from MBP (Fig. 2) for a total of 967 individual oysters. Miyagi and Hiroshima oysters were collected in 2003, Midori River oysters in 2004, naturalized oysters in the USA, Canada and New Zealand in 2006. MBP samples consist of all the parents used to create each of the progeny cohorts studied collected on the day of spawning (Fig. 2). For all samples, small pieces of fresh mantle tissue or adductor muscle from live animals were stored in 95% ethanol until DNA could be extracted.

2.2 DNA extraction and AFLP genotyping

DNA was extracted using Quiagen DN-Easy kits (Qiagen, Valencia CA, USA) for all samples except for the Hiroshima and Miyagi samples, which were exracted using Nucleo Spin Tissue Kits (Takara, Otsu, Shiga Japan), both according to the manufacturer's instructions. Prior to AFLP analysis, we confirmed the species identifications of all samples using the multiplex polymerase chain reaction (PCR) assay of Wang and Guo (2008) following their published protocol.

AFLP markers were typed following Li and Guo's (2004) modifications of Perkin Elmer's AFLP protocol, which is based in turn on the protocol of Vos et al. (1995). The digestion-ligation reaction consisted of about 200 ng of extracted DNA, 1.1 μl 10X T_4 DNA ligase buffer with ATP, 1.1 μl 0.5 M NaCl, 0.55 μl of 1 mg ml^{-1} bovine serum albumin (BSA), 1 U *Mse*1, 5 U *Eco*R1 (New England Biolabs, Beverly MA, USA), and 1.0 U T_4 DNA ligase (New England Biolabs, Beverly, MA, USA), 50 pmol Mse1 adaptor, 5 pmol *Eco*R1 adaptor, and water to bring the final volume to 11 μl. The digestion-ligation reaction mixtures were incubated at 16 °C for 16 h and then diluted with 150 μl $TE_{0.1}$ (20 mM Tris-HCl, 0.1 mM EDTA, pH 8.0).

Pre-selective PCR amplifications were performed using primer pairs complementary to the ligated adaptor sequences

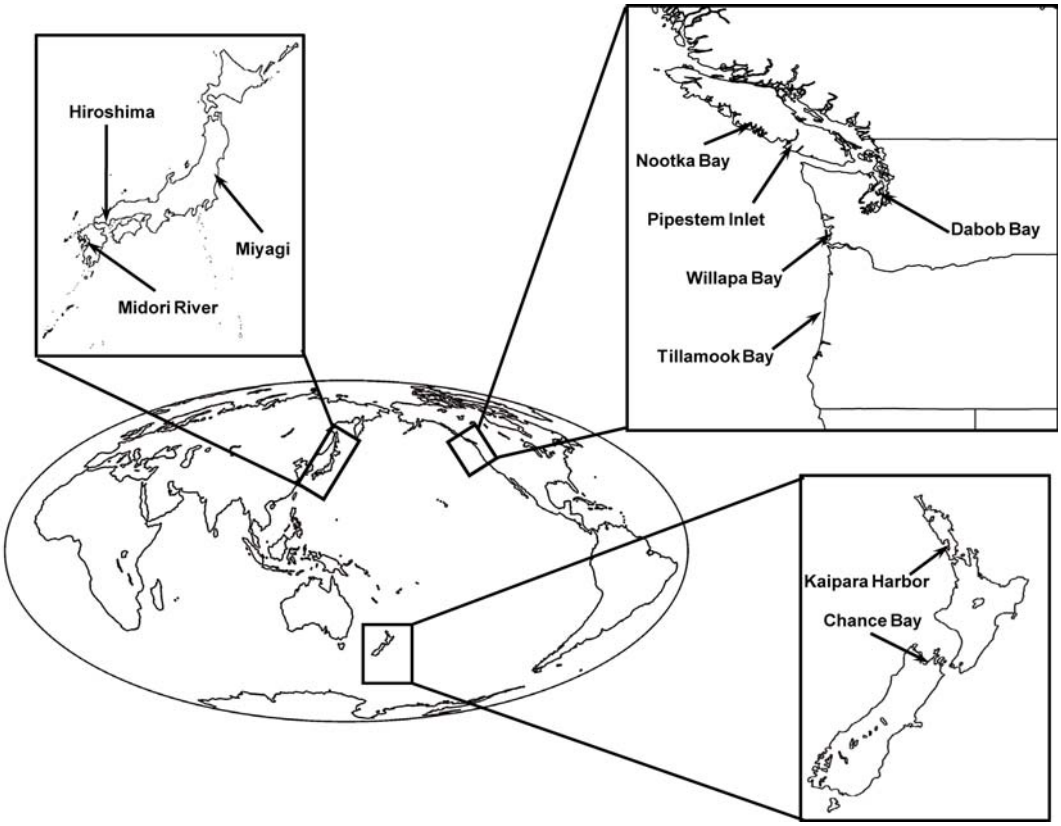

Fig. 1. Field sampling locations.

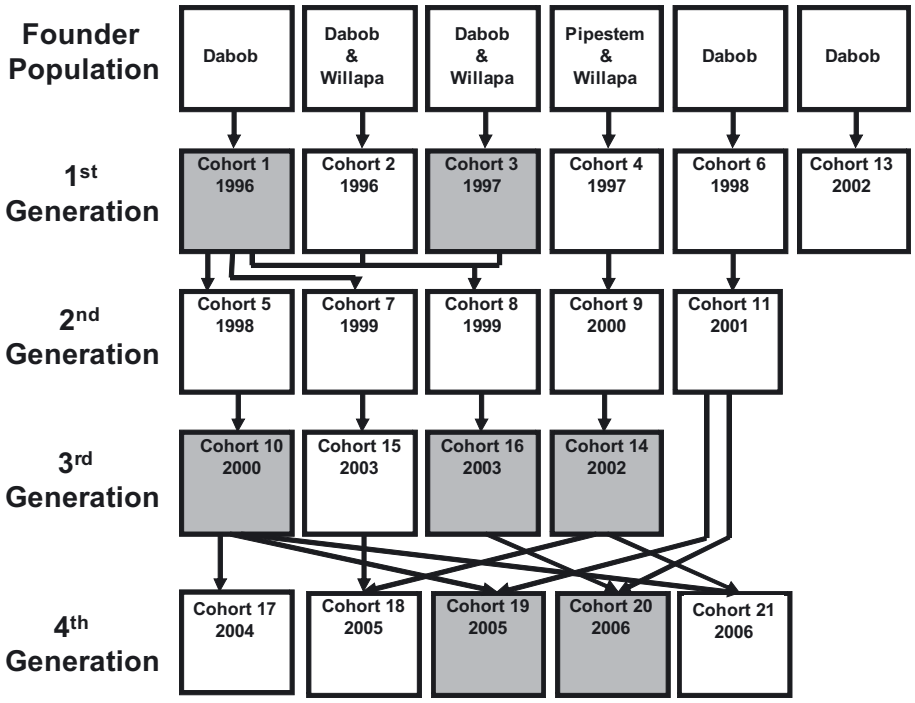

Fig. 2. Pedigree of the Molluscan Broodstock Program 1996–2006. Sampled cohorts are shaded gray.

with one additional nucleotide to reduce the number of amplified fragments (Eco+A; Mse+C). Selective primers consisted of the pre-selective primers with two or three additional selective bases added to the 3′ end for EcoR1 and Mse1 primers respectively. Four such primer combinations were used in this study (Eco+AAC/Mse+CTAG; Eco+ACC/Mse+CACT; Eco+ACC/Mse+CTAT; and Eco+ACT/Mse+CACG).

The PCR reaction mixtures for pre-selective and selective amplification were the same, with a total reaction volume of 10 μl consisting of 2 μl 5X Taq DNA polymerase buffer (Promega, Madison WI, USA) 0.6 μl 25 mM MgCl, 0.2 μl 10mM dNTP mixture, 0.5 μl 10 μl M of each primer, 0.1 μl 5 U/μl Taq polymerase, and 2.5 μl template DNA (1:20 dilution of the digestion-ligation products or pre-selective PCR product). The PCR protocol for pre-selective amplification was 5 min of denaturing at 94 °C, 72 °C for 2 min, followed by 20 cycles of 94 °C for 30 s, 56 °C for 1 min, 72 °C for 1 min, and a final extension at 72 °C for 5 min. Selective amplification was initiated with 5 min of denaturing at 94 °C, followed by 12 "touchdown" cycles of denaturing at 94 °C for 30 s, 30 s of annealing at temperatures beginning at 66 °C and decreasing by 1 °C each cycle, and 72 °C for 1 min, which were followed by an additional 25 cycles consisting of 94 °C for 30 s, 30 s at 56 °C, 72 °C for 1 min, and a final extension at 72 °C for 5 min. Fragments from selective amplifications were visualized using an ABI 3730XL Genetic Analyzer by first mixing each sample with 0.1 μl of LIZ size standard (Applied Biosystems, Foster City CA, USA) and 5 μl deionized formamide, followed by denaturing at 95 °C for 5 min, and then rapid cooling on ice. Raw data were analyzed with GeneMapper Analysis Software (Applied Biosystems, Foster City CA, USA). Fragments ranging from 50 to 500 base pairs were scored as dominant markers (i.e. absence = aa, presence = AA or Aa).

2.3 Data analysis

I used the GenAlEx add-in for Microsoft Excel (Peakall and Smouse 2006) to tabulate the number of polymorphic loci (i.e. band present frequency between 5 and 95%) both overall and within each of the 17 samples. I used AFLPSurv (Vekemans et al. 2002) to estimate the expected locus-specific and average heterozygosities both overall and within populations under the assumption of Hardy-Weinberg equilibrium using a Bayesian approach developed by Zhitovsky (1999), to estimate pair-wise genetic distances (Nei's D) and pair-wise F_{st} between all samples, and to estimate the multi-locus genetic divergence between all pairs of individuals using the approaches of Lynch and Milligan (1994). I tested for differences in the mean locus-specific expected heterozygosity between all populations and between wild Japanese populations, naturalized populations, and domesticated cohorts using analysis of variance in which the 506 locus-specific expected heterozygosities were treated as independent observations of heterozygosity within each of the samples using PROC GLM in SAS version 9.2 (SAS 2004).

The Lynch and Milligan method estimates relatedness coefficients (r_{ab} = mean probablility over all scored loci that a pair of individuals share alleles identical by descent) between all pairs of individuals, but because r_{ab} ranges from zero to one, $1-r_{ab}$ is an appropriate estimate of pair-wise "non-relatedness" or genetic divergence. Estimates of $1 - r_{ab}$ are probabilistic and must be estimated relative to specific reference allele frequencies. As a consequence, the use of different reference allele frequency distributions produces different results related to different population-level phenomena. The two appropriate reference frequency distributions in this context are the allele frequencies when all populations are pooled and the sample-specific allele frequencies within each population, which provide estimates of the pair-wise genetic divergence between individuals relative to the overall gene pool and relative to the appropriate within-sample gene pools respectively. Therefore, I ran these analyses first on the entire dataset to estimate genetic divergence using the overall allele frequencies and again on each of the 17 samples separately to use population-specific reference allele frequencies.

Because each individual's multi-locus genotype enters into numerous pair-wise estimates of genetic divergence between individuals, large subsets of such estimates within the dataset are not statistically independent. I therefore tested for differences in the mean genetic divergence coefficients between samples or between the means of wild Japanese populations, naturalized populations, and domesticated cohorts (including a priori contrasts between the three population types in the latter analysis) using permutation tests that do not assume independent observations with a specificied distribution. For these tests, I first reduced the AFLP-Surv output consisting of all possible pair-wise genetic divergence values between individuals regardless of their sample of origin to only estimates involving pairs of individuals within samples. I next used a SAS macro developed by Cassell (2002) to generate 1000 pseudo-datasets in which the values of $1 - r_{ab}$ were randomly reassigned to samples without replacement and the SAS GLM procedure was invoked to calculate F-ratios testing the hypothesis that the mean pair-wise genetic divergence between individuals differed among populations (or population types) for each of the permuted datasets. The resulting distribution of 1000 F-ratios for each hypothesis tested represents the distribution of F under the null hypothesis of no differences between samples (or sample types), and I estimated the p-value associated with each hypothesis test as the proportion of the null distribution of F-ratios greater than or equal to that of the non-permuted data.

I also used AFLP-Surv to generate 10 000 bootstrapped matrices of Nei's genetic distance between all pairs of samples and constructed a consensus neighbor-joining tree, estimated bootstrap support for each node in that tree, and visualized the tree using the NEIGHBOR, CONSENSE and DRAWTREE modules of the PHYLIP program (Felsenstein 1989, 1993).

3 Results

The four selective primer pairs amplified 506 scorable fragments. Approximately 98% (495 of 506) were polymorphic in the full sample, but only ~13–18% were polymorphic within samples (Fig. 3a; Table S1). Samples had between 1 and 25 private bands (i.e. bands that occurred only in that sample; Fig. 3b; Table S1) with domesticated MBP cohorts in general having fewer private bands than either wild Japanese or

naturalized populations. The proportion of polymorphic bands was slightly higher in wild and cultured populations than in naturalized populations, but because there is no level of sampling that provides replication, this was not tested statistically. Expected heterozygosity followed a similar pattern (Fig. 3c; Table S1), but one-way ANOVA of all 17 populations found no significant differences ($p = 0.169$). However, when populations were classified as wild, naturalized or domesticated, ANOVA revealed significant differences between population types (Fig. 4a; Table S2; $p = 0.0191$), and pair-wise contrasts testing for differences between the three population types found that naturalized populations had lower mean heterozygosity than domesticated ones ($p = 0.0066$), but the difference between Japanese wild and naturalized populations was not significant ($p = 0.0839$) nor was the difference between Japanese wild and domesticated ($p = 0.706$).

Mean coefficients of genetic divergence estimated using either the entire sample of all populations as reference allele frequencies or sample-specific within sample reference allele frequencies varied significant among populations (Fig. 3d, e; Table S1; $p < 0.001$ in both cases), but the patterns of variation were different – in fact opposite. Coefficients of genetic divergence estimated using all populations as a reference were higher in naturalized and wild populations than in domesticated populations (Fig. 3d; 4b; $p < 0.001$), but using the within-sample allele frequencies as a reference, all pair-wise contrasts were significant with wild and populations having substantially lower mean divergence coefficients than naturalized and cultured populations (Figs. 3e, 4c; $p < 0.001$).

Expected heterozygosity partitioned into components of 92% within populations (Hw) and 8% between populations (Hb). Overall Fst was estimated as 0.0764 and differs significantly from zero based on permutation tests conducted using AFLPSurv ($p = 0.007$). Table 1 lists the number of individuals in each sample as well as all pair-wise values of Nei's D and F_{st} between the 17 samples.

A neighbor-joining tree based on Nei's genetic distance between populations (Fig. 5) reveals two main branches. One includes all of the naturalized populations except for Tillamook Bay and the Midori River wild population, and the other contains all of the MBP cohorts, the Myagi and Hiroshima wild populations, and the naturalized Tillamook Bay population. These branches have high bootstrap support.

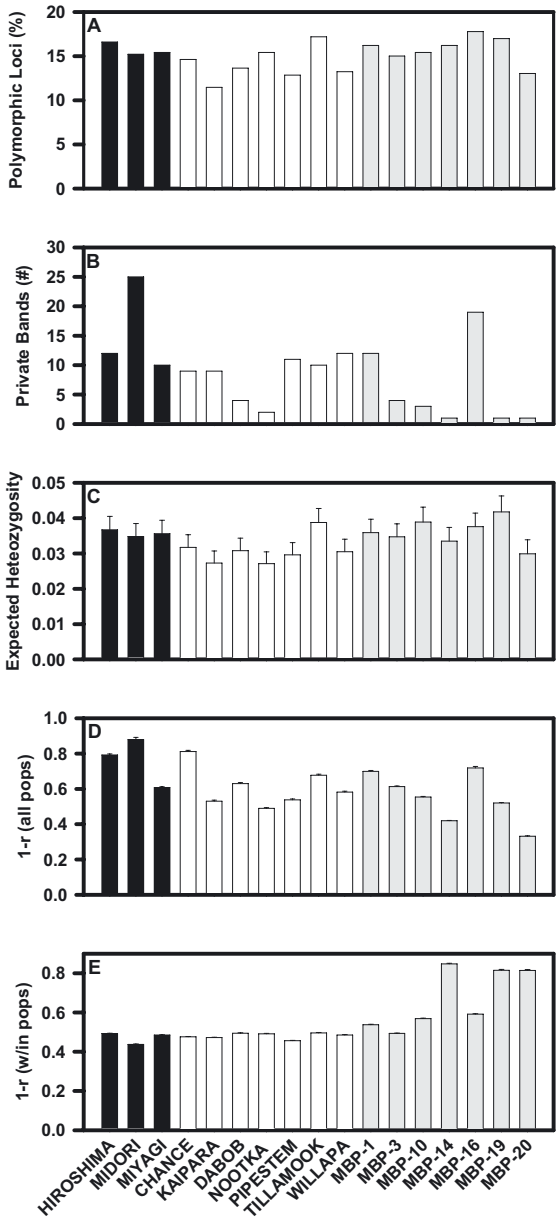

Fig. 3. Values of various parameters for each sample: A. Percentage of markers polymorphic, B. Number of private bands, C. Expected heterozygosity, D. Mean pair-wise divergence between individuals ($1-r$) based on allele frequencies from the entire sample, E. Mean pair-wise divergence between individuals ($1-r$) based on sample-specific allele frequencies. Black bars are wild Japanese populations; white bars naturalized populations, and gray bars domesticated cohorts. Error bars are standard errors of the mean either across all loci (C) or across all possible pairs of individuals w/in a sample (D, E).

4 Discussion

I used AFLP markers to examine whether and how molecular genetic variation has been affected by the introduction and naturalization of Pacific oysters from Japan to novel environments on the West Coast of North America and New Zealand and by the development of domesticated and artificially selected strains from these populations. Similar to previous studies that used allozyme markers (English et al. 2000) and a combination of allozymes and microsatellites (Appleyard and Ward 2006) to compare naturalized populations in Australia to Japanese wild populations, I found that despite a slight decrease in the number of polymorphic loci in naturalized populations relative to their wild progenitors (Fig. 3a), naturalization has had very little impact on expected heterozygosity (Figs. 3c, 4a). The expected heterozygosity of naturalized populations did not differ significantly from wild Japanese populations. However, in contrast with the studies of Australian selectively bred Pacific oysters (Appleyard and Ward 2006), domesticated and selectively bred cohorts in the US have higher expected heterozygosity than naturalized populations, most likely because the MBP mating

Table 1. Sample size (n), Nei's genetic distance (D) between all pairs of samples (above diagonal), and F_{st} (below diagonal).

Sample	n	HIROSHIMA	MIDORI	MIYAGI	CHANCE	KAIPARA	DABOB	NOOTKA	PIPESTEM	TILLAMOOK	WILLAPA	MBP-1	MBP-3	MBP-10	MBP-14	MBP-16	MBP-19	MBP-20
HIROSHIMA	50	–	0.0025	0.0000	0.0041	0.0041	0.0045	0.0045	0.0045	0.0004	0.0045	0.0008	0.0007	0.0021	0.0014	0.0010	0.0021	0.0028
MIDORI	50	0.0627	–	0.0024	0.0005	0.0004	0.0008	0.0009	0.0009	0.0025	0.0007	0.0027	0.0026	0.0039	0.0044	0.0023	0.0049	0.0052
MIYAGI	50	0.0011	0.0618	–	0.0039	0.0040	0.0039	0.0039	0.0038	0.0000	0.0039	0.0006	0.0004	0.0014	0.0009	0.0006	0.0018	0.0023
CHANCE	50	0.1031	0.0139	0.1016	–	0.0000	0.0003	0.0004	0.0003	0.0039	0.0003	0.0044	0.0042	0.0050	0.0062	0.0036	0.0066	0.0070
KAIPARA	50	0.1097	0.0135	0.1091	0.0012	–	0.0006	0.0005	0.0005	0.0040	0.0005	0.0044	0.0043	0.0055	0.0064	0.0036	0.0068	0.0073
DABOB	50	0.1141	0.0237	0.1023	0.0100	0.0191	–	0.0002	0.0000	0.0035	0.0001	0.0043	0.0040	0.0044	0.0058	0.0035	0.0060	0.0066
NOOTKA	50	0.1208	0.0262	0.1087	0.0130	0.0163	0.0054	–	0.0000	0.0037	0.0000	0.0040	0.0038	0.0046	0.0057	0.0033	0.0062	0.0072
PIPESTEM	50	0.1163	0.0255	0.1023	0.0097	0.0161	0.0000	0.0000	-	0.0035	0.0000	0.0042	0.0038	0.0045	0.0056	0.0032	0.0060	0.0067
TILLAMOOK	51	0.0106	0.0604	0.0000	0.0954	0.1051	0.0894	0.0975	0.0904	–	0.0035	0.0007	0.0005	0.0013	0.0010	0.0007	0.0018	0.0022
WILLAPA	49	0.1148	0.0200	0.1017	0.0097	0.0153	0.0039	0.0013	0.0000	0.0890	–	0.0042	0.0038	0.0045	0.0057	0.0032	0.0060	0.0066
MBP–1	64	0.0207	0.0697	0.0171	0.1120	0.1201	0.1114	0.1091	0.1104	0.0178	0.0042	–	0.0003	0.0012	0.0017	0.0006	0.0022	0.0029
MBP–3	64	0.0191	0.0678	0.0114	0.1079	0.1185	0.1052	0.1071	0.1033	0.0138	0.0038	0.0003	–	0.0011	0.0011	0.0003	0.0018	0.0025
MBP–10	64	0.0499	0.0924	0.0342	0.1211	0.1387	0.1093	0.1194	0.1136	0.0307	0.0045	0.0012	0.0011	–	0.0021	0.0011	0.0022	0.0040
MBP-14	87	0.0374	0.1113	0.0235	0.1544	0.1695	0.1489	0.1549	0.1454	0.0250	0.0057	0.0017	0.0011	0.0021	–	0.0016	0.0020	0.0035
MBP-16	64	0.0265	0.0582	0.0164	0.0910	0.0979	0.0891	0.0897	0.0853	0.0164	0.0032	0.0006	0.0003	0.0011	0.0016	-	0.0021	0.0025
MBP-19	64	0.0483	0.1100	0.0418	0.1481	0.1608	0.1372	0.1492	0.1388	0.0405	0.0060	0.0022	0.0018	0.0022	0.0020	0.0021	–	0.0035
MBP-20	60	0.0755	0.1353	0.0636	0.1807	0.1994	0.1731	0.1968	0.1794	0.0583	0.0066	0.0029	0.0025	0.0040	0.0035	0.0025	0.0035	–

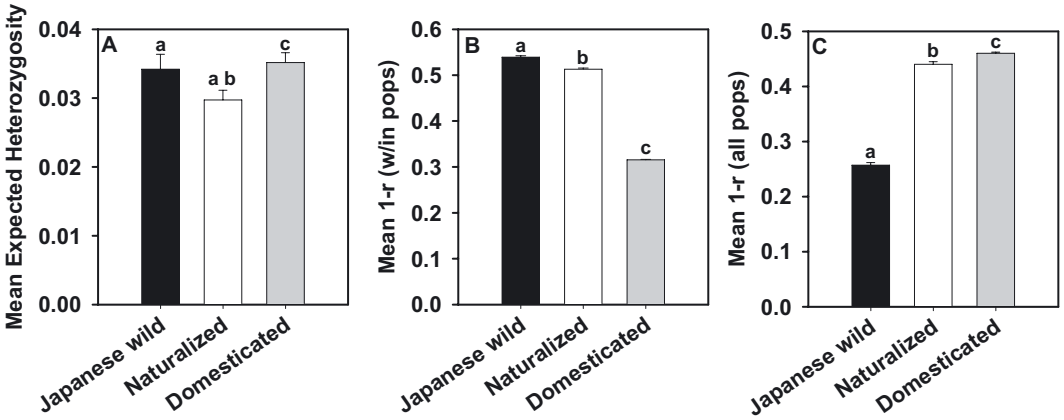

Fig. 4. A. Mean expected heterozygosity for wild, naturalized and domesticated (i.e. MBP) samples , B. mean pair-wise genetic divergence ($1 - r$) between individuals based on sample-specific allele frequencies, and C. mean pair-wise genetic divergence ($1 - r$) between individuals based on allele frequencies from the entire sample. Error bars are standard errors. Differences between bars labeled with different letters are statistically significant ($p < 0.05$).

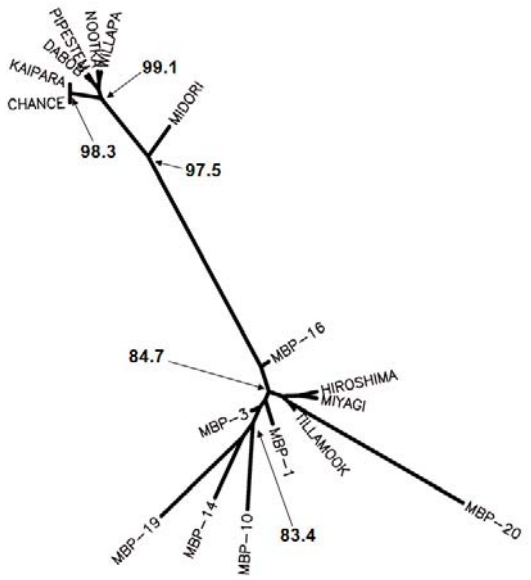

Fig. 5. Neighbor-joining tree based on Nei's genetic distance. Numbers are bootstrap values for all nodes with >80% support.

scheme consciously avoids consanguineous matings and the equalization of parental contributions to these closed populations has effectively controlled inbreeding, at least in the first few generations of selective breeding. However both naturalization and domestication have, on average (with the notable exception of MBP cohort 16), reduced the number of private bands in domesticated strains and influenced the mean pair wise divergence between individuals within both naturalized and domesticated populations. When pair-wise genetic divergence is estimated relative to overall allele frequencies, despite some variation among samples (Fig. 3a), wild populations in Japan have significantly higher mean genetic divergence between individuals than either naturalized or domesticated populations (Fig. 4b), indicating that if the entire group of samples is viewed as a single gene pool, the probability of matings between relatives is higher in both naturalized and domesticated populations than in wild populations. On the other hand, when genetic divergence between individuals is estimated relative to population-specific allele frequencies, domesticated populations have significantly higher mean coefficients of genetic divergence than both wild and naturalized populations (Fig. 4c), again indicating that if the various samples are viewed as distinct gene pools, consanguineous matings in the first few generations since they were separated from their naturalized parent populations are less probable than in naturalized populations as a consequence of effective inbreeding management practices in the breeding program.

While it may seem paradoxical that domesticated cohorts initiated using limited numbers of founding parents from naturalized populations would be more heterozygous and have higher mean genetic divergence between individuals than their source populations (Figs. 4a and c respectively), this result is consistent with previous studies of naturalized populations of Pacific oyster. Hedgecock and Coworkers demonstrated that as a consequence of high temporal variance in reproductive success, the variance effective population size of the Dabob Bay population is extremely low (Hedgecock et al. 1992; Hedgecock 1994), which would decrease both heterozygosity and average pair-wise genetic divergence between individuals. This is most likely a consequence of environmental conditions in Pacific Northwest estuaries being at or near the physiological limits for successful reproduction and larval development in *C. gigas*, making recruitment highly stochastic with few individuals reproducing successfully in any given year (i.e. "sweepstakes" recruitment). In contrast, wild populations living in their native range under conditions that are presumably more favorable for reproduction and larval development do not always exhibit sweepstakes recruitment (e.g. Taris et al. 2009), and in domesticated MBP populations the contributions of parents are equalized. In both of these cases, higher heterozygosity and average genetic divergence would be expected.

Surprisingly, colonizing new habitats in both the Pacific Northwest and New Zealand appears to have shifted the

genetic composition of all naturalized populations away from that of their Miyagi source population towards a more "Midori River-like" configuration and domestication has consistently reversed this tendency. A neighbor joining tree based on Nei's genetic distance between populations (Fig. 5) unambiguously grouped all but one naturalized population (Tillamook), including those in New Zealand, with the Midori River wild population rather than with their progenitors in the Myagi region of Japan. The Tillamook population is a special case in that it is a very recent colonization that resulted from the successful spawning of farmed oysters. Bootstrap support for this major division is very high.

Even more surprisingly, all seven domesticated MBP cohorts grouped with the Miyagi and Hiroshima wild populations rather than with the naturalized populations from which they were derived, again with very high boostrap support.

These analyses cannot rigorously test whether the observed patterns of genetic differentiation are a consequence of genetic drift or of natural and artificial selection during naturalization and domestication respectively. And because the relationships among these "populations" consist of a complex combination of intentional massive-scale and long-distance human translocations that ceased decades ago, current gene flow between populations mediated by geographical and oceanographic factors, and a mixture of natural and artificial selection, these data are not suitable for genome scan analyses to identify specific loci under selection by comparing locus-specific genetic divergence measures to null distributions based mainly on current patterns of gene flow in the absence of selection (Bonin et al. 2007; Excoffier et al. 2009). Even so, under pure genetic drift, a neighbor-joining tree based on genetic distances is expected to mirror the historical pattern of population splitting through naturalization and domestication, but this expectation is not supported by our data. All but one naturalized population in both North America and New Zealand grouped together, and none branch directly from their known Miyagi progenitors (Fig. 5). The one exception (Tillamook) is a very recent naturalization event derived from domesticated stocks – an exception that actually supports the rule.

Similarly, the tight grouping of all seven domesticated MBP cohorts and the Tillamook population and the long and well supported branch between them and the naturalized populations from which they were derived (Fig. 5) are also unexpected under random genetic drift. The two founding cohorts I studied (MBP-1 and MBP-3; Fig. 2) were independently created by randomly sampling naturalized populations (n = 64; Table 1), so the long branch lengths between these cohorts and the very similar Dabob Bay and Willapa Bay populations from which they were derived (Fig. 5), are inconsistent with random genetic drift. Second-generation MBP cohorts derived from these two base or founder populations (MBP-10, MBP-16) did not diverge so dramatically from their parental cohorts, and the other second generation cohort studied (MBP-14) is genetically similar to them even though it can be traced back to a third random sample of founders from a different combination of naturalized populations (Cohort 4, Fig. 2). The two third-generation cohorts (MBP-19, MBP-20; Fig. 2), however are more divergent from the previous generation (MBP-10; MBP-14, MBP-16) as indicated by the long branch lengths.

While speculative, it is reasonable to hypothesize that natural selection and domestication selection have played important roles in adapting this species to both novel natural environments and to the artificial hatchery environment. This could explain the strong divergence between naturalized populations and their progenitors in Japan, the genome-level similarity between naturalized populations in similar environmental conditions, and a very similar pattern of non-random divergence following domestication and artificial selection. The genome-wide similarity between US domesticated Pacific oysters and native populations in Hiroshima and Miyagi also makes these populations interesting as potential sources of novel germplasm for hatchery-based culture in the US, but further characterization at the phenotypic level is necessary.

5 Conclusion

Naturalization and domestication have altered the genetic diversity of Pacific oysters in surprisingly consistent ways. An analysis of AFLP genotype frequencies shows that all but one naturalized population in both the North American Pacific Northwest and New Zealand derived from the Miyagi region of Japan have evolved to more closely resemble the native populations in the Midori River/Ariake Sea region and the only exception is a very recent colonization event in which domesticated stocks have successfully reproduced. Domestication and selective breeding, in contrast, have produced captive germplasm populations more genetically similar to native populations in the Hiroshima and Miyagi regions of Japan than to the naturalized populations in the US from which they were derived. All of the domesticated strains examined have Miyagi-like allele frequencies that differ substantially from the Midori River-like naturalized populations in the US Pacific Northwest and British Columbia. The consistency of these patterns also invites speculation, but does not conclusively demonstrate that these changes are driven by natural and artificial selection. Further investigation is required to determine whether the Miyagi and Hiroshima populations have favorable characteristics for aquaculture production in the US.

Acknowledgements. The author would like to thank Jim Krenz and Sean Matson for collecting samples from US naturalized populations, Chris Langdon for samples from Midori region, Masashi Sekino for samples from the Miyagi region, Brian Kingzett and Jennifer Dawson for samples from British Columbia, and Nick King and Serean Adams for samples from New Zealand. Chris Langdon provided archived tissue samples from the Molluscan Broodstock Program. Jim Krenz did the vast majority of the DNA extraction and AFLP genotyping work and Crystal Rink did the rest. This research was supported by USDA-ARS programmatic funding to the Shellfish Genetics Program (CRIS Project #5358-31000-001-00D). Any use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable. USDA is an equal opportunity provider and employer.

Supporting information

Table S1: Data for Fig. 3.
Table S2: Data for Fig. 4.

References

Appleyard S.A., Ward R.D., 2006, Genetic diversity and effective population size in mass selection lines of Pacific oyster (*Crassostrea gigas*). Aquaculture 254, 148–159.

Bonin A., Ehrich D., Manel S., 2007, Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. Mol. Ecol. 16, 3737–3758.

Breese W.P., Wick W.Q., 1974, Oyster farming: culturing, harvesting, and processing a product of the Pacific coast area. Extension Marine Advisory Program, Marine Science Publication SG Number 13, Newport, OR, Oregon State University, p. 8.

Cassell D.L., 2002, A Randomization-test Wrapper for SAS®PROCs, Proceedings of the Twenty-Seventh Annual SAS®Users Group International Conference, Cary, NC, SAS Institute Inc., http://www2.sas.com/proceedings/sugi27/p251-27.pdf.

Clark J.E., Langmo R., 1979, Oyster seed hatcheries on the US west coast: an overview. Mar. Fish. Rev. 41, 10–16.

English L.J., Maguire G.B., Ward R.D., 2000, Genetic variation of wild and hatchery populations of the Pacific oyster, *Crassostrea gigas* (Thunberg), in Australia. Aquaculture 187, 283–298.

Excoffier L., Hofer T., Foll M., 2009, Detecting loci under selection in a hierarchically structured population. Heredity 103, 285–298.

Feely R.A., Sabine C.L, Lee K., Berelson W., Kleypas J., Fabry V.J., Millero F.J., 2004, Impact of Anthropogenic CO2 on the CaCO3 System in the Oceans. Science 305, 362–366.

Felsenstein J., 1989, PHYLIP – Phylogeny Inference Package (version 3.2). Cladistics 5, 164–166.

Felsenstein J., 1993, PHYLIP (Phylogeny Inference Package) version 3.5c, Distributed by the author. Department of Genetics, University of Washington, Seattle.

Hedgecock D., 1994, Does variance in reproductive success limit effective population sizes of marine organisms? In: Beaumont A.R. (Ed.), Genetics and Evolution of Aquatic Organisms. London, Chapman Hall, p. 122–134.

Hedgecock D., Chow V., Waples R.S., 1992, Effective population numbers of shellfish brood stocks estimated from temporal variances in allelic frequencies. Aquaculture 108, 215–232.

Hershberger W.K., Perdue J.A., Beattie J.H., 1984, Genetic selection and systematic breeding in Pacific oyster culture. Aquaculture 39, 237–245.

Jones G., Jones B., 1983, Methods for setting hatchery produced oyster larvae, Marine Resources Branch, Ministry of Environment, British Columbia, p. 61.

Langdon C., Evans F., Jacobson D., Blouin M., 2003, Yields of cultured Pacific oysters *Crassostrea gigas* Thunberg improved after one generation of selection. Aquaculture 220, 227–244.

Li L., Guo X., 2004, AFLP-based Genetic linkage maps of the Pacific oyster *Crassostrea gigas* Thunberg. Mar. Biotechnol. 6, 26–36.

Lynch M., Milligan B.G., 1994, Analysis of population genetic-structure with RAPD markers. Mol. Ecol. 3, 91–99.

Newkirk G.F., 1978, A discussion of possible sources of inbreeding in hatchery stock and associated problems. Proc. World Aquac. Soc. 9, 93–100.

Peakall R., Smouse P.E., 2006, genalex 6: genetic analysis in Excel, Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.

SAS, 2004, SAS OnlineDoc®, Version 9.1 http://support.sas.com/documentation/onlinedoc. Cary, NC, SAS Institute Inc.

Taris N., Boudry P., Bonhomme F., Camara M.D., Lapegue S., 2009, Mitochondrial and nuclear DNA analysis of genetic heterogeneity among recruitment cohorts of the European flat oyster *Ostrea edulis*. Biol. Bull. 217, 233–241.

Troost K., 2010, Causes and effects of a highly successful marine invasion: case-study of the introduced Pacific oyster *Crassostrea gigas* in continental NW European estuaries. J. Sea Res. 64, 145–165.

Vekemans X., Beauwens T., Lemaire M., Roldan-Ruiz I., 2002, Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Mol. Ecol. 11, 139–151.

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M., 1995, AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23, 4407–4414.

Wang H., Guo X., 2008, Identification of *Crassostrea ariakensis* and related oysters by multiplex species-specific PCR (polymerase chain reaction), J. Shellfish Res. 27, 481–487.

Zhivotovsky L.A., 1999, Estimating population structure in diploids with multilocus dominant DNA markers. Mol. Ecol. 8, 907–913.