Journal of Shellfish Research, Vol. 23, No. 2, 379–385, 2004.

WIDESPREAD NULL ALLELES AND POOR CROSS-SPECIES AMPLIFICATION OF MICROSATELLITE DNA LOCI CLONED FROM THE PACIFIC OYSTER, *CRASSOSTREA GIGAS*

DENNIS HEDGECOCK,* GANG LI, SOPHIE HUBERT, KATHERINE BUCKLIN AND VANESSA RIBES

University of Southern California, Department of Biological Sciences, Los Angeles, California 90089

ABSTRACT Non-amplifying, PCR-null alleles are detected at 49 (51%) of 96 microsatellite DNA markers tested for Mendelian segregation in three families of the Pacific oyster *Crassostrea gigas* Thunberg. The average frequency of null alleles among F_1 hybrid grandparents is 0.093. The frequency of null alleles suggests a high level of sequence polymorphism in PCR primer-binding sites and yields a conservative estimate of one single nucleotide polymorphism (SNP) every 82 base pairs. Among 86 markers tested on congeneric species, 83 (96.5%) are likely to be useful markers for the Portuguese oyster *Crassostrea angulata*, 71 (82.6%) for the Kumamoto oyster *C. sikamea*, 31 (36.0%) for the Suminoe oyster *C. ariakensis*, and only 11 (12.8%) for the Eastern oyster *C. virginica*. PCR product-yield and mean numbers of alleles per locus also decline significantly across this series of congeneric species, which separated from the Pacific oyster <1, ~2, ~4, and >5 million years ago, respectively. Decline in cross-specific PCR yield does not depend on microsatellite repeat-motif but is correlated with the frequency of null alleles across loci. The high nucleotide diversity suggested by these observations for the oyster may be a by-product of high fecundity, consistent with G. C. Williams' (1975) Elm-Oyster evolutionary model and experimental evidence for a high mutational load. Microsatellite loci should be identified *de novo* for each species of cupped oyster, and their inheritance should be validated before use in population analyses. Homology of microsatellite loci among related species should be confirmed by sequencing of flanking regions.

KEY WORDS: Pacific oyster, microsatellite DNA, null alleles, cross-specific amplification, nucleotide polymorphism, *Crassostrea gigas*

INTRODUCTION

With the completion or impending completion of genome sequences for the human, fruit fly, nematode, mouse, and other eukaryotic genetic models, much attention is being paid to DNA sequence polymorphism and its potential use in understanding the genetic basis of complex phenotypes, such as disease susceptibility (Zwick et al. 2000). Whereas nucleotide diversity is becoming very well described for model organisms, which are all low-fecundity species ($<10_3 - 10_4$ eggs per female), little is known about DNA polymorphism in high-fecundity species (>106 eggs per female). We might expect highly fecund species to have high nucleotide diversity, owing to large population sizes, and perhaps higher mutation rates. Indeed, G. C. Williams (1975) argued with his Elm-Oyster Model that highly fecund species with high early mortality (Type-III survivorship) should reproduce sexually, show tremendous variation in individual fitness, and carry a large load of recessive deleterious mutations.

A large load of recessive deleterious mutations has recently been confirmed for the European flat oyster *Ostrea edulis* (Bierne et al. 1998) and the Pacific oyster *Crassostrea gigas* Thunberg (Launey & Hedgecock 2001, Bucklin 2002). Oysters naturally carry dozens of highly deleterious recessive mutations, which explain widespread observations of heterosis for fitness-related traits in bivalve mollusc species, at the whole organism and genetic-marker levels, and distortions of Mendelian segregation ratios at marker loci (Launey & Hedgecock 2001). On a practical level, discovery of genetic load in bivalves suggests that marker inheritance and linkage should be confirmed early in larval development, before selection can substantially distort genotypic proportions. Typing 11-day-old larvae and using double-hybrid crosses to reduce homozygosity by descent and inbreeding depression in mapping families, Hubert & Hedgecock (2004) have produced the first low-density microsatellite DNA marker maps for the Pacific oyster.

Microsatellite DNA markers are short tandem repeats of nucleotide motifs, 2 to 6 base pairs (bp) in length, which are distributed throughout the genome in prokaryotes and eukaryotes (Chambers & MacAvoy, 2000). Because they are highly polymorphic, microsatellites have been widely used as genetic markers for studies of linkage, kinship, and population structure (Goldstein & Schlötterer 1999, Chambers & MacAvoy 2000). Presently, the DNA sequences of 369 microsatellite containing clones from *C. gigas* are deposited in GenBank. Of these, 123 have been developed into PCR-amplifiable markers with confirmed inheritance (Magoulas et al. 1998, Huvet et al. 2000, McGoldrick et al. 2000, Li et al. 2003, Sekino et al. 2003) and 100 have been placed on linkage maps for the Pacific oyster (Hubert & Hedgecock 2004). Here, we present data on polymorphism of 96 of these microsatellite DNA markers and show that there seems to be little or no dependence of polymorphism on repeat-motif or motif complexity.

In developing microsatellite markers for constructing a linkage map of the Pacific oyster, we also uncovered two lines of evidence that nucleotide diversity and rate of sequence evolution in cupped oysters may be extremely high. Because high nucleotide diversity has important implications for future genetic and genomic studies with oysters, we present these findings here. First, we show that there is a high frequency of nonamplifying PCR null alleles, which likely result from polymorphism in the nonrepetitive flanking sequences to which PCR primers are designed to anneal. The inheritance of these null alleles is confirmed in multigenerational families, which were derived from the same population from which the microsatellite markers were cloned. Second, we show a dramatic decay in ability to amplify these markers across a series of four congeneric species that diverged from the Pacific oyster from <1 to >5 million years ago. This decay in cross-species amplification is independent of microsatellite repeat-motif but is correlated with the frequency of null alleles across loci. These observations contrast sharply with reports of success in amplifying microsatellites from very divergent vertebrate taxa (Garza et al. 1995, Pépin et al. 1995, Schlötterer et al. 1991, FitzSimmons et al. 1995, Rico et al.

*Corresponding author. Email: dhedge@usc.edu

1996) and species groups of *Drosophila* flies (Colson et al. 1999, Huttunen & Schlötterer 2002).

MATERIALS AND METHODS

Microsatellite Markers

Inheritance and polymorphism of 96 microsatellite markers was determined for three, multigenerational families that were used to construct linkage maps for the Pacific oyster (see Hubert & Hedgecock 2004). Of the markers that were mapped, 79 come from Li et al. (2003) and 17 come from previous publications (13 listed in Table 1 plus *cmrCgi61*, *cmrCgi141*, *um2Cgi10*, and *um2Cgi48*; Magoulas et al. 1998; Huvet et al. 2000; McGoldrick et al. 2000; see Hubert & Hedgecock 2004). Eighty-six of the 96 markers were further tested in cross-specific PCR amplifications.

Polymorphism and Nonamplifying Null-Allele Frequencies in C. Gigas

The *C. gigas* used for this study were F_2 or F_3 hybrid parents of the three mapping families, $(7 \times 6) \times (5 \times 2)$, $(2 \times 5) \times (7 \times 9)$, $(7 \times 9) \times (2 \times 5)$, their F_1 grandparents (or, for hybrid line 7×6, the F_1 great-grandparents), and other individuals from inbred lines under investigation (Bucklin 2002, Hubert & Hedgecock 2004). These lines were derived from the same Dabob Bay, WA, population, from which the microsatellite DNA clones were obtained (Li et al. 2003). DNA of the F_1 individuals was available from a previous study (Launey & Hedgecock 2001). Whereas families 2×5 and 5×2 share great grandparents (from inbred lines 92-2 and 89-5, respectively), families 7×6 and 7×9 are descended from lines 89-7 and 93-7, respectively, and are unrelated.

We typed the eight F_1 grandparents or great-grandparents of these four families to estimate the observed proportion of heterozygous individuals per locus, H_o, and the number of alleles per locus, n_a. Null alleles, which are included in n_a, were identified and confirmed by segregation analysis in mapping families. The minimum number of independent null alleles per locus (*#null* in Table 1 and Table 1 of Li et al. 2003) was tallied by tracing allele pedigrees back to inbred great-grandparents (for families 2×5 and 5×2) or, when allele pedigrees were incomplete (for 7×9 and 7×6), by assuming that multiple null alleles at a locus were identical by descent. Differences in H_o (arcsine square-root transformed), n_a, and *#null* among classes of microsatellite motifs were tested by ANOVA. Repeat-motifs were classified as di-, tri, or tetranucleotide repeats or alternatively cross-classified into categories of motif-complexity (simple or compound vs. pure or interrupted by nonmotif nucleotides; following Chambers & MacAvoy 2000, with compound and complex motifs combined).

TABLE 1.

Cross-species amplification success and number of null alleles for 13, previously published, Pacific oyster microsatellite loci.

Locus[a]	*C. angulata*	*C. sikamea*	*C. ariakensis*	*C. virginica*	*#null* in *C. gigas*[b]
cmrCg1[c]	+[e]	+	–	–	0
cmrCg3[c]	+	+	–	–	0
cmrCg5[c]	+	+	+	–	1
imbCg44[d]	+	+	–	–	3
imbCg49[d]	+	–	–	–	1
imbCg108[d]	+	–	–	–	0
ucdCg1[c]	+	+	+	+	0
ucdCg2[c]	+	+	–	+	0
ucdCg3[c]	+	–	–	–	0
ucdCg6[c]	+	–	–	–	0
ucdCg14[c]	+	+	–	–	1
ucdCg18[c]	?	+	+	+	2
ucdCg28[c]	+	+	+	–	1

[a] Loci names modified slightly from the original published name to conform to conventional nomenclature.

[b] Number of independent null-alleles in F_1 hybrid grand- or great grandparents

[c] McGoldrick et al. 2000; Launey & Hedgecock 2001

[d] Magoulas et al. 1998

[e] +, amplification; –, insufficient amplification; ?, slight amplification

Cross-species Amplifications

For cross-species comparison, we obtained DNA from five or six individuals of three other Asian species and the American species of *Crassostrea*, as described by Li et al. (2003). For each of 86 markers, PCR amplifications were done simultaneously for the five species, using a single PCR reaction mix, followed by a single acrylamide gel separation and fluorescent scan. PCR product was quantified by peak optical density per pixel (OD) in bands of similar size and appearance as bands observed in *C. gigas* on the same gel, using FMBIO software (Hitachi Genetic Systems). Owing to variable DNA template concentrations and a significant positive regression, within *C. gigas*, of OD on initial DNA concentration ($F_{1, 1135} = 9.95$, $P = 0.0017$), we used a natural log transformation, $adjOD = \ln[(OD + 10)/C]$, where C is template DNA concentration. Regression of *adj*OD, on C was not significant ($F_{1, 1135} = 1.02$, $P = 0.31$).

Variation in *adjOD* was analyzed with linear models (GLM procedure of SAS version 7, SAS Institute, Inc., Cary, NC) of completely randomized block design, with 5 species × 3 repeat-motif or 4 cross-classified categories of motif-complexity × 2 different observers working with different DNA extractions (blocks). There were 3051 observations in these ANOVAs for repeat motif and 2593 observations for motif-complexity categories. Significance of block and main effects was tested by the within-block variance of species-motif combinations, using Type III sums of squares. Relationship between null-allele counts within *C. gigas* and ability to amplify markers from congeneric species was determined by stepwise regression, with *adjOD* dependent on species and null-allele count.

Allelic diversity per locus was estimated by counting the number of unique bands (n_a) observed per marker, per species (over all markers and individuals, $n = 380$). Six 7×6 F_3 hybrid individuals were included in this analysis as an additional comparison. Analysis of these data must account for varying numbers of individuals successfully amplified, per species and PCR reaction (n_i), and an expected, positive correlation between n_a and n_i. The number of individuals amplified was used as a covariate in a 2-factor ANCOVA of species and repeat motif (6×3). Least-squares estimates of the n_a for each species, adjusted to a constant $n_i = 5$, were obtained with estimate statements in PROC GLM.

A qualitative and partially subjective assessment was made of whether each marker worked sufficiently well in another species to advise its use as a genetic marker for that species. A marker was judged likely to work for another species if it amplified from at least four of the five or six individuals surveyed and it produced bands of sufficient strength for genotypes to be reliably determined. These assessments are listed under the "cross species"

column of Table 1 of Li et al. (2003), for the 79 loci presented there, and are summarized here in Table 1, for previously published markers, as "+," for amplification, "–," for insufficient amplification, and "?," for slight amplification.

RESULTS

Heterozygosity and Null-allele Frequencies

The 96 microsatellite markers used for determination of polymorphism in *C. gigas* comprise 58 di-, 22 tri- and 16 tetra-nucleotide repeat-motif. Alternatively they can be classified into 4 crossed categories of motif-complexity: 41 are simple-pure, 23, simple-interrupted, 11, compound-pure and 21, compound-interrupted (Table 2). Interrupted repeat arrays comprise 45.8% (44/96) of the loci tested. These markers vary widely in degree of polymorphism, with numbers of alleles (n_a), ranging from 2 to 10 (mean, 5.77 ± 0.14) and observed heterozygosity (H_o), ranging from 0.125 to 1.0 (mean, 0.748 ± 0.047; see Table 2). Neither measure of polymorphism, however, varies significantly by repeat-motif (di-, tri-, and tetra-nucleotide motifs; for n_a, $F_{2,93} = 0.90$, $P = 0.41$ and for H_o, $F_{2,93} = 0.81$, $P = 0.45$) or motif-complexity category (simple-pure, compound-pure, simple-interrupted, compound-interrupted; for n_a, $F_{2,93} = 1.77$, $P = 0.16$, and for H_o, $F_{2,93} = 1.33$, $P = 0.27$) in our sample of F_1 hybrids.

Non-amplifying PCR-null alleles are found at 49 (51%) of the 96 loci. Minimum counts of independent PCR-null alleles per locus, in the F_1 grandparents or great grandparents of the mapping families, range from 0 to 4 and sum, over all loci, to 75 (Fig. 1). The average number of independent alleles per locus is 8.6, and the average frequency of null alleles is 0.093. Neither null-allele frequency nor null-allele count (see Table 2) varies among repeat-motifs ($F_{2,93} = 1.53$, $P = 0.22$) or motif-complexity categories ($F_{3,92} = 0.83$, $P = 0.48$).

Cross-species Amplifications

Among 86 markers tested on congeneric species, 83 (96.5%) are likely to be useful for *C. angulata*, 71 (82.6%) for *C. sikamea*, 31 (36.0%) for *C. ariakensis*, and only 11 (12.8%) for *C. virginica*. These qualitative observations of amplification are supported by a quantitative analysis of PCR yields. Variance in amount of PCR product, as measured by peak optical density, is corrected for differences in initial concentration of template DNA (adjusted OD). Variance in *adjOD* is analyzed by 2-way, randomized, complete-block linear models with fixed effects. Both models are highly significant, the first accounting for 41.2% of variance in adjusted OD and the second accounting for 41.3% of variance. In both models, the block effect (different observers working with different DNA extractions) is nonsignificant. Species, on the other hand, has a highly significant effect on adjusted OD in both analyses ($F_{4,14} = 81.71$, $P < 0.0001$, for the repeat-motif analysis, and $F_{4,19} = 36.13$, $P < 0.0001$, for the motif-complexity analysis). Repeat-motif has only a mildly significant effect on adjusted OD ($F_{2,14} = 3.99$, $P < 0.043$), motif-complexity has no effect on adjusted OD, and species-by-motif interaction terms are nonsignificant. Mean adjusted OD declines from 12.6 in the focal species, *C. gigas*, and 12.4 in its nearest relative, *C. angulata*, which are not significantly different, to 10.9, 8.1, and 6.7 in the nonfocal species *C. sikamea*, *C. ariakensis*, and *C. virginica*, which are significantly different from *C. gigas* and *C. angulata* and from each other (Fig. 2).

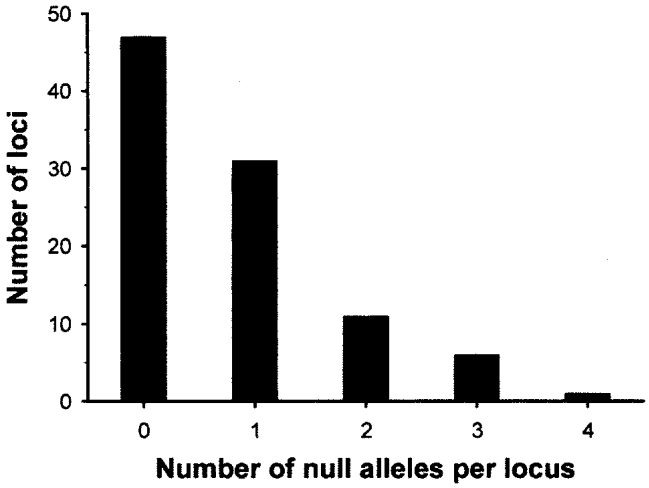

Figure 1. Distribution of 96 microsatellite markers, according to the number of potentially independent null alleles in 4 lineages of Pacific oysters. Forty-nine markers (51%) have at least 1 null allele, and the total number of potentially independent null alleles is 75.

Regression of adjusted OD on species and null-allele count per locus (*#null*) in *C. gigas* is significant for both variables. Species explains 35% of variance in the adjusted OD, whereas null-allele count per locus, though significant, accounts for only 0.3% of variance ($adjOD = 14.772 - [1.481 \cdot species] - [0.210 \cdot \#null]$; $F_{1,3012} = 1647.6$ for *species*, $P < 0.0001$ and $F_{1,3012} = 13.04$, $P < 0.0003$ for *#null*).

Variance in allelic diversity is analyzed with 2-way linear models, in which the main effects are species (six, including an F_3 hybrid group of *C. gigas*) and either repeat-motif or motif-complexity categories. The dependent variable, n_a, is standardized by the covariate n_i, the number of individuals amplified. Both models are significant, accounting for 43.8% of variance in the analysis of species and repeat motifs and 41.6% in the analysis of species and motif-complexity categories. In both models, the covariate is highly significant ($F_{1,394} = 61.4$, $P < 0.0001$, and $F_{1,388} = 43.74$, $P < 0.0001$, respectively) and positively related to allelic diversity (slopes of 0.43 and 0.37, respectively), as expected. Spe-

TABLE 2.

Observed heterozygosity, numbers of alleles, and numbers of independent null alleles, by (A) repeat motif and (B) motif-complexity categories, for 96 microsatellite markers in *C. gigas*.

A. Repeat number

Motif (number of loci)	H_{obs} ± SE	n_a ± SE	#*null* ± SE
Di-nucleotide (58)	0.769 ± 0.028	5.93 ± 0.26	0.948 ± 0.128
Tri-nucleotide (22)	0.704 ± 0.046	5.27 ± 0.42	0.545 ± 0.207
Tetra-nucleotide (16)	0.729 ± 0.054	5.88 ± 0.50	0.688 ± 0.243
Overall (96)	0.748 ± 0.022	5.77 ± 0.20	0.812 ± 0.100
B. Motif complexity			
Simple, pure (41)	0.742 ± 0.033	5.54 ± 0.31	0.683 ± 0.153
Compound, pure (11)	0.758 ± 0.064	6.54 ± 0.59	1.182 ± 0.295
Simple, interrupted (23)	0.689 ± 0.045	5.30 ± 0.41	0.783 ± 0.204
Compound, interrupted (21)	0.817 ± 0.047	6.33 ± 0.43	0.905 ± 0.214
Overall (96)	0.748 ± 0.022	5.77 ± 0.20	0.812 ± 0.100

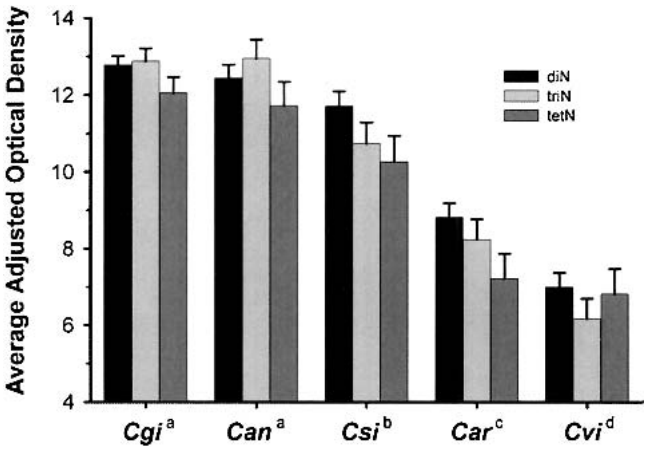

Figure 2. Average adjusted optical density (least-squares means and standard errors) of amplified microsatellite DNA markers with di-, tri-, and tetra-nucleotide repeat-motifs, among 5 species of *Crassostrea*, *C. gigas* (*Cgi*, $n = 1137$), *C. angulata* (*Can*, $n = 516$), *C. sikamea* (*Csi*, $n = 430$), *C. ariakensis* (*Car*, $n = 484$), and *C. virginica* (*Cvi*, $n = 484$). Superscripts on species names indicate significant differences; bars give the upper 95% confidence limits on least-squares means. diN, di-nucleotide repeat; triN, trinucleotide repeat; tetN, tetra-nucleotide repeat.

cies is also highly significant in both models ($F_{5, 394} = 16.08$, $P < 0.0001$, and $F_{5, 388} = 17.43$, $P < 0.0001$, respectively). Repeat-motif is significant ($F_{2, 394} = 4.21$, $P = 0.016$), whereas motif complexity is barely significant ($F_{3, 388} = 2.78$, $P = 0.041$). The interaction terms are not significant in these analyses. Mean number of alleles, adjusted to a common sample size of 5 individuals, declines from 4.54 and 4.58 in the *C. gigas* and *C. angulata* (not significantly different), to 3.42, 2.64, and 2.09 in *C. sikamea*, *C. ariakensis*, and *C. virginica*, respectively (Fig. 3). The 7×6 F_3 hybrid *C. gigas* have 4.27 alleles on average, fewer but not significantly fewer than noninbred *C. gigas* or *C. angulata*. Mean n_a for trinucleotides, 3.13, is significantly less than the mean

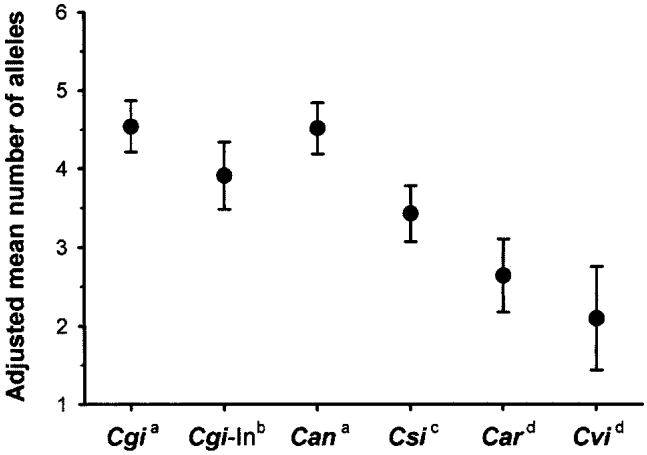

Figure 3. Mean number of alleles, standardized to 5 individuals, for microsatellite DNA markers, amplified from 5 species of *Crassostrea* and an F_3 hybrid group of *C. gigas* (*Cgi-In*): *C. gigas* (*Cgi*, $n = 85$; *Cgi-In*, $n = 52$), *C. angulata* (*Can*, $n = 85$), *C. sikamea* (*Csi*, $n = 78$), *C. ariakensis* (*Car*, $n = 50$), and *C. virginica* (*Cvi*, $n = 30$). Superscripts on species names show significant differences. Bars show 95% confidence intervals on least-squares means.

for tetra-nucleotide repeats (3.85) or di-nucleotide repeats (3.51), respectively.

DISCUSSION

Polymorphism in C. gigas

The diversity of repeat motif and motif-complexity of markers used in this study should help reduce bias in polymorphism or distribution across the oyster genome (Tóth et al. 2000, Katti et al. 2001). Interruptive mutations, which occur in almost half of the microsatellite loci in *C. gigas* (see Table 2), are hypothesized to be the first step in the "death" of microsatellites because they prevent slipped-strand mispairing and stabilize the repeat (Taylor et al. 1999). This scenario is not supported here, because neither heterozygosity nor number of alleles per locus varies among repeat-motif or motif-complexity categories. However, the power of this test is uncertain, because we measure heterozygosity and allele diversity in only four pairs of sibling F_1 hybrids. Nevertheless, we do compare a large number of loci in each repeat-motif category. Ideally, one would like to base comparisons of polymorphism among microsatellite motifs on large samples from natural populations, but the widespread occurrence of null alleles confounds estimation of heterozygosity and allelic diversity in natural populations.

Null Alleles at Microsatellite DNA Markers

Null alleles at allozyme loci have been detected in experimental crosses of several bivalve species (see references in Gaffney 1994). Gaffney (1994) reported null alleles at 10 of 11 allozyme loci in the coot clam *Mulinia lateralis*, with an average frequency per locus of 0.04. Since the advent of PCR, nonamplifying null alleles at DNA markers have been reported in the American oyster, the Pacific oyster, and the geoduck clam (Hu & Foltz 1996, Launey & Hedgecock 1999, 2001, McGoldrick et al. 2000, Vadopalas & Bentzen. 2000, Reece et al. 2001). Among 47 family tests of Mendelian transmission of microsatellites in Pacific oyster families, 17% of parental alleles were nonamplifying (McGoldrick et al. 2000). Similarly, of 94 segregation ratios tested in seven families by Launey & Hedgecock (2001), 15 showed unexpected progeny phenotypes that were best explained by null alleles. More recently, Sekino et al. (2003) report population data for nine microsatellite DNA markers in the Pacific oyster; of eight loci that are polymorphic, four show a deficiency of heterozygotes that is significant at the nominal 5% level. The mean frequency of null alleles at these four loci is 0.111.

Of 96 microsatellite DNA markers assayed in this study, 49 (51%) have at least one nonamplifying null allele in the three families examined (see Fig. 1). This is especially surprising, because these families were derived from the same natural population from which the microsatellite libraries were originally cloned. The average frequency of null alleles for oyster microsatellite loci, 0.093, is twice as high as the frequency of allozyme null alleles. If we make the simple assumption that each PCR-null allele is caused by a single nucleotide polymorphism (SNP) in a primer-binding site, then, given 3997 nucleotide base pairs in the PCR primers for 96 loci, we infer a SNP density of $49/3997 = 0.0123$, one SNP every 82 base pairs. This is likely to be an underestimate, because not all nucleotides in primer binding sites are essential to primer binding. An estimate based on the assumption that the 75 null alleles in independent lineages are different, 0.0188, is signifi-

cantly greater, but this may be an overestimate if many of the nulls segregating among inbred lines are allelic. The frequency of non-amplifying null alleles at microsatellite DNA loci in this oyster population suggests high nucleotide polymorphism in *C. gigas*.

In population studies, PCR-null alleles produce an excess of homozygotes relative to Hardy-Weinberg (random mating) equilibrium genotypic proportions. Although PCR-null alleles may be visualized by redesigning primers to more conservative parts of the flanking region (Jones et al. 1998), primer redesign does not guarantee that all PCR-null alleles will be eliminated from all populations of potential interest. It will be necessary to exercise caution in applying microsatellite markers to studies of natural bivalve populations.

Decline in Cross-species Amplification of Microsatellite Markers

The congeneric species used in this study are thought to have diverged from the Pacific oyster from <1 to >5 million years ago (Mya). Divergence times, however, must be inferred largely from molecular differences, without benefit of a fossil record or other means for calibrating rates of molecular evolution.

The Portuguese oyster *C. angulata* is the closest relative of the focal species *C. gigas*. Indeed, the Portuguese oyster is believed to be synonymous with the Pacific oyster, based on similarity of allozymes, indistinguishable larval and adult shells, complete cross-fertility, and normal meiosis in hybrids (see Boudry et al. 1998, Ó Foighil et al. 1998, and references therein). Mitochondrial DNA analysis implicates Taiwan as the source of introduction of this oyster to Portugal, likely by 16th century Portuguese traders (Boudry et al. 1998). An average 2.3% difference in mitochondrial cytochrome oxidase I (COI) nucleotide sequence suggests, however, that the Taiwanese and Japanese populations of this species may have diverged several hundred thousand years ago (Ó Foighil et al. 1998). The next closest relative of *C. gigas* is the Kumamoto oyster *C. sikamea*, which diverged perhaps 1.4–1.8 Mya, based on mitochondrial 16S rDNA sequences and allozymes (Banks et al. 1994). This divergence may also be estimated as ~2.3 Mya, based on the 9.3% average nucleotide difference for the COI sequences determined by Ó Foighil et al. (1998) and assuming a 2% per million years rate of evolution. In the same manner, an average of 14.5% nucleotide difference between the COI sequences of *C. gigas* and *C. ariakensis* suggests a divergence of ~3.6 Mya, whereas the separation of the American oyster *C. virginica* from the Asian *Crassostrea* clade (*C. gigas*, *C. ariakensis,* and *C. belcheri*) might have occurred more than 5 Mya. The latter seems to be an underestimate based on Littlewood's (1994) finding that sequences for 28S rDNA suggest that *C. virginica* is a sister group, along with *Saccostrea commercialis*, to the Asian *Crassostrea* clade.

Thus, ability to amplify microsatellite markers developed for the Pacific oyster from the DNA of congeneric *Crassostrea* species declines precipitously over an evolutionary time span of only 5–10 million years (see Fig. 2). Fewer than one out of eight Pacific oyster microsatellite markers are likely to be useful for the American oyster. Of course, we may have underestimated PCR success in nonfocal species because of 3 limitations in the survey. First, only a single PCR reaction and gel was run for each primer set, though on multiple individuals and always with the same *C. gigas* controls. Second, comparisons were performed under the PCR condition optimized for the focal species. Third, homology of PCR products across congeneric species was assumed for fragments in the same size range as those observed in *C. gigas*, without being confirmed by sequencing. Despite these limitations, we observe a clear signal of PCR decay, perhaps owing to the large number of comparisons, 86 markers, and more than 3000 observations.

The slight but highly significant, negative correlation between null-allele frequency within *C. gigas* and PCR product yield from closely related species implies that failure of cross-species amplification results from evolution of flanking sequences. Fixation of nucleotide substitutions and insertions/deletions causes widespread PCR-null alleles within *C. gigas*, and this same process, over evolutionary time scales, results in loss of ability to amplify homologous loci from related species.

Decline in the ability to amplify Pacific oyster microsatellites from congeneric species is paralleled by a decline in allelic diversity for those markers that do amplify (Fig. 3). The validity of this trend is supported by the observation that an F_3 hybrid family of Pacific oysters has fewer alleles than the various F_1 hybrid or first generation inbred Pacific oysters or the unrelated individuals of *C. angulata* surveyed. The decline in allelic diversity across the congeneric species surveyed is consistent with ascertainment bias, the expectation that, owing to a relationship between allele size and polymorphism, loci selected to be polymorphic in a focal species are likely to be less polymorphic in other species (Ellegren et al. 1995; Goldstein & Pollock 1997). For the limited data in Table 1 of Li et al. (2003), allele size seems to be uncorrelated with either heterozygosity or allele-diversity, leaving the cause of ascertainment bias uncertain.

Homologous microsatellite loci have been successfully amplified with the same PCR primers from very divergent animal taxa—humans and chimpanzees (4–6 Mya), cow and goat (14–17 Mya), cetaceans (35–40 *Mya*), marine and freshwater turtles (300 Mya), and fish (470 *Mya*), (Schlötterer et al. 1991; Garza et al. 1995; Pépin et al. 1995; FitzSimmons et al. 1995; Rico et al. 1996). These studies suggest that microsatellite flanking regions are well conserved in vertebrate taxa, at least in selected cases.

Studies of cross-specific amplification of microsatellites in *Drosophila* provide invertebrate comparisons. Again, many studies have selected markers that could be amplified in related *Drosophila* species to reconstruct phylogenies (e.g., Noor et al. 2001, Noor M, personal communication). Two studies, however, seem to have tested cross-specific amplifications for a large, randomly selected set of microsatellite loci. Colson et al. (1999) report that 86 (80.4%) of 107 microsatellite loci characterized from *D. melanogaster* can be amplified from *D. simulans* and *D. sechellia*. Based on the nucleotide divergence of *Adh* sequences and a rate of evolution calibrated for Hawaiian species, which are endemic to islands of known age, the latter two species are estimated to have diverged from the focal species about 2.3 ± 0.65 Mya (Russo et al. 1995). These authors point out that this estimate has large standard errors and may not be inconsistent with previous estimates of 2–5 Mya. The results of Colson et al. (1999) are thus comparable with the 82.6% amplification success of *C. gigas* microsatellites from *C. sikamea*, which are believed to have diverged ~2 Mya. However, in a study of the *Drosophila virilis* group, Huttunen & Schlötterer (2002) report that, of 42 microsatellite markers developed from *Drosophila virilis*, 34 (81%) and 32 (76%) can be amplified from *D. montana* and *D. flavomontana*, respectively. Based, again, on divergence of *Adh* sequences and the Hawaiian calibration of evolutionary rate, Nurminsky et al. (1996) estimate that the *virilis* and *montana* clades diverged 9.0 ± 0.07 Mya. The success of cross-specific amplification in the *virilis* group is

clearly much higher than what we observe for the *C. gigas versus C. ariakensis* (36.0%) or *C. gigas versus C. virginica* (12.8%) comparisons. Compared with vertebrates or *Drosophila*, then, cross-specific amplification of Pacific oyster, microsatellite DNA markers show an unusually precipitous decline across a series of closely to progressively more distantly related species.

Nucleotide Diversity and High Fecundity

The evidence for high nucleotide polymorphism in the Pacific oyster is interesting in light of recent experimental evidence for a high mutational load in this species (Launey & Hedgecock 2001, Bucklin 2002). High mutational load, nucleotide diversity, and rapid sequence evolution may all be by-products of sexual reproduction in highly fecund organisms having very high rates of mortality in early life stages (Williams 1975). Female Pacific oysters routinely spawn tens of millions of eggs per year, with an upper limit of about 100 million per year (Galtsoff 1964). Species with very high fecundity likely generate more mutations than low fecundity species, owing to the large number of cell divisions required to produce millions or billions of gametes (cf. similar to the argument for male-driven evolution in humans, Li et al. 2002). An association of high mutation rate with high fecundity was previously proposed to explain the distribution of mitochondrial DNA variant haplotypes in the Pacific oyster (Beckenbach 1994). An extremely variable position of *C. virgincia* in a bivalve phylogeny based on 18S rDNA was also attributed to a high substitution rate of this species (Steiner & Muller 1996). Whether the mutation rate per-cell-division, as well as the number of cell divisions, is higher in oysters than in low-fecundity species is so far unknown. The high rate of somatic cell aneuploidy in bivalve molluscs, which correlates with reduced individual growth and probably fitness in the Pacific oyster (Leitão et al. 2001), suggests that important cellular and molecular mechanisms may be more labile in these animals than in vertebrates.

CONCLUSION

The evidence obtained from developing and mapping microsatellite DNA markers suggests that nucleotide sequences in the Pacific oyster are mutating and diverging rapidly. This nucleotide diversity has a dramatic impact on the population genetics and evolution of microsatellite markers. Nucleotide substitutions and insertions/deletions in microsatellite flanking regions produce a high frequency of null-alleles, even within the source population from which microsatellites were cloned, and contribute to failure in amplification of microsatellite markers from closely related, congeneric species. As a result, microsatellite loci should probably be identified *de novo* for each new species and validated with progeny testing before use in population analyses. Homology of microsatellite loci among related species should be confirmed by sequencing before use in phylogenetic analysis.

ACKNOWLEDGMENTS

We dedicate this manuscript to the memory of our friend and colleague, Will Borgeson, who reared and cared for oyster stocks used in this study. The authors thank Drs. P. Boudry, IFREMER, La Tremblade, France, and S. K. Allen, Jr. and K. Reece, Virginia Institute of Marine Science, for providing oyster samples. This research was supported by USDA National Research Initiative Competitive Grants Program (agreement number 99-35205-8260).

LITERATURE CITED

Banks, M. A., D. J. McGoldrick, W. Borgeson & D. Hedgecock. 1994. Gametic incompatibility and genetic divergence of Pacific and Kumamoto oysters, *Crassostrea gigas* and *C. sikamea. Mar. Biol.* 121:127–135.

Beckenbach, A. T. 1994. Mitochondrial haplotype frequencies in oyster: neutral alternatives to selection models. In: B. Golding, editor. Non-neutral evolution: theories and molecular data. New York: Chapman and Hall. pp. 188–198.

Bierne, N., S. Launey, Y. Naciri-Craven & F. Bonhomme. 1998. Early effect of inbreeding as revealed by microsatellite analyses of *Ostrea edulis* larvae.

Boudry, P., S. Heurtebise, B. Collet, F. Cornette & A. Gérard. 1998. Differentiation between populations of the Portuguese oyster, *Crassostrea angulata* (Lamark) and the Pacific oyster, *Crassostrea gigas* (Thunberg), revealed by mtDNA RFLP analysis. *J. Exp. Mar. Biol. Ecol.* 226:279–291.

Bucklin, K. A. 2002. Analysis of the genetic basis of inbreeding depression in the Pacific oyster *Crassostrea gigas*. Ph.D. dissertation, University of California, Davis.

Chambers, G. K. & E. S. MacAvoy. 2000. Microsatellite: consensus and controversy. *Comp. Biochem. Physiol. Part B* 126:455–476.

Colson, I., S. J. MacDonald & D. B. Goldstein. 1999. Microsatellite markers for interspecific mapping of *Drosophila simulans* and *D. sechellia. Mol. Ecol.* 8:1951–1955.

Ellegren, H., C. R. Primmer & B. C. Sheldon. 1995. Microsatellite evolution: directionality or bias in locus selection. *Nature Genetics* 11:360–362.

FitzSimmons, N. N., C. Moritz & S. S. Moore. 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Mol. Biol. Evol.* 12:432–440.

Gaffney, P. M. 1994. Heterosis and heterozygote deficiencies in marine bivalves: more light? In: A. R. Beaumont, editor. Genetics and evolution of aquatic organisms. London: Chapman and Hall. pp. 146–153.

Galtsoff, P. S. 1964. The American oyster. *Fish. Bull. U.S. Fish. Wildlife Ser.* 64:1–480.

Garza, J. C., M. Slatkin & N. B. Freimer. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.* 12:594–603.

Goldstein, D. B. & D. D. Pollock. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *J. Hered.* 88:335–342.

Goldstein, D. B. & C. Schlötterer. editors. 1999. Microsatellites: evolution and applications. Oxford University Press, Oxford.

Hu, Y.-P. & D. W. Foltz. 1996. Genetics of scnDNA polymorphisms in juvenile oysters (*Crassostrea virginica*) part I: characterizing the inheritance of polymorphisms in controlled crosses. *Mol. Mar. Biol. Biotechnol.* 5:123–129.

Hubert, S. & D. Hedgecock. 2004. Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas. Genetics*.

Huttunen, S. & C. Schlötterer. 2002. Isolation and characterization of microsatellites in *Drosophila virilis* and their cross species amplification in members of the *D. virilis* group. *Mol. Ecol. Notes* 2:593–597.

Huvet, A., P. Boudry, M. Ohresser, C. Delsert & F. Bonhomme. 2000. Variable microsatellites in the Pacific oyster *Crassostrea gigas* and other cupped oyster species. *Anim. Genet.* 31:71–72.

Jones, A. G., C. A. Stockwell, D. Walker & J. C. Avise. 1998. The molecular basis of a microsatellite null allele from the white sands pupfish. *J. Hered.* 89:339–342.

Katti, M. V., P. K. Ranjekar & V. S. Gupta. 2001. Differential distribution of simple sequence repeats in Eukaryotic genome sequences. *Mol. Biol. Evol.* 18:1161–1167.

Launey, S. & D. Hedgecock. 1999. Genetic load causes segregation distortion in oysters: mapping at 6 hours. Plant & animal genome VII, abstract W14, available at: http://www.intl-pag.org/7/abstracts/ag7031.html.

Launey, S. & D. Hedgecock. 2001. High genetic load in the Pacific oyster *Crassostrea gigas. Genetics* 159:255–265.

Leitão, A., P. Boudry & C. Thiriot-Quievreux. 2001. Negative correlation between aneuploidy and growth in the Pacific oyster, *Crassostrea gigas*: ten years of evidence. *Aquaculture* 193:39–48.

Li, G., S. Hubert, K. Bucklin, V. Ribes & D. Hedgecock. 2003. Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas. Mol. Ecol. Notes* 3:228–232.

Li, W.-H., S. Yi & K. Makova. 2002. Male-driven evolution. *Current Opinion in Genetics & Development* 12:650–656.

Littlewood, D. T. 1994. Molecular phylogenetics of cupped oysters based on partial 28S rDNA gene sequences. *Mol. Phylogen. Evol.* 3:221–229.

Magoulas, A. N., B. Gjetvaj, V. Terzoglou & E. Zouros. 1998. Three polymorphic microsatellites in Japanese oyster *Crassostrea gigas* (Thunberg). *Anim. Genet.* 29:69–70.

McGoldrick, D., D. Hedgecock, L. J. English, P. Baoprasertkul & R. D. Ward. 2000. The transmission of microsatellite alleles in Australian and north American stocks of the Pacific oyster (*Crassostrea gigas*): selection and null alleles. *J. Shellfish Res.* 19:779–788.

Noor, M. A. F., R. M. Kliman & C. A. Machado. 2001. Evolutionary history of microsatellites in the *obscura* group of *Drosophila. Mol. Biol. Evol.* 18:551–556.

Nurminsky, D. I., E. N. Moriyama, E. R. Lozovskaya & D. L. Hartl. 1996. Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplications of the *alcohol dehydrogenase* gene. *Mol. Biol. Evol.* 13:132–149.

Foighil, D., P. M. Gaffney, A. E. Wilbur & T. J. Hilbish. 1998. Mitochondrial cytochrome oxidase I gene sequences support an Asian origin for the Portuguese oyster *Crassostrea angulata. Mar. Biol.* 131:497–503.

Pépin, L., Y. Amigues, A. Lepingle, J. L. Berthier, A. Bensaid & D. Vaiman. 1995. Sequence conservation of microsatellites between *Bos taurus* (cattle), *Capra hircus* (goat) and related species—examples of use in parentage testing and phylogeny analysis. *Heredity* 74:53–61.

Reece, K. S., W. L. Ribeiro, C. L. Morrison & P. M. Gaffney. 2001. *Crassostrea virginica* microsatellite marker development, testing and preliminary linkage analysis. Plant & animal genome IX, abstract W10, available at: http://www.intl-pag.org/pag/9/abstract/W10_07.html.

Rico, C., I. Rico & G. Hewitt. 1996. Four-hundred and seventy million years of conservation of microsatellite loci among fish species. *Proc. R. Soc. Lond. B* 263:549–557.

Russo, C. A., N. Takezaki & M. Nei. 1995. Molecular phylogeny and divergence times of Drosophilid species. *Mol. Biol. Evol.* 12:391–404.

Schlötterer, C., B. Amos & D. Tautz. 1991. Conservation of polymorphic simple sequence loci in cetacean species. *Nature* 354:63–65.

Sekino, M., M. Hamaguchi, F. Aranishi & K. Okoshi. 2003. Development of novel microsatellite DNA markers from the Pacific oyster *Crassostrea gigas. Mar. Biotechnol.* 5:227–233.

Steiner, G. & M. Muller. 1996. What can 18S rDNA do for bivalve phylogeny? *J. Mol. Evol.* 43:58–70.

Taylor, J. S., J. M. H. Durkin & F. Breden. 1999. The death of a microsatellite: a phylogenetic perspective on microsatellite interruptions. *Mol. Biol. Evol.* 16:567–572.

Tóth, F., Z. Gáspári & J. Jurka. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* 10:967–981.

Vadopalas, B. & P. Bentzen. 2000. Isolation and characterization of di- and tetranucleotide microsatellite loci in geoduck clams, *Panopea abrupta. Mol. Ecol.* 9:1435–1436.

Williams, G. C. 1975. Sex and evolution. Princeton: Princeton University Press. pp. 200.

Zwick, M. E., D. J. Cutler & A. Chakravarti. 2000. Patterns of genetic variation in Mendelian and complex traits. *Ann. Rev. Genomics Hum. Genet.* 1:387–407.