Original Article

MARINE BIOTECHNOLOGY

Development and Characterization of EST-SSR Markers in the Eastern Oyster *Crassostrea virginica*

Yongping Wang, Ximing Guo

Haskin Shellfish Research Laboratory, Institute of Marine and Coastal Sciences, Rutgers University, 6959 Miller Avenue, Port Norris, NJ 08349, USA

Received: 24 February 2007 / Accepted: 26 March 2007 / Published online: 9 June 2007

Abstract

Simple sequence repeat (SSR) markers were developed from expressed sequence tags (ESTs) in the eastern oyster (*Crassostrea virginica*). ESTs of the eastern oyster were downloaded from GenBank and screened for SSRs with at least eight units of dinucleotide or five units of tri-, tetra-, penta-, and hexa-nucleotide repeats. The screening of 9101 ESTs identified 127 (1.4%) SSR-containing sequences. Primers were designed for 88 SSR-containing ESTs with good and sufficient flanking sequences. Polymerase chain reaction (PCR) amplification was successful for 71 primer pairs, including 19 (27%) pairs that amplified fragments longer than expected sizes, probably due to introns. Sixty-six pairs that produced fragments shorter than 800 bp were screened for polymorphism in five oysters from three populations via polyacrylamide gels, and 53 of them (80%) were polymorphic. Fifty-three polymorphic SSRs were labeled and genotyped in 30 oysters from three populations via an automated sequencer. Five of the SSRs amplified more than two fragments per oyster, suggesting locus duplication. The remaining 48 SSRs had 2 alleles per individual, including 11 with null alleles. In the 30 oysters analyzed, the SSRs had an average of 9.3 alleles per locus, ranging from 2 to 24. Forty-three loci segregated in a family with 100 progeny, with nine showing significant deviation from Mendelian ratios (three after Bonferroni correction). Seventy percent of the loci were successfully amplified in *C. rhizophorae* and 34% in *C. gigas*. This study demonstrates that ESTs are valuable resources for the development of SSR markers in the eastern oyster, and EST-derived SSRs are more transferable across species than genomic SSRs.

Keywords: *Crassostrea* — expressed sequence tags — linkage mapping — microsatellite — oyster — population genetics — simple sequence repeats

Introduction

Oysters are one of the most important aquaculture species in the United States and throughout the world. With a yield of 4.3 million tons in 2002, oysters are the second largest species group in world aquaculture production after carps and other cyprinids (FAO 2004). Despite the importance and long history of oyster culture, oyster breeding is still at its early stages, especially in the application of molecular breeding technologies (Guo 2004). Genetic improvement of oysters ultimately depends on our understanding of the oyster genome and the availability of necessary tools for genetic improvement. Genetic markers are important tools for aquaculture genetics and breeding. They are essential for stock identification, pedigree analysis, and genomic mapping (Liu and Cordes 2004). Genomic mapping and mapping of economically important quantitative trait loci (QTL) are necessary for molecular-based breeding such as marker-assisted selection. Genomic mapping requires a large number of genetic markers.

Several types of genetic markers are available for genetic and genomic analysis including allozyme, restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and simple sequence repeat (SSR) markers

Correspondence to: Ximing Guo; E-mail: xguo@hsrl.rutgers.edu

 DOI: 10.1007/s10126-007-9011-7 • Volume 9, 500–511 (2007) • © Springer Science + Business Media, LLC 2007

(Karl and Avise 1992; Sekino et al. 2003; Yu and Guo 2003; Liu and Cordes 2004). Allozyme and RFLP markers are useful for population genetics studies, but their usefulness in genomic analysis and breeding is limited because of the small number of loci available and low levels of polymorphism. RAPD markers suffer from low reproducibility and are rarely used for advanced genomic studies. In aquaculture species, AFLPs and SSRs are two of the most popular markers for genomic mapping. In most aquaculture species including oysters, the first genetic maps were constructed with AFLP markers (Kocher et al. 1998; Young et al. 1998; Yu and Guo 2003). While AFLPs are efficient and popular markers for linkage mapping, they are anonymous and dominant markers that are less informative and transferable than SSRs. SSRs or microsatellites are tandem repeats of short sequences of one to six base pairs that are abundant in all eukaryotic genomes. They are codominant, highly polymorphic, and ideal for mapping studies. However, SSRs are expensive to develop and use, and large numbers of SSR markers are needed for most aquaculture species. In the eastern oyster (*Crassostrea virginica*), one of the most important aquaculture molluscs of the United States, only about 28 SSR markers are currently available (Brown et al. 2000; Reece et al. 2004). The number of available SSRs is grossly inadequate for genomic mapping in this species. Although genetic maps have been constructed and disease-resistance genes mapped in the eastern oyster using AFLP markers (Yu and Guo 2003, 2006), codominant SSR markers are needed to make the map and mapping information widely useful. Thus, there is a great need for SSR markers in the eastern oyster.

Traditionally, SSR markers are developed by cloning and sequencing from enriched genomic libraries, a process that is complicated and costly. Recently, large numbers of expressed sequence tags (ESTs) have become available in many aquaculture species, providing a new avenue for the development of SSR markers. ESTs contain SSR sequences in both coding and noncoding regions (Temnykh et al. 2001), and SSRs have been successfully developed from ESTs in many species (Thiel et al. 2003; Gao et al. 2004; Nicot et al. 2004; Serapion et al. 2004; Perez et al. 2005; Wang et al. 2005). EST-derived SSRs are part of or adjacent to functional genes and therefore provide an efficient means of gene mapping.

More than 9000 ESTs are now available for the eastern oyster, and this resource has not been explored for the development of SSR markers. In this study, we tested the feasibility of developing EST-SSRs using the eastern oyster EST database. Here we report the successful development and characterization of 53 EST-SSR markers for the eastern oyster.

Materials and Methods

Identification of SSR-Containing ESTs. ESTs of the eastern oyster were downloaded from the EST database at GenBank (http://www.ncbi.nlm.nih.gov/dbEST). All sequences were downloaded to a text file in FASTA format. The ESTs were screened for SSRs using the software MISA (MicroSAtellite, http://pgrc.ipk-gatersleben.de/misa/). For this study, the criteria for SSRs were set as sequences having at least eight repeats of dinucleotide and five repeats for all other repeats (tri-, tetra-, penta-, and hexanucleotide).

Primer Design. All SSR-containing ESTs were individually inspected for suitability for primer design. SSR-containing ESTs that contain sufficient flanking sequences of good quality (no unknown bases) were selected for primer design. Primers were designed using the software PRIMER 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), with an optimal annealing temperature of 60°C and a fragment size between 100 and 300 bp. A GC clamp was added at the 3′ primer end when possible. Primers were synthesized with Proligo (Sigma-Aldrich, Saint Louis, MO).

Primer Validation. Primer pairs were first validated for polymerase chain reaction (PCR) in five oysters from three populations: two oysters from a wild Delaware Bay (DB) population; two oysters from a hatchery population (NE) that originated in Long Island Sound and had been maintained at the Haskin Shellfish Research Laboratory, Rutgers University, NJ for about ten generations (selected for disease resistance); and one wild oyster from Louisiana (LA). DNA was extracted from adductor muscle or mantle/gill tissue from each individual with the E.Z.N.A. mollusc DNA kit (Omega Bio-tek, GA) using supplied protocols. PCR was carried out in a 10-μl solution containing 1× PCR buffer with 1.0 to 2.5 mM $MgCl_2$, 1.0 mg/ml of bovine serum albumin (BSA), 0.2 mM dNTPs, 0.025 U of *Taq* DNA polymerase, 1 μM of each primer, and 10 to 30 ng of oyster genomic DNA. All primers were amplified using the following PCR profile: an initial denaturing for 2 min at 95°C followed by 30 to 35 cycles of 95°C for 30 s, 55 to 60°C for 1 min, and 72°C for 2 min; and a final extension at 72°C for 10 min. PCR was conducted on either a GeneAmp 9700 thermocycler (Perkin Elmer, Weiterstadt, CA) or a PTC-200 DNA engine (MJ Research Inc., Watertown, MA). PCR products were separated on 2% agarose gels in 0.5× Tris-boric acid-

EDTA (TBE, 1×, 0.09 M Tris, 0.09 M boric acid, 20 mM EDTA at pH 8.0) containing 0.5 μg/ml of ethidium bromide and visualized under ultraviolet (UV) illumination. When necessary, PCR was optimized by adjusting annealing temperatures and $MgCl_2$ concentration.

Determining Polymorphism and Mendelian Inheritance. Once SSR primer pairs were validated as producing stable and specific PCR fragments, polymorphism among the above five oysters was determined using denaturing polyacrylamide gels (4% to 8% polyacrylamide, AA:BIS=19:1, with 7 M urea and in 0.5 × TBE). PCR was performed using the same PCR condition as defined in the preceding text. PCR products were denatured at 95°C for 5 min and immediately cooled on ice before being loaded on preheated polyacrylamide gels and run for 3 to 5 h at 150 V. PCR fragments were visualized using ethidium bromide staining and ultraviolet illumination to determine polymorphism.

Loci that were polymorphic on polyacrylamide gels were genotyped and characterized in 30 oysters from three populations using an automated genetic analyzer. The 30 oysters consisted of 10 from DB, 10 from NE, and 10 from LA populations as described in the preceding text.

Primers for polymorphic loci were labeled with WellRED dyes (Proligo LLC, Sigma-Aldrich) and genotyped using a Beckman CEQ 8000 automated genetic analyzer (Beckman Coulter). PCR was conducted as described in the preceding text, except for reduced concentrations of the labeled primer ranging from 0.05 to 0.1 μM. PCR products were mixed with 30 μl of deionized formamide and 0.4 μl of size standard and loaded on the genetic analyzer in 96-well plates. To reduce cost, several (up to eight) SSRs with different fragment sizes and colors were combined in a single well for genotyping with CEQ 8000. As different dyes varied in signal strength (D4>D3>D2), the amounts of PCR products were adjusted so that all loci would show about the same intensity. Allele size was determined via the genetic analyzer, and genotypes for each oyster were recorded.

Mendelian inheritance of all polymorphic markers was tested in a full-sib family (HB4) with two parents and 100 one-year-old progeny. All segregating loci scored were checked for goodness of fit to the expected Mendelian ratios using the χ^2 test.

Cross-Species Amplification. To assess the transferability of the EST-SSRs obtained from *C. virginica*, all functional SSRs were tested in four other species of *Crassostrea: C. rhizophorae, C. gigas, C. ariakensis,* and *C. sikamea*. Six to seven oysters were used in each species. Whenever possible, individuals from different geographic populations were used to cover possible intraspecific variation. *C. rhizophorae* were the first-generation progeny of a Caribbean population produced and maintained at the Harbor Branch Oceanographic Institute, Florida. *C. gigas* samples consisted of four oysters from a Rutgers (NJ) stock originated from Washington State and three oysters from Oregon. Four of the *C. ariakensis* were from Guangdong in southern China and three from Shandong in northern China. Six *C. sikamea* were from Taylor Shellfish Company (Shelton, WA). *C. ariakensis* from China and *C. sikamea* from Washington were genetically identified using diagnostic markers (Wang and Guo 2007). PCR was conducted using conditions described in the preceding text. Successful cross-species amplification was verified as producing clear and specific bands on 2% agarose gels.

Gene Identity of Characterized EST-SSRs. To obtain putative function of genes represented by the SSRs developed here, GenBank homology search was conducted for all ESTs represented by the characterized SSRs using BLASTX and BLASTN (http://www.ncbi.nlm.nih.gov/BLAST). Significant homology was assumed at an E-value < 2.0E-8.

Results

Type and Frequency of EST-SSRs. A search of the EST database at GenBank found 9101 ESTs for the eastern oyster. The screening of the 9101 ESTs (or 5.2 million bp) using MISA identified 127 (or 1.4%) sequences that contained at least eight di- or five other nucleotide repeats. Eleven ESTs contained more than one SSR, bringing the total number of SSRs to 139 (1.5%) or approximately one in every 37.5 kb. Ten were compound SSRs that had more than one repeat type. Di- and trinucleotide repeats were found at about the same frequency. Among the 139 SSRs, 65 (46.8%) were dinucleotide repeats, 60 (43.2%) were trinucleotide, and only 14 (10.0%) were tetranucleotide repeats (Table 1).

Among dinucleotide repeats, AG/CT was the most frequent motif and accounted for 33.1% of all SSRs or 70.8% of all dinucleotide repeats (Table 1). Among the trinucleotide repeats, ACT/AGT was the most frequent motif, accounting for 16.6% of all SSRs or four times more than the other seven trinucleotide repeats. SSRs consisted of GC-only repeats were not observed.

Primer Design and Validation. Among the 127 SSR-containing sequences, 39 did not qualify for primer design because the flanking sequences were

Table 1. **Number and frequency of different SSR-repeat types identified from 9101 *C. virginica* ESTs**

SSR Motif	*Number*	*Frequency*
AG/CT	46	33.1
AT/AT	14	10.1
AC/GT	5	3.6
ACT/AGT	23	16.6
AAC/GTT	9	6.5
AGG/CCT	8	5.8
ACG/CGT	7	5.0
AAT/ATT	5	3.6
AAG/CTT	3	2.2
ACC/GGT	3	2.2
AGT/ACT	2	1.4
AAAC/GTTT	4	2.8
AGAT/ATCT	3	2.2
AACT/AGTT	2	1.4
ACAG/CTGT	2	1.4
AAAG/CTTT	1	0.7
AATC/GATT	1	0.7
ACAT/ATGT	1	0.7
Total	139	100.0

SSRs are defined as having at least eight di- and five tri- and tetranucleotide repeats.

too short or of poor quality (too many unknown bases). Primers were designed for the remaining 88 SSR primers. After some optimization, 71 of the 88 primer pairs (80.7%) were successfully amplified in five oysters from three populations. The other 17 primer pairs (19.3%) failed to amplify at various annealing temperatures and Mg^{2+} concentrations, and were excluded from further analysis (Figure 1). Among the 71 working primer pairs, 52 (73.2%) produced PCR products at the expected sizes, and 19 (26.8%) produced PCR products that were considerably longer than the expected sizes, probably due to introns. PCR fragments from five primer pairs were longer than 800 bp, a size that is not suitable for automated genotyping, and were excluded from this study. Thus, a total of 66 primer pairs were checked for polymorphism using polyacrylamide gels.

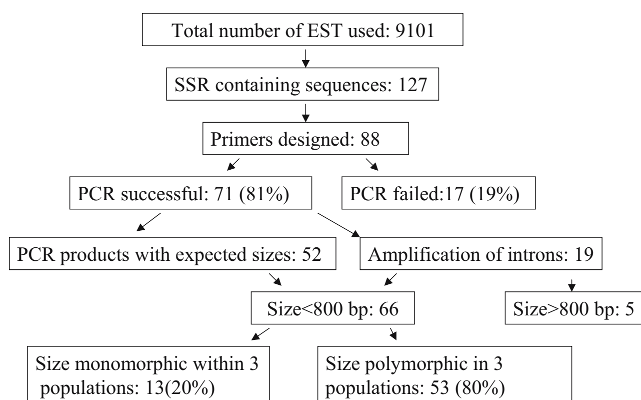

Figure 1. Schematic summary of stepwise results for the development of EST-SSR markers in *Crassostrea virginica*.

SSR Polymorphism. As tested in the five oysters from three populations, 53 of the 66 EST-SSR loci (80.3%) were polymorphic, and the other 13 (19.7%) were monomorphic and might not be true microsatellites. The primer sequences and PCR conditions for the 53 polymorphic loci are presented in Table 2.

The 53 polymorphic loci were further characterized in a panel of 30 oysters using labeled primers and a Beckman CEQ 8000 genetic analyzer. Of the 53 polymorphic SSRs, 48 produced two alleles per individual as expected for single-locus markers. The remaining five SSRs produced more than two alleles per oyster, suggesting that they belonged to duplicated loci.

Polymorphism varied among loci. Of the 48 single-locus SSRs, allele number in the 30 oysters sampled ranged from 2 to 24 (Figure 2), averaging 9.3 alleles per locus. Non-amplifying individuals that could not be explained by poor DNA quality were observed at 11 loci. They were considered homozygous for null alleles. Among the 30 oysters analyzed, the number of non-amplifying individuals ranged from 3 to 21 (or 10% to 70%) across the 11 loci. Notably, one of the loci, RUCV15, was amplified only in the LA population, but not in the two Atlantic populations (DB and NE). At eight SSRs, the peaks of some alleles were considerably lower compared to that of other alleles (Table 3), which was probably due to poor amplification by one of the primers and could be regarded as “mild cases” of null alleles.

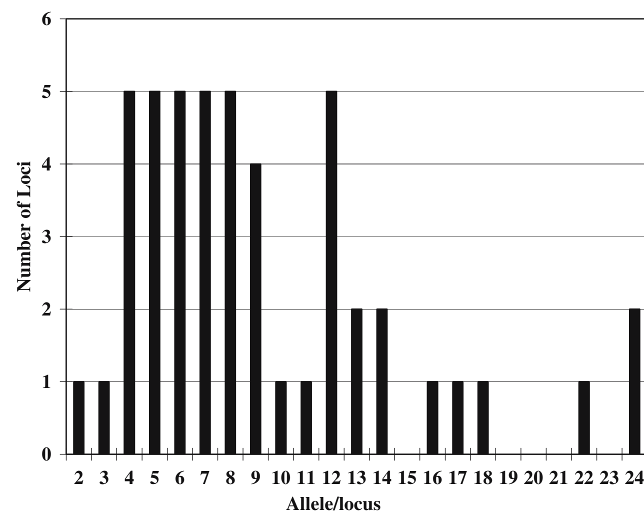

Figure 2. Distribution of allele number per locus at 53 EST-SSR loci in *C. virginica*.

Table 2. Locus name, sequence identity, motif type, primer sequences, PCR conditions, and allele number (in 30 oysters) of 53 EST-SSRs developed in *Crassostrea virginica*

SSR	*GenBank ID*	*Motif*	*Primer sequence (5′–3′)*	*$MgCl_2$ (mM)*	*T_m (°C)*	*Observed size*	*Expected size*	*Allele no.*	*Note*
RUCV1	39726421	(TC)14	AGTCAAGAACTATACAAATTTACGCT CTCACAGACCATGAAAATGGGCTGTT	1.5	60	145–221	180	17	
RUCV3	31908953	(TTGA)8	AGTTATCCATTCTGTTGTGGAAGTGA GTTTGTCCCGACAACATACCGCCATT	1.5	60	268–318	305	12	
RUCV6	31907922	(TC)18	GCATGATACAAGATGGTGAGGTCGAT GATACTTACCTTATATGTAGCTCTGA	1.5	60	165–212	178	13	
RUCV7	31907493	(CT)9	GACCACGCGTCCGCCTTCGCTGTGTT GAATACACCACCAGAATCCTCTGGAA	1.5	55	220–253	231	12	7/30 nulls
RUCV8B	31907084	(CT)11(GT)4tt (GT)3ctctc(GATA)5 (AG)4g(GA)12	TACGACAACGCATGTACG ATCTGCAAGTTCTTCTTTGG	1.5	55	265–326	317	13	
RUCV10	31906358	(GA)32	GAAGTTAATATGGATCCGTGCTTGTA TTATCTTTTGTATAGGGTGAGGGCAA	1.5	60	120–168	160	9	9/30 nulls
RUCV11	31906314	(CT)12y(TC)9	TGCCGGTCGTTCTTTCAGGTATGTTC TTTCTGAAGGGACACTGATAGTGAGT	1.5	60	136–180	160	12	
RUCV12	31906050	(TA)16	TGAGAAATCCCAGAAATCCTGAGAGT CTTTATGTAATTAAAGCTTGCCAGGA	2.5	58	110–155	138	16	>1 locus
RUCV13	31905597	(TGA)5a(GAT)5	CTAAACCTATCCAACAGTCTATGGCC TTTTATTTCTGGTGTTACAGGAAGAT	2.5	58	380–386	218	6	
RUCV15	31905486	(AT)11...(AT)8	TAAGCGCACCTGACCTAACTAGTCCA CCATTTCGCAAGACACTTTCTATGCA	1.5	60	221–243	238	8	10/30 nulls
RUCV18	31904364	(AT)9	TACTTTAATTGCATGCATGTGGTTGT GTCGGTCTGCTTGATCTGTGAAGGTT	1.5	60	133–153	146	7	
RUCV19	31904291	(GA)9	GTTTGTTCTTCAAATCAGTATGAAGT CATTGCACAAGTACATGGTCATGTCA	1.5	60	103–202	112	19	>1 locus
RUCV20	31904207	(CAAT)10	GGCTGAAATATTGCCGATACGGCGTA AGGGGAAAGCAAGAAGTGGTGCAGGA	1.5	60	90–146	116	17	>1 locus
RUCV21	31903865	(TG)9	CCATTATCAACTGGTAGCAGCTGGCA CAAAGCAAGTCTTAGTGGCCCTTCCT	1.5	60	153–167	161	7	
RUCV22	31903758	(AT)47	AAAATTCGCCCTGTTCGTGTTTCATT AAGCGCCTTAGACACTCGTTTGCACA	2.5	58	123–182	188	9	7/30 nulls
RUCV23	31901688	(GA)28g(GA)7	GCAAGATGGGGATGATCAACCTGCAT GGACATCGGATCCCAGTGTCGGTTGA	1.5	55	216–331	253	18	
RUCV24	31901270	(GA)14at(GA)6	AAAAGGGAATTTTGTTACACAATCCA AAAAACAAAATAATGAATACATTGGC	1.5	55	107–173	141	14	
RUCV25	31901222	(AAC)10	GTACAACAGCCTCAGAGCCAATGGCA TCTTAGTTGTGGCGCTGCCGGTTGGT	1.5	55	376–400	132	8	
RUCV26	14581348	(CT)10	TTTTTGAACAAAGGCTTTTGGATTCA GCACAATCATACCAAAGTTTCCGCAT	2.5	58	200–250	194	6	21/30 nulls
RUCV27	14581203	(GA)10	GCTGATCGGGATGGCGAGAGAGTGAC TGAAAACATGCACGTCCGGACAACAT	1.5	60	190–234	197	24	
RUCV28	14581194	(GA)10	GGAGGCCCAAGAACTGCGAGGGGACC TTGAAAACATGCACGTCCGGCAACAT	1.5	55	221–265	225	12	
RUCV29	14581154	(GA)17	TCATAATATTCTTTAAAGCGTCATGA TTGTCTCAAATCCTTTTCATCGGACT	1.5	60	130–191	155	24	
RUCV41	51876887	(GCT)5	CTTGCATCTGTTCCCTCTGG AGAGACCAAATCCCATGTCG	1.5	60	323–340	148	4	
RUCV42	51876864	(CCT)5	CTGCTCCACGCTAAGTCTCC TATTTCGGTGTGGAGGAAGG	1.5	60	527–560	216	6	
RUCV43	51876830	(GAT)6	TCACTCCATTTCTCACACTGG TCCACAAGAAAATGGAAAGG	2.5	60	366–379	230	6	
RUCV44	51876783	(GAA)5	CGAGAGCTTTTAGCGTTTGG CCAAAAGAGACCCCACAGC	1.5	60	211–214	211	2	

RUCV45	51876782	(GA)19	TGTTTAGTCATGGCAGTGTGC GTGACTTCATTTTGAGCCTTTTACC	1.5	60	190–337	200	22	
RUCV46	51876568	(GA)17	GTCGTGCAAGTTGACATTCC TCCACCTCTATTTCATGTTGTCC	1.5	60	100–149	119	9	
RUCV47	51876434	(GCT)5	CTTGCATCTGTTCCCTCTGG AGAGACCAAATCCCATGTCG	1.5	60	323–340	148	4	
RUCV51	51568970	(GAC)5	GAAAAGGCACAGGAGAATGG TCGAACGGTTTTGTTGAAGG	2.5	55	167–187	167	7	
RUCV52	51568628	(GAG)6	AATAAGCCCTGGTCATGTCG CATCATCCCCTTCGTCTTCC	1.5	60	103–115	111	4	
RUCV53	51568519	(TAT)5	ATACGTATCTTTTGTGTTTGATGC TGCAAAACATGCATATTACAACC	1.5	60	234–240	234	5	
RUCV55	51568312	(CCA)5	GCCCATGCCCTACAGATAGC CCATTTCATTGGAACAGTCG	1.5	60	146–194	153	4	
RUCV56	51568204	(CAAA)7	TGAGTGAAAAATTCTCGAGTGG CATTAGCAAATTTGTTCACAGTAACC	1.5	60	116–154	123	9	18/30 nulls
RUCV58	51568032	(AT)10	TTGGAATTTTAATAAATGTCAGAAAGG AACTTAGTTAACATTTTGGAATTGG	2.5	60	243–276	246	8	3/30 nulls
RUCV60	31907658	(GT)9	CAAGTTATGATAAGAGTGACAGG CATACACAGAAACACACATACAG	1.5	58	114–124	119	7	
RUCV61	51567664	(GA)15	CAGCCAACATCACTTTGAGG CTGTGCCGGTACAATCTGC	1.5	60	347–397	343	14	
RUCV62	51567610	(GAT)6	GATTTGCCCTACAGGACTCG CCGATTCCTTATCATCTGTGG	2.5	55	367–376	192	5	6/30 nulls
RUCV63	31905756	(GA)24c(AG)11	TGATACTTTCGTATTGCTTG GATTGTAATTTATTTGAACATT	1.5	55	300–350	346	10	16/30 nulls
RUCV64	51566684	(GA)17	CATATATGGCACTAATATCG CTTAACAAACTCCTCCTTGC	1.5	55	91–145	109	13	>1 locus
RUCV66	31903945	(ACT)6	ACCATCAGCAACACAGAACG GGGTCCCAAGTGTTGTCG	1.5	60	242–340	317	12	
RUCV68	31908118	(CAA)5	TCTTGGAATGACAAGCAAGC CCAGGGGTCAACAGTTTCC	1.5	60	321–342	338	11	
RUCV69	31907800	(GAT)5	GACAAGAGGCCGTAGAATGC ATTGCAGTCGATCGTTTTCC	1.0	60	328–353	219	4	
RUCV73	14581283	(GAT)6	GGGAGCATTATTGCCTAAACC TTCGATAATCACAGAAGGATGG	1.5	60	443–484	331	5	
RUCV74	31906202	(GAT)5	AGAACTGGCACAGTGGTTCC TGGAAACCGCACTTTCTACC	1.0	55	155–167	161	5	
RUCV75	31905842	(GAT)5	CTTATTTGCTGCCCTTGTCC TCTTCCATGTCAGCAAGTCG	1.5	60	300–312	209	5	
RUCV79	31905234	(GAT)5	AGATCTGGCACAGTGGTTCC ACTGCTGTTGTTGCTGTTGG	1.5	60	244–250	247	3	8/30 nulls
RUCV80	31904553	(CTA)7	GTCGGCAACCTCCTACTGC TGACCACCACAATGATGACC	1.5	60	139–151	143	6	
RUCV85	31902108	(AAC)5	TCCATGGCCAGAGAATAAGG CCAATCCAACACAGTTGACG	1.5	60	128–184	176	8	7/30 nulls
RUCV89	14580649	(ACT)5	CGGACAGCTGATGAGTAACC AGTCGGTCATGTGATTGTGC	2.5	60	211–229	221	8	
RUCV97	14581013	(GA)24...(AG)18	AGCCATGATTGAGGAATTGG ATCCCCTAAAGTGCGACTGG	1.5	55	242–323	256	16	
RUCV91	31907287	(GAG)6	AGCAGCACTGACAATACTCTGC TCTGGGCGTAAGATTTGACC	1.5	60	557–654	330	7	
RUCV94	51876253	(ACA)7	CAAGCACTTGTTACATTGAGG GTAATCACCACACACTTTGC	1.5	60	500–750	227	13	>1 locus

Population genetics parameters were not calculated, as the 30 individuals were not from a single population.

Segregation Analysis. All 53 polymorphic SSRs were tested for Mendelian segregation in the HB4 family with 100 progeny (Table 3). Null alleles were deduced at some loci where the unexpected progeny genotypes could be explained only by null alleles in the parents. Null alleles were inferred at 13 of 53 loci (24.5%) in the HB4 family.

Of the 53 SSRs tested, 43 segregated in the family including three duplicated SSRs—RUCV46, RUCV69, and RUCV94—that amplified more than two alleles per oyster. These duplicated loci could be analyzed as only one of the loci was segregating and the other loci were fixed. At RUCV46 (BC × AA), for example, two extra alleles (D at 129 and E at 156 bp) were present in both parents and all the progeny, and could be excluded for segregation analysis. Similarly at RUCV69 (AB × AA), the extra allele (C at 341 bp) was present in both parents and showed up in all the progeny.

As indicated by the χ^2 contingency test, 34 of 43 (79.1%) SSRs segregated in Mendelian ratios, while nine (20.9%) SSRs showed significant segregation distortion ($P<0.05$) (Table 3). After sequential Bonferroni corrections, only three SSR loci showed significant segregation distortion: RUCV8B, RUCV22, and RUCV26.

Cross-Species Amplification. Cross-species amplification varied among species. Of the 53 *C. virginica* EST-SSRs tested, 37 (69.8%) could be amplified in *C. rhizophare*, a closely related species from the Atlantic (Table 4). In comparison, the number of *C. virginica* SSRs that could be amplified in Asian species was considerably lower: 18 (34.0%) in *C. gigas* and 14 (26.4%) in *C. ariakensis* and *C. sikamea*. Cross-species amplification of SSRs is summarized in Table 4.

Gene Function. GenBank (BLASTX) searches indicated that 17 of 53 SSR-bearing ESTs (32.1%) matched to genes of known functions at E values < 2.0E-8, while the other 36 (67.9%) either had no significant matches or matched to unknown genes (Table 5). Among the unknowns, locus RUCV11 matched Cvi2i20, a previously described microsatellite locus of *C. virginica* (Reece et al. 2004). RUCV15 was an (AT)n repeat from the major noncoding region (MNR) of the mitochondrial genome (P.M. Gaffney, personal communication). This is also the locus that amplified in the Gulf population, but not in the Atlantic populations.

The 17 genes included the tetraspanin family gene, agrin precursor, heat shock protein 90, two ribosomal proteins, and several other proteins. Several loci matched to the same EST sequence or different fragments of the same gene. They included RUCV27, 28 and 85 matching the DUF614 family gene; RUCV 43 and 73 matching LOC495205 protein; RUCV41 and 47 matching ribosomal protein P2; and RUCV68, 74 and 79 matching to ribosomal protein 28S. Most of these genes were segregating in HB4 and could potentially be mapped (Table 3).

Discussion

This study demonstrates that ESTs are valuable resources for the development of type I SSR markers in the eastern oyster, as has been shown in other organisms (Thiel et al. 2003; Gao et al. 2004; Nicot et al. 2004; Serapion et al. 2004; Perez et al. 2005; Wang et al. 2005). This is probably the first time that EST-SSRs have been developed in ~~molluscs~~. Using the criteria we defined (at least eight di- and five other nucleotide repeats), we found 1.4% of the ESTs contain SSRs, or about one SSR every 37.5 kb. The rate is lower than what has been reported in other species, probably because of the conservative standard we used. In wheat and rice, for example, SSRs are found at every 9.2 and 3.4 kb, respectively (Cardle et al. 2000; Gupta et al. 2003). Using a similar lower threshold (at least five di-, four tri- and tetranucleotides), we found one SSR every 9.7 kb in *C. virginica*, which is about the same as in wheat and rice. We expect that more SSRs can be developed from the eastern oyster ESTs. Interestingly, the most abundant trinucleotide repeat ACT is also the most productive selection bases for AFLPs in the same species (Yu and Guo 2003).

It should be pointed out that none of the loci has been resequenced to verify the nature of the length polymorphism observed here. It is possible that some of the polymorphisms, especially at less polymorphic loci, are caused by indels rather than true SSRs. The highly polymorphic loci with many alleles (Table 3) are likely true SSRs.

Amplification Success and Intron Frequency. In this study, PCR amplification was successful for 81% of the primer pairs designed from ESTs, a rate that is higher than that observed for genomic SSRs: 67% in the same species (Reece et al. 2004) and 47% in *C. gigas* (Li et al. 2003). The difference may be explained by variations in primer design or by lower levels of sequence variation at the priming sites for EST-SSRs.

Among primer pairs that amplified, 19, or 27%, amplified introns, which is similar to the 22%

Table 3. Segregation analysis of 53 C. virginica EST-SSRs in a full-sib family

Locus	*F*	*M*	*N*	*Progeny genotype*	*Expected ratio*	*Observed*	*P value*
RUCV1	BC	AB	98	AB:BB:AC:BC	1:1:1:1	14:22:39:23	0.0038*
RUCV3	AO	BO	100	AB:AO:BO:OO	1:1:1:1	36:27:20:17	0.0358*
RUCV6	BO	AB	98	AB:BB+BO:AO	1:2:1	28:55:15	0.0855
RUCV7	AB	AB	100	AA:AB:BB	1:2:1	27:47:26	0.8270
RUCV8B	**A**B	BB	92	AB:BB	1:1	24:66	0.0001*
RUCV10	AO	OO	97	AO:OO	1:1	44:53	0.4958
RUCV11	AC	BD	97	AB:AD:BC:CD	1:1:1:1	22:25:26:24	0.9482
RUCV12	BO	AO	99	AB:AO:BO:OO	1:1:1:1	24:25:17:33	0.1576
RUCV13	BD	AC	97	AB:CD:AD:BC	1:1:1:1	27:26:27:17	0.4045
RUCV15	BB	AA	100	AB	1	100	—
RUCV18	AB	AA	99	AB:AA	1:1	59:40	0.0562
RUCV19	AD	BC	94	AB:AC:BD:CD	1:1:1:1	20:16:25:330	0.0768
RUCV20	AC	AB	95	AA:AB:AC:BC	1:1:1:1	28:18:25:24	0.6428
RUCV21	BC	AA	97	AB:AC	1:1	42:55	0.1869
RUCV22	AO	BO	100	AB:AO:BO:OO	1:1:1:1	2:5:46:47	0.0001*
RUCV23	AB	AC	99	AA:AC:AB:BC	1:1:1:1	24:26:22:27	0.8964
RUCV24	AC	BC	100	AB:AC:BC:CC	1:1:1:1	18:20:20:42	0.0014*
RUCV25	AA	BB	100	AB	1	100	—
RUCV26	BO	AO	100	AB:AO:BO:OO	1:1:1:1	10:24:23:43	0.0006*
RUCV27	AB	CD	96	AC:AD:BC:BD	1:1:1:1	30:20:22:24	0.5061
RUCV28	AB	CD	94	AC:AD:BC:BD	1:1:1:1	26:22:24:22	0.9258
RUCV29	AB	CD	97	AC:AD:BC:BD	1:1:1:1	21:20:30:26	0.4453
RUCV41	BB	AA	99	AB	1	99	—
RUCV42	AA	AA	96	AA	1	96	—
RUCV43	AA	BC	95	AB:AC	1:1	42:53	0.2591
RUCV44	CC	AB	96	AC:BC	1:1	41:55	0.1530
RUCV45	BD	AC	96	AB:CD:AD:BC	1:1:1:1	25:25:24:22	0.9691
RUCV46	BC	AA	100	AB:AC	1:1	45:55	0.3173
RUCV47	BB	AA	100	AB	1	100	—
RUCV51	AO	BB	100	AB:BO	1:1	50:50	1.0000
RUCV52	AB	AB	95	AA:AB:BB	1:2:1	21:42:32	0.1480
RUCV53	AA	AA	97	AA	1	97	—
RUCV55	AA	AA	97	AA	1	97	—
RUCV56	AO	OO	100	AO:OO	1:1	42:58	0.1096
RUCV58	BO	AC	98	AB:BC:AO:CO	1:1:1:1	18:21:28:31	0.2169
RUCV60	BB	AB	96	AB:BB	1:1	43:53	0.3074
RUCV62	BC	AB	96	AB:BB:AC:BC	1:1:1:1	15:28:19:34	0.0261*
RUCV61	BB	AB	97	AB:BB	1:1	47:50	0.7607
RUCV63	AB	CD	98	AC:AD:BC:BD	1:1:1:1	21:24:29:24	0.7180
RUCV64	AB	BC	98	AB:AC:BB:BC	1:1:1:1	18:26:36:18	0.0301*
RUCV66	BO	AB	100	AB:BB+BO:AO	1:2:1	14:51:35	0.0119*
RUCV68	OO	AA	100	AO	1	100	—
RUCV69	AB	AA	98	AB:AA	1:1	50:48	0.8399
RUCV73	AA	BC	95	AB:AC	1:1	43:52	0.3558
RUCV74	BB	AA	100	AB	1	100	—
RUCV75	BC	AA	99	AB:AC	1:1	59:40	0.0562
RUCV79	BB	AO	99	AB:AO	1:1	52:47	0.6157
RUCV80	AA	AA	100	AA	1	100	—
RUCV85	BC	AC	99	AB:BC:AC:CC	1:1:1:1	31:26:20:22	0.4139
RUCV89	AA	AB	97	AA:AB	1:1	40:57	0.0843
RUCV91	AC	AB	93	AA:AB:BC:AC	1:1:1:1	24:24:21:24	0.9618
RUCV94	AB	AO	92	AO+AA:AB:BO	2:1:1	45:16:31	0.0848
RUCV97	BC	AB	97	AB:BB:AC:BC	1:1:1:1	28:24:24:21	0.7963

*Designates significant deviation from expected Mendelian ratios at $P<0.05$ revealed by χ^2 test. Bold letters in parental genotype designate weak alleles; O represents inferred null alleles.

observed in barley (Thiel et al. 2003). However, the 27% intron frequency should be a conservative estimate, as primers that failed to amplify may be caused by problems in primer design, as well as by the occurrence of large introns. Therefore, the 27% should be viewed as a lower limit, and the upper limit should be around 41% (assuming all non-amplification was caused by large introns). There

Table 4. Cross-species amplification of 53 *C. virginica* EST-SSRs

Locus	*C. rhizophorae*	*C. gigas*	*C. ariakensis*	*C. sikamea*
RUCV1	–	–	–	–
RUCV3	–	–	–	–
RUCV6	–	–	–	–
RUCV7	++	–	–	–
RUCV8B	++	–	–	–
RUCV10	–	–	–	–
RUCV11	–	–	–	–
RUCV12	++	–	–	–
RUCV13	++	–	–	–
RUCV15	–	++	+	+
RUCV18	++	–	–	–
RUCV19	–	–	–	–
RUCV20	+	–	–	–
RUCV21	++	–	–	–
RUCV22	++	–	–	+
RUCV23	++	–	–	–
RUCV24	++	++	++	++
RUCV25	++	–	–	–
RUCV26	–	++	++	++
RUCV27	+	–	–	–
RUCV28	+	+	+	+
RUCV29	–	–	–	–
RUCV41	++	+	+	–
RUCV42	++	–	–	–
RUCV43	++	–	–	–
RUCV44	–	–	–	–
RUCV45	++	–	–	–
RUCV46	+	+	–	–
RUCV47	++	++	++	+
RUCV51	++	+	++	++
RUCV52	+	+	+	–
RUCV53	++	–	–	–
RUCV55	–	–	–	–
RUCV56	++	–	–	–
RUCV58	++	++	–	–
RUCV60	++	–	–	–
RUCV61	++	++	++	++
RUCV62	–	–	–	++
RUCV63	++	++	++	+
RUCV64	–	–	–	–
RUCV66	++	–	–	–
RUCV68	++	++	+	+
RUCV69	++	+	+	–
RUCV73	++	–	–	–
RUCV74	+	++	++	++
RUCV75	++	–	–	–
RUCV79	–	–	–	–
RUCV80	++	++	–	–
RUCV85	–	++	+	++
RUCV89	++	–	–	++
RUCV91	++	–	–	–
RUCV94	++	+	–	–
RUCV97	–	–	–	+
Strong	31	11	7	8
Weak	6	7	7	6
Total	37 (69.8%)	18 (34.0%)	14 (26.4)	14 (26.4%)

++, Strong amplification; +, weak; –, no amplification.

are no comparable data in oysters and marine bivalves. Most of the amplified introns are less than 800 bp, but this observation is biased because large introns may not be amplified.

Polymorphism and Null Alleles. EST-SSRs have been evaluated in several studies and tend to be less polymorphic than those from genomic DNA (Thiel et al. 2003). In this study, the majority (80%) of EST-

Table 5. Functional genes and proteins represented by characterized EST-SSRs in *C. virginica*

Locus	*Sequence ID*	*Gene function*	*E-value*
RUCV7	gi|72009233|ref|XP_783607.1|	Predicted: similar to sterile alpha motif domain containing 8,	2.00E–37
RUCV13	gi|68395729|ref|XP_694278.1|	Predicted: hypothetical protein XP_689186 [*Danio rerio*].	1.00E–38
RUCV19	gi|66734438|gb|AAY53608.1|	Tetraspanin family protein [*Branchiostoma belcheri tsingtaunese*]	4.00E–27
RUCV27	gi|48479185|gb|AAT44867.1|	DUF614 protein [*Branchiostoma belcheri tsingtaunese*]	6.00E–14
RUCV28	gi|48479185|gb|AAT44867.1|	DUF614 protein [*Branchiostoma belcheri tsingtaunese*]	2.00E–08
RUCV41	gi|56199560|gb|AAV84269.1|	Ribosomal protein P2-like [*Culicoides sonorensis*]	1.00E–20
RUCV43	gi|54038523|gb|AAH84597.1|	LOC495205 protein [*Xenopus laevis*]	2.00E–10
RUCV47	gi|56199560|gb|AAV84269.1|	Ribosomal protein P2-like [*Culicoides sonorensis*]	1.00E–20
RUCV55	gi|399021|sp|P25304|AGRN_RAT	Agrin precursor	3.00E–09
RUCV61	gi|15294073|gb|AAK95213.1|	FLJ33790 protein [*Homo sapiens*]	3.00E–08
RUCV62	gi|27803586|gb|AAO21341.1|	Heat shock protein gp96 [*Strongylocentrotus purpuratus*]	1.00E–40
RUCV63	gi|19344028|gb|AAH25712.1|	Odd-skipped related 1 [*Homo sapiens*]	1.00E–38
RUCV68	gi|15294073|gb|AAK95213.1|	40S ribosomal protein S28 [*Ictalur…*	9.00E–19
RUCV73	gi|54038523|gb|AAH84597.1|	LOC495205 protein [*Xenopus laevis*]	2.00E–10
RUCV74	gi|15294073|gb|AAK95213.1|	40S ribosomal protein S28 [*Ictalur..*	1.00E–18
RUCV79	gi|12239511|gb|AAG49498.1|	Ribosomal protein S28 [*Cricetulus griseus*]	8.00E–18
RUCV85	gi|48479185|gb|AAT44867.1|	DUF614 protein [*Branchiostoma belcheri tsingtaunese*]	5.00E–16

SSRs were polymorphic in the three populations studied, which is comparable to the number of polymorphic loci (7 out of 10) observed for genomic-SSRs in the same species (Brown et al. 2000). In the bay scallop *Argopecten irradians*, only 8 out of 21 (38%) EST-SSRs are polymorphic (Roberts et al. 2005).

The number of alleles per locus displayed by the EST-SSRs in this study is high and variable, ranging from 2 to 24 per locus. Such a level of allele diversity is similar to or higher than that observed for genomic SSRs in the same species (Brown et al. 2000; Reece et al. 2004). Such a comparison needs to be viewed with caution, as different populations and sample sizes are used in different studies. Nevertheless, the level of polymorphism observed here is adequate for genetic studies.

Null alleles are non-amplifying alleles, which are caused by polymorphism at the priming sites. Because of exceptionally high levels of DNA sequence variation, null alleles are very common in oysters (Huvet et al. 2000; McGoldrick et al. 2000; Launey et al. 2002; Li et al. 2003; Hedgecock et al. 2004; Reece et al. 2004). In this study, homozygotes for null alleles were found at 11 of the 53 loci, with frequencies ranging from 0.1 to 0.7. Assuming Hardy–Weinberg equilibrium, the null-genotype frequencies translate to null-allele frequencies of 0.32 to 0.84. At some other loci, some alleles were considerably lower in peak height than others. These alleles may be mild cases of null allelism. Cares must be taken during data analysis, so that none of the low-peak alleles are overlooked.

In the HB4 family, 13 of the 53 loci had at least one null allele. In most cases, the null alleles can be accurately deduced and do not affect segregation and mapping analyses (McGoldrick et al. 2000). In some cases, however, the null alleles may make it impossible to separate homozygotes from heterozygotes (e.g., RUCV6; Table 3).

Null alleles are problematic for population studies, as they may complicate genetic analysis of population structure. Null alleles are probably a major cause for the heterozygote deficiency observed from SSR analysis of populations. The high null-allele frequency would limit the use of some of the loci in population studies, although null alleles can be reduced by improvements in primer design (Reece et al. 2004). No null genotypes were observed for the other 37 SSR loci, which may make them useful for population analysis.

Locus Duplication. Among the 53 loci, five loci (CV12, CV19, CV20, CV64, and CV94) produced more than two alleles per oyster, suggesting that they belong to duplicated loci. The duplicated loci may still be used for segregation analysis when only one of the loci is segregating. Multiple loci were also found in other organisms such as wheat (Akhunov et al. 2003; Gupta et al. 2003). In bay scallops, five of the nine EST-SSRs studied amplified multiple bands, suggesting a higher level of locus duplication (Wang et al. 2006). Although preliminary, these observations support the genome-duplication hypothesis that a whole genome-duplication occurred during the evolution of bivalves, and oysters and scallops represent the diploid and tetraploid lineages, respectively (Wang and Guo 2004).

Segregation Distortion. Nine of the 43 loci (21%) showed significant deviation from expected Mendelian

ratios (P<0.05), and only three (7%) remained significant after Bonferroni corrections. The level of segregation distortion observed in this study is about the same as what has been reported for genomic SSRs in the same species, 29% before and 11% after Bonferroni corrections (Reece et al. 2004), but lower than the 36% before and 21% after corrections reported for *C. gigas* (Launey and Hedgecock 2001). A comparable level of segregation distortion (20%) has been reported in *Ostrea edulis* (Naciri et al. 1995). Segregation distortion is common in oysters. It is probably caused by selective mortality of certain genotypes (McGoldrick et al. 2000; Launey and Hedgecock 2001; Reece et al. 2004) and possibly by weak or null alleles. Although they may complicate linkage analysis (Naciri et al. 1995; Launey and Hedgecock 2001), distorted loci can be mapped, and the mapping of distorted markers may help to reveal genes that are important to survival (Yu and Guo 2003, 2006).

Cross-Species Transferability. Because of high sequence divergence, oyster SSRs amplify poorly across species. Only 12.8% of genomic SSRs from *C. gigas* can cross-amplify in *C. virginica* (Hedgecock et al. 2004). In this study, 34% EST-derived SSRs from *C. virginica* amplified in *C. gigas*. This result suggests that EST-SSRs are more (2.7 times) transferable across species than genomic SSR. This is not surprising, as expressed sequences are expected to be more conserved across species than random genomic sequences (Holton et al. 2002). Ten of the *C. virginica* SSRs from this study worked well in three other *Crassostrea* species, suggesting that the flanking regions of these loci are highly conserved. Higher transferability of EST-SSRs has also been reported in other studies (Holton et al. 2002; Gupta et al. 2003; Thiel et al. 2003). The increased cross-species transferability is an added advantage for EST-SSRs. Most of the *C. virginica* EST-SSRs can be amplified in *C. rhizophorae*, providing useful markers for the latter species.

In conclusion, 53 polymorphic SSRs are developed from ESTs in the eastern oyster. This represents a significant increase in the number of SSRs for this species, as currently there are only about 28. Most of the EST-SSRs are highly polymorphic and segregate in Mendelian ratios. They should be valuable in genome mapping and population studies. The EST-SSRs have two major advantages over genomic SSRs. First, as EST-SSRs are part of or adjacent to functional genes, they can be used for the mapping and functional analysis of candidate genes. Second, because ESTs are more conserved than average genomic sequences, EST-SSRs may be more stable and transferable across species. The successful development of EST-SSR in the eastern oyster should encourage similar efforts in other molluscs for which a large number of ESTs are available.

Acknowledgments

We thank Dr. Arnaud Tanguy for assistance with primer design for some of the loci and Dr. Patrick Gaffney for helpful comments. This study was supported by grants from New Jersey Commission on Science and Technology (02-2042-007-11) and NOAA Sea Grant Oyster Disease Research Program (R/OD-2003, NJMSC-6520–0011 and 12) and the Gulf Oyster Industry Program. This is publication no. 2007–50 of IMCS and NJSG–07–667.

References

Akhunov ED, Goodyear AW, Geng S, Qi LL, Echalier B, Gill BS, Miftahudin, Gustafson JP, Lazo G, Chao SM, Anderson OD, Linkiewicz AM, Dubcovsky J, Rota ML, Sorrells ME, Zhang DS, Nguyen HT, Kalavacharla V, Hossain K, Kianian SF, Peng JH, Lapitan NLV, Gonzalez-Hernandeiz JL, Anderson JA, Close DWCJ, Dilbirligi M, Gill KS, Walker-Simmons MK, Steber C, McGuire PE, Qualset CO, Dvorak J (2003) The organization and rate of evolution of the wheat genome are correlated with recombination rates along chromosome arms. Genome Res 13, 753–763

Brown BL, Franklin DE, Gaffney PM, M MH, Dendanto D, Kornfield I (2000) Characterization of microsatellite loci in the eastern oyster, *Crassostrea virginica*. Mol Ecol 9, 2217–2219

Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156, 847–854

FAO (2004) The State of World Fisheries and Aquaculture. (Rome: FAO) p 153

Gao LF, Jing RL, Huo NX, Li Y, Li XP, Zhou RH, Chang XP, Tang JF, Ma ZY, Jia JZ (2004) One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. Theor Appl Genet 108, 1392–1400

Guo X (2004) Oyster breeding and the use of biotechnology. Bull Aquacul Assoc Canada 104–2, 26–33

Gupta PK, Rustgi S, Sharma S, Singh R, Kumar N, Balyan HS (2003) Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Mol Gen Genomics 270, 315–323

Hedgecock D, Li G, Hubert S, Bucklin K, Ribes V (2004) Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. J Shellfish Res 23, 379–385

Holton TA, Christopher JT, McClure L, Harker N, Henry RJ (2002) Identification and mapping of polymorphic SSR markers from expressed gene sequences of barley and wheat. Mol Breed 9, 63–71

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

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. Science 256, 100–102

Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B (1998) A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). Genetics 148, 1225–1232

Launey S, Hedgecock D (2001) High genetic load in the pacific oyster *Crassostrea gigas*. Genetics 159, 255–265

Launey S, Ledu C, Boudry P, Bonhomme F, Naciri-Graven Y (2002) Geographic structure in the flat oyster *Ostrea edulis* L as revealed by microsatellite polymorphism. J Hered 93, 331–338

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

Liu ZJ, Cordes JF (2004) DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238, 1–37

McGoldrick D, Hedgecock D, English LJ, Baoprasertkul P, Ward RD (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

Naciri Y, Vigouroux Y, Dallas J, Desmarais E, Delsert C, Bonhomme F (1995) Identification and inheritance of (GA/TC)n and (AC/GT)n repeats in the European flat oyster *Ostrea edulis* (L). Mol Mar Biol Biotechnol 4, 83–89

Nicot N, Chiquet V, Gandon B, Amilhat L, Legeai F, Leroy P, Bernard M, Sourdille, P (2004) Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor Appl Genet 109, 800–805

Perez F, Ortiz J, Zhinaula M, Gonzabay C, Calderon J, Volckaert FA (2005) Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei*, *Litopenaeus stylirostris*, and *Trachypenaeus birdy*. Mar Biotechnol 7, 554–569

Reece KS, Ribeiro WL, Gaffney PM, Carnegie RB, Allen SK (2004) Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*): confirmation of null alleles and non-Mendelian segregation ratios. J Hered 95, 346–352

Roberts S, Romano C, Gerlach G (2005) Characterization of EST derived SSRs from the bay scallop, *Argopecten irradians*. Mol Ecol Notes 5, 567–568

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

Serapion J, Kucuktas H, Feng J, Liu Z (2004) Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*). Mar Biotechnol 6, 364–377

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellite in rice (*Oryza sativa* L): frequency, length variation, transposon association, and genetic marker potential. Genome Res 11, 1441–1452

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L). Theor Appl Genet 106, 411–422

Wang Y, Guo X (2004) Chromosomal rearrangement in Pectinidae revealed by rRNA loci and implications in bivalve evolution. Biol Bull 207, 247–256

Wang H, Guo X (2007) Identification of *Crassostrea ariakensis* and related oysters by multiplex species-specific PCR. J Shellfish Res (in press)

Wang H, Li F, Xiang J (2005) Polymorphic EST–SSR markers and their mode of inheritance in *Fenneropenaeus chinensis*. Aquaculture 249, 107–114

Wang L, Song L, Xu W, Zhang H (2006) Screening microsatellite markers from EST sequences of bay scallop *Argopecten irradians*. High Technol Lett 12, 97–102

Young WP, Wheeler PA, Coryell VH, Keim P, Thorgaard GH (1998) A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148, 839–850

Yu Z, Guo X (2003) Genetic linkage map of the eastern oyster *Crassostrea virginica* Gmelin. Biol Bull 204, 327–338

Yu Z, Guo X (2006) Identification and mapping of disease-resistance QTLs in the eastern oyster, *Crassostrea virginica* Gmelin. Aquaculture 254, 160–170