JOURNAL OF THE
WORLD AQUACULTURE SOCIETY

doi: 10.1111/jwas.12393

Estimates of Linkage Disequilibrium and Effective Population Size in Wild and Selected Populations of the Pacific Oyster Using Single-nucleotide Polymorphism Markers

Xiaoxiao Zhong

Key Laboratory of Mariculture Ministry of Education, Ocean University of China, Qingdao 266003, China

Qi Li[1]

Key Laboratory of Mariculture Ministry of Education, Ocean University of China, Qingdao 266003, China
Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China

Lingfeng Kong and Hong Yu

Key Laboratory of Mariculture Ministry of Education, Ocean University of China, Qingdao 266003, China

Abstract

The level and extent of linkage disequilibrium (LD) and effective population size (*Ne*) were studied in three selected lines and a wild population of Pacific oyster, *Crassostrea gigas*, using 61 single-nucleotide polymorphisms. Significant differences were detected in the average r^2 between the selected lines and wild population for both syntenic and nonsyntenic loci with LD beyond population-specific critical values ($P<0.05$). Moreover, the proportions of syntenic and nonsyntenic loci with the expected LD level in the wild population were lower than that in the selected lines. Taken together, the LD level of the selected lines was higher than that of the wild population. The extent of LD analysis showed that a short range of LD (0–0.23 cM) was detected in the four populations, and the decay distance was lower in the wild population than in the selected lines. *Ne* values ranged from 47.6 to 58.5 in the selected lines and ranged from 527.9 to 709.6 with infinite upper limits in the wild population. Further variance analysis of LD demonstrated that genetic drift and epistatic selection might account for the increased LD levels in selected lines. The LD information will be valuable for further association study and marker-assisted selection in oysters.

KEYWORDS

***Crassostrea gigas*, linkage disequilibrium, SNP**

The Pacific oyster, *Crassostrea gigas*, is the major cultivated oyster species worldwide. In China, *C. gigas* was originally introduced from Japan in 1979, and now the major production places are Liaoning and Shandong provinces in the north and Fujian and Guangdong provinces in the south (Li et al. 2011). In 2014, China produced 4.35 million tons of oysters, and *C. gigas* accounted for about one third of the total production (DOF 2015). Although there is a long history of oyster aquaculture in China, oyster seeds are produced exclusively in hatcheries and the broodstock remains largely unselected (Li et al. 2006). In 2006, a selective breeding program for *C. gigas* was initiated in Weihai, Shandong Province, to establish selected lines for faster growth by mass selection and obtained an average increase in growth of 10% per generation (Li et al. 2011). Furthermore, several selective breeding projects have been initiated to establish improved breeds of *C. gigas* in other countries (Ward et al. 2000;

[1] Correspondence to: qili66@ouc.edu.cn

Nell 2001; Dégremont et al. 2010; Wang et al. 2012). Quantitative trait locus (QTL) detection and marker-assisted selection (MAS) could be an effective adjunct to traditional selective breeding in oysters. Recently, linkage map and QTL mapping analyses have been extensively developed in *C. gigas* (Li and Guo 2004; Hubert and Hedgecock 2004; Hedgecock et al. 2007; Sauvage et al. 2010; Guo et al. 2012; Zhong et al. 2014; Hedgecock et al. 2015; Wang et al. 2016). However, most QTL mapping analyses have been based on low-density linkage maps using biparental segregating populations with small numbers. This may lead to limited recombination events and make it difficult to detect closely linked markers for MAS (Zhao et al. 2014). An alternative approach to traditional linkage-based QTL mapping is association analysis, also known as linkage disequilibrium (LD) mapping, which studies the relationship between phenotypes and genetic variants to establish marker-trait associations. Its most distinct advantage over traditional QTL mapping is that the association mapping does not require specially designed populations and allows us to acquire higher QTL mapping resolution using historical recombination events in populations (Ersoz et al. 2007; Zhu et al. 2008; Zhao et al. 2014).

LD is a nonrandom association of alleles at different loci, with frequencies higher than that expected from random combination in a population (Kimura 1956). LD may be caused by or reflect epistatic selection, but LD can also be caused by random genetic drift, population mixing, and genetic hitchhiking (Rieger et al. 2012). The feasibility and efficacy of association mapping strongly rely on the extent of LD, which determines the number of markers needed to perform an association analysis and experimental design (Flint-Garcia et al. 2003). Long-range LD indicates that a small number of markers will be required, which is suitable for efficient genome-wide mapping. On the other hand, short-range LD indicates that many more markers will be needed to find associations but that fine-scale mapping of the underlying genes may then be possible (Barnaud et al. 2010). In recent years, a number of investigations have studied the level of LD in humans, domestic animals, and several important aquatic organisms, which may be helpful to design the strategy of association mapping (Farnir et al. 2000; Conrad et al. 2006; Du et al. 2007). By genotyping 49 microsatellites, spanning four chromosomes, in 96 unrelated rainbow trout, *Oncorhynchus mykiss*, Rexroad and Vallejo (2009) showed that LD (squared correlation coefficient r^2) between syntenic loci decayed rapidly at distances greater than 2 cM. By genotyping 307 single-nucleotide polymorphisms (SNPs) and 146 microsatellites in 16 Atlantic salmon, *Salmo salar*, Moen et al. (2008) showed that r^2 dropped to half its maximum value (above baseline) within 15 cM, and r^2 values were higher than 0.2 above baseline for unlinked markers at distances less than 5 cM. By genotyping 53 microsatellites in 88 wild and 96 cultured Pacific oysters, *C. gigas*, Guo et al. (2016) showed that three significant LD ($r^2 > 0.1$) were observed in the cultured populations, that significant LD extended up to 12.7 cM, and that no significant LD was observed for syntenic marker pairs in the wild population. Several studies have reported that the extent of LD was variable in different populations of the same species because many factors could influence the levels and extent of LD, including population size and structure, mating pattern, selection, genetic drift, migration, admixture, mutation, gene conversion, epistasis, and recombination (Ardlie et al. 2002; Li and Merilä 2011). Therefore, it is also crucial to investigate LD information among different populations of *C. gigas*.

The aim of this study was to measure the level and extent of LD and to estimate effective population sizes (*Ne*) in wild and selected populations of *C. gigas* using SNP markers. The LD information will be valuable for devising the strategy of association mapping to identify markers and genes responsible for economic traits in *C. gigas*.

Materials and Methods

Sample Collections and DNA Extraction

Three selected lines and one wild population of *C. gigas* were used in our study. The three selected lines of *C. gigas* were established by

mass selection in Laizhou, Shandong Province, China, in 2007, using three cultivated populations from Korea (Stock K), Japan (Stock J), and China (Stock C) (Li et al. 2011). In June 2007–2013, 30–60 pairs of individuals from the top end of the shell height distribution of each line in every generation were used as parental oysters for the next generation (Jiang et al. 2013). In June 2012, 91 (50 ♀×41 ♂) and 107 (59 ♀×48 ♂) individuals from the fifth-generation selected lines of Japan and Korea were selected as parental oysters for the sixth-generation selected lines of Japan and Korea (JS6 and KS6), respectively. In June 2013, 85 (50 ♀×35 ♂) individuals from the fifth-generation selected lines of China were selected as parental oysters for the sixth-generation selected lines of China (CS6). The parental oysters from each line were sexed and separated. Gametes were rinsed into separate buckets after stripping the gonad. Equal amounts of gametes were mixed well by estimating concentrations using microscopes. Sperm fertilized egg at a ratio of 50:1 with 10^7 oocytes for each female (Li et al. 2011; Jiang et al. 2013). After fertilization, embryos were transferred into a 20-m^3 tank, and 200 million D-larvae were randomly collected from each hatching tank and stocked into 20-m^3 larval-rearing tanks after 24-h incubation. When eyed larvae grew to a size of 320–340 μm on Days 18–22, spat collectors (strings of scallop shells) were placed into the tanks. After 5–7 d, all eyed larvae metamorphosed to spats. Spats attached to the collectors were kept in the concrete tanks until they grew to a size of 500–600 μm. To avoid the admixture of spats from natural populations of oysters, spats from the selected lines were transferred to an outdoor nursery tank (200 m^3) for 1–2 mo, and then the spats from the selected lines were transported to Weihai Bay in Shandong Province, China, mostly after August 15 (Li et al. 2011). The spawning season of wild oysters usually starts from June or July in this area. One-year-old oysters of the three, sixth-generation selected lines (CS6, 48 individuals; JS6, 48 individuals; and KS6, 48 individuals) were randomly selected for this study. The wild *C. gigas* samples were collected from Rushan, Shandong Province, China (RS, 48 individuals). There is an abundant resource of oysters in Rushan, and natural recruitment exists in the site sampled. The spawning season of wild oysters usually starts from June or July in Rushan. Hatchery-propagated oysters are usually kept in the area used for aquaculture from October to April of the next year, and all cultured oysters are collected to breeding farms before April. Therefore, natural populations of oysters might not be heavily impacted by hatchery-propagated populations in this area. The adductor muscle was collected and stored at −80 C. DNA was extracted using the phenol–chloroform method according to Li et al. (2002).

SNP Genotyping

A total of 61 SNPs were selected from the sex-averaged linkage map, which was constructed by Zhong et al. (2014), not including five SNPs (JY129, CgSNP688, CgSNP507, CgSNP625, and CgSNP385) that were monomorphic in the four populations. SNPs were genotyped using the high-resolution melting method on the LightCycler 480 real-time polymerase chain reaction machine (Roche Diagnostics, Burgess Hill, UK) as previously described by Zhong et al. (2013).

Data Analysis

The level and extent of LD were estimated, for samples taken from the sixth selected generation, by calculating the squared correlation (r^2) of allele frequencies between all pairs of markers, defined as:

$$r^2 = \frac{D^2}{\left[p\,(1-p)\,q\,(1-q)\right]}$$

where p is the frequency of allele A at the first locus, q is the frequency of allele B at the second locus, and D = frequency (AB) – pq (Hill and Robertson 1968; Hill 1977). The LD parameter r^2 between all pairs of loci was calculated using Haploview (Barrett et al. 2005). SNPs with minor allele frequency (MAF) ≥ 0.01, $P \geq 0.001$ Hardy-Weinberg equilibrium (HWE), and genotyping rate ≥ 75% (the default setting in Haploview) in a population were selected

for further LD analysis. LD was estimated separately for unlinked loci on different chromosomes and for linked loci on the same chromosome (nonsyntenic r^2 and syntenic r^2, respectively). The comparable statistics of r^2 between selected lines and the wild population were conducted using SPSS (Chicago, IL, USA) 16.0 with Mann–Whitney *U* test (Sokal and Rohlf 1995). The LD decay was examined by fitting a logarithmic trend line to the plot of syntenic r^2 versus genetic distance (cM), using SPSS 16.0. Map positions were based on the sex-averaged linkage map constructed by Zhong et al. (2014) (Table S1, Supporting Information). We set two critical values of r^2 for LD analysis. The population-specific critical value of r^2 was derived from the distribution of the unlinked r^2. Unlinked r^2 estimations should be transformed to approximate a normally distributed random variable, and the parametric 95th percentile of that distribution was regarded as a population-specific critical value of r^2 (Breseghello and Sorrells 2006). Moreover, a strict critical value of r^2 was set at 0.25. The extent of LD was considered to be the distance at which the LD dropped below the threshold of r^2. In addition, LD (r^2) plots were also generated using Haploview.

Ne values were calculated using the LD method in the program LDNE v1.31 (Waples and Do 2008; Waples and Do 2010). The methodology estimates the correlation among alleles at unlinked loci (r^2), which can be related to *Ne* by the formula (Hill 1981):

$$Ne = \frac{1}{3\left(r^2 - \frac{1}{S}\right)}$$

where *S* is sample size. LDNE uses a modified version of this equation with bias correction (Waples 2006). *Ne* was calculated sequentially excluding minor alleles at the 0.01, 0.02, and 0.05 frequency levels, and 95% confidence intervals (CIs) were estimated using jackknife approaches. The *Ne* was also estimated using a temporal method with the program GONE (Coombs et al. 2012). The fifth-generation parents and samples selected from the sixth-generation selected lines were used for the temporal *Ne* analyses. Temporal *Ne* values were calculated using three genetic drift estimators: *Fs* (Jorde and Ryman 2007), *Fc* (Nei and Tajima 1981), and *Fk* (Pollak 1983). Parameters were set as 10,000 iterations of the correction factor, Plan II sampling, two cohorts, no age structure, and 95% CIs.

The molecular variances of LD were analyzed using POPGENE (version 1.32) (Yeh et al. 1999) based on Ohta's method (Ohta 1982a, 1982b). The variances of LD are divided into several components: D^2_{IT} (total variance of disequilibrium), D^2_{IS} (expected variance of LD within a colony), D'^2_{IS} (variance of the correlation of Ai and Bj of one gamete in a colony relative to that of the total population), D^2_{ST} (variance of the correlation of genes of the two loci (Ai and Bj) of different gametes of one colony relative to that of the total population), and D'^2_{ST} (variance of LD of the total population). Ohta (1982a, 1982b) presented this equation:

$$D^2_{IT} = D'^2_{IS} + D'^2_{ST}$$

The analyses were conducted with four data sets (all populations, CS6 and RS, JS6 and RS, and KS6 and RS).

Results

The level of LD was measured for both syntenic and nonsyntenic loci in the four populations (Table 1). Syntenic loci are two genetic loci presumed to be linked to the same chromosome or linkage group; nonsyntenic loci are two loci linked to different chromosomes or linkage groups. The number of syntenic marker pairs ranged from 285 to 319, and the number of nonsyntenic marker pairs ranged from 1093 to 1334 in the four populations. Different numbers of syntenic and nonsyntenic marker pairs were detected in the four populations because different numbers of SNPs were used for LD analyses (Table S2).

The population-specific critical r^2 values ranged from 0.0668 (RS) to 0.0790 (KS6) in the four populations (Table 1). The proportion of the syntenic marker pairs with LD (r^2) beyond

TABLE 1. *Summary of linkage disequilibrium (LD) analysis in wild and selected populations of* Crassostrea gigas *based on the population-specific critical* r^2 *values.*

Population	Marker pairs	Total number of marker pairs	Population-specific critical value of r^2	Number and percentage (%) of marker pairs with LD (r^2) ≥ critical r^2	Average r^2 value (±SD) of marker pairs with LD (r^2) ≥ critical r^2
CS6	Syntenic marker pairs	319	0.0777	38, 11.91	0.169 ± 0.167
	Nonsyntenic marker pairs	1334		96, 7.20	0.119 ± 0.041
RS	Syntenic marker pairs	315	0.0668	25, 7.94	0.144 ± 0.166
	Nonsyntenic marker pairs	1281		84, 6.56	0.107 ± 0.058
JS6	Syntenic marker pairs	301	0.0786	38, 12.62	0.175 ± 0.163
	Nonsyntenic marker pairs	1184		84, 7.09	0.119 ± 0.041
KS6	Syntenic marker pairs	285	0.0790	35, 12.28	0.177 ± 0.164
	Nonsyntenic marker pairs	1093		80, 7.32	0.120 ± 0.046

population-specific critical values ranged from 7.94 (25/315, RS) to 12.62% (38/301, JS6), and the proportion of the nonsyntenic markers varied from 6.56 (84/1281, RS) to 7.32% (80/1093, KS6) using the population-specific critical r^2 as thresholds. Meanwhile, the average r^2 values of syntenic marker pairs with LD (r^2) beyond population-specific critical values ranged from 0.144 (RS) to 0.177 (KS6), and significant differences were observed between the three selected lines and the wild population, respectively ($P < 0.05$). The average r^2 values of nonsyntenic marker pairs with LD (r^2) beyond population-specific critical values varied from 0.107 (RS) to 0.120 (KS6), and significant reductions were detected in the wild population compared with the three selected lines ($P < 0.05$).

The frequency distributions of r^2 values observed for syntenic and nonsyntenic marker pairs in wild and selected populations are substantially similar in the four populations (Fig. 1). More than 40% of the r^2 values are distributed in the 0–0.01 interval, and RS accounts for the highest proportion in this interval. Moreover, the lowest proportion is detected in RS within the 0.06–1.00 interval in the four populations.

Decline of LD with map distance was estimated for all the linkage groups combined. The LD of syntenic loci decreased with increasing map distance both for the three selected lines and the wild population (Fig. 2). The data suggest the levels of LD decay rapidly at distances greater than 0.23 cM in the three selected lines, and that LD extends to a shorter distance of 0.0029 cM in the wild population using the

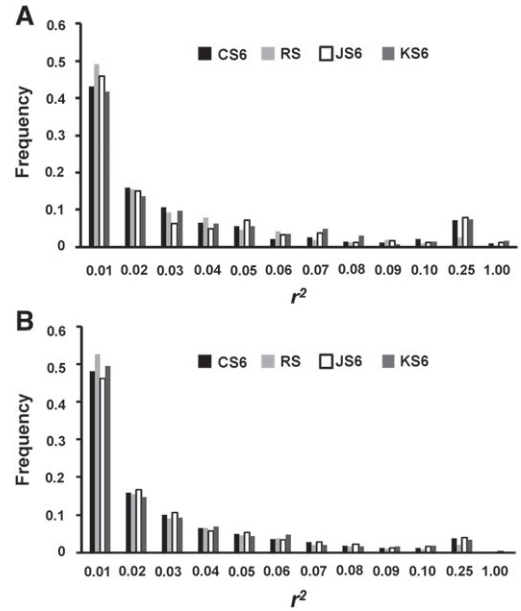

FIGURE 1. *Frequency distribution of* r^2 *values observed for (A) syntenic and (B) nonsyntenic marker pairs in the wild and selected populations of* Crassostrea gigas*. Wild population (RS) and three selected populations (CS6, JS6, and KS6) are shown by different colors.*

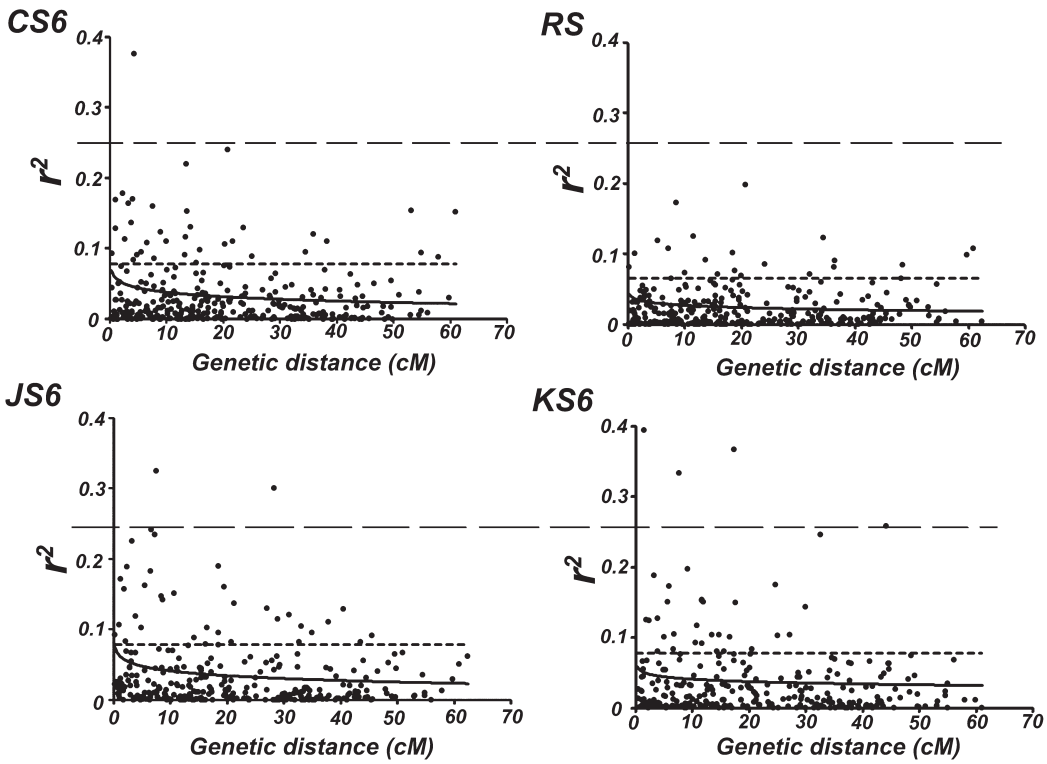

FIGURE 2. *Decline of linkage disequilibrium (r^2) with genetic distance (cM) in the wild and selected populations of* Crassostrea *gigas. Solid black lines show the logarithmic trend line, representing the linkage disequilibrium decay. Black dotted lines indicate the fixed r^2 value of 0.25 (strict critical r^2 value) and the baseline r^2 values (population-specific critical r^2 value) based on the 95th percentile of the distribution of unlinked r^2 values.*

population-specific critical r^2 as thresholds. Moreover, the trend curves were below the strict threshold of $r^2 = 0.25$ in the four populations, also suggesting the rapid decay of LD in *C. gigas*. However, several significant LDs between syntenic loci ($r^2 \geq 0.25$, log of the likelihood odds ratio ≥ 3) were still detected in the four populations (Table 2). Three significant LDs were observed at 1.422, 4.319, and 11.611 cM in the CS6, respectively. In the JS6, four significant LDs extended to 1.422, 7.61, 11.611, and 28.308 cM, respectively. In the KS6, five significant LDs extended to 1.422, 7.61, 11.611, 17.258, and 44.108 cM, respectively. Nevertheless, only two significant LDs were detected at 1.422 and 11.611 cM in the RS, respectively. Most significant LDs between syntenic loci were detected on linkage group 7. Two significant LDs between syntenic loci were observed on linkage group 7 in all four populations (i.e., CgSNP869-CgSNP870 and CgSNP633-CgSNP632 marker pairs) (Fig. 3).

Ne values were estimated for the four populations as shown in Table 3. With the single-sample method, the estimated *Ne* values varied from 47.6 to 55.5 with finite 95% CIs in the three selected lines, whereas the *Ne* values varied from 542.1 to 709.6 with infinite upper limits in the wild population. With the temporal method, the estimated *Ne* values ranged from 49.3 to 58.5 with finite 95% CIs in the three selected lines, and the values ranged from 527.9 to 708.3 with infinite upper limits in the wild population.

The result of variance analysis of LD is shown in Table 4. For the within-subpopulation components, D^2_{IS} ranged from 0.00099 to 0.00111, and D'^2_{IS} ranged from 0.01255 to

TABLE 2. *Significant linkage disequilibrium of syntenic loci in the four populations based on the strict critical* r^2 *value* ($r^2 = 0.25$).

Population	Locus 1 (linkage group [LG])	Locus 2 (LG)	Genetic distance (cM)	r^2	Log of the likelihood odds ratio
CS6	CgSNP484 (LG6)	CgSNP33 (LG6)	4.319	0.376	4.84
	CgSNP869 (LG7)	CgSNP870 (LG7)	11.611	0.783	14.53
	CgSNP633 (LG7)	CgSNP632 (LG7)	1.422	0.871	11.40
RS	CgSNP869 (LG7)	CgSNP870 (LG7)	11.611	0.501	6.70
	CgSNP633 (LG7)	CgSNP632 (LG7)	1.422	0.825	12.65
JS6	CgSNP588 (LG3)	CgSNP686 (LG3)	28.308	0.300	4.31
	CgSNP746 (LG7)	CgSNP451 (LG7)	7.610	0.324	3.33
	CgSNP869 (LG7)	CgSNP870 (LG7)	11.611	1.000	20.63
	CgSNP633 (LG7)	CgSNP632 (LG7)	1.422	0.536	6.28
KS6	CgSNP746 (LG7)	CgSNP451 (LG7)	7.610	0.333	4.42
	CgSNP285 (LG7)	CgSNP870 (LG7)	44.108	0.258	3.54
	CgSNP616 (LG7)	CgSNP432 (LG7)	17.258	0.366	4.79
	CgSNP869 (LG7)	CgSNP870 (LG7)	11.611	1.000	22.12
	CgSNP633 (LG7)	CgSNP632 (LG7)	1.422	0.393	4.10

0.02112. For the between-subpopulation components, D^2_{ST} varied from 0.01215 to 0.02037, and D'^2_{ST} varied from 0.00044 to 0.00074. For the total population component, D^2_{IT} ranged from 0.01326 to 0.02156. Moreover, the data showed the percentages of within-subpopulation components of the total variance ranged from 94.65 to 97.96%, whereas the percentages of between-subpopulation components of the total variance varied from 2.04 to 5.35%.

Discussion

LD, simply defined, is the nonrandom association of alleles at different loci. Previous studies generally considered $r^2 = 0.10$ or 0.25 as the critical value for the extent of LD analysis in aquaculture species (Guo et al. 2016; Rexroad and Vallejo 2009). To describe the relationship between r^2 and genetic distance, two methods of establishing critical r^2 values were investigated. Critical values of r^2 were determined based on a fixed value of 0.25 and from the parametric 95th percentile of the distribution of the unlinked markers (population-specific critical values). The calculation of population-specific critical value for the extent of LD analysis was proposed by Breseghello and Sorrells (2006). The point where the trend line defined by the regression of syntenic LD intersects this baseline (population-specific critical value) defines the extent of LD. The main advantage of this method is that the unlinked LD distribution incorporates the effects of population structure and selection in the experimental samples (Breseghello and Sorrells 2006).

Association mapping of genes or QTLs underlying traits of interest relies on LD between molecular markers and trait loci on same chromosomes, showing physical linkage (Flint-Garcia et al. 2003; Zhao et al. 2014). Our results revealed that the average r^2 and proportion of syntenic marker pairs with LD (r^2) beyond population-specific critical values were higher than these values for nonsyntenic marker pairs, respectively. This demonstrated that physical linkage may play a major role in influencing LD, suggesting that association analysis may be applicable to this species. The higher LD level was observed for syntenic marker pairs because recombination is relatively rare between loci on the same chromosome. However, nonsyntenic marker pairs mean loci located on different chromosomes with different selective pressures and independent segregation, resulting in lower LD level (Flint-Garcia et al. 2003).

In this study, the proportions of syntenic and nonsyntenic marker pairs with LD (r^2) beyond the population-specific critical values in the selected lines were larger than these values in the wild population, respectively. Moreover,

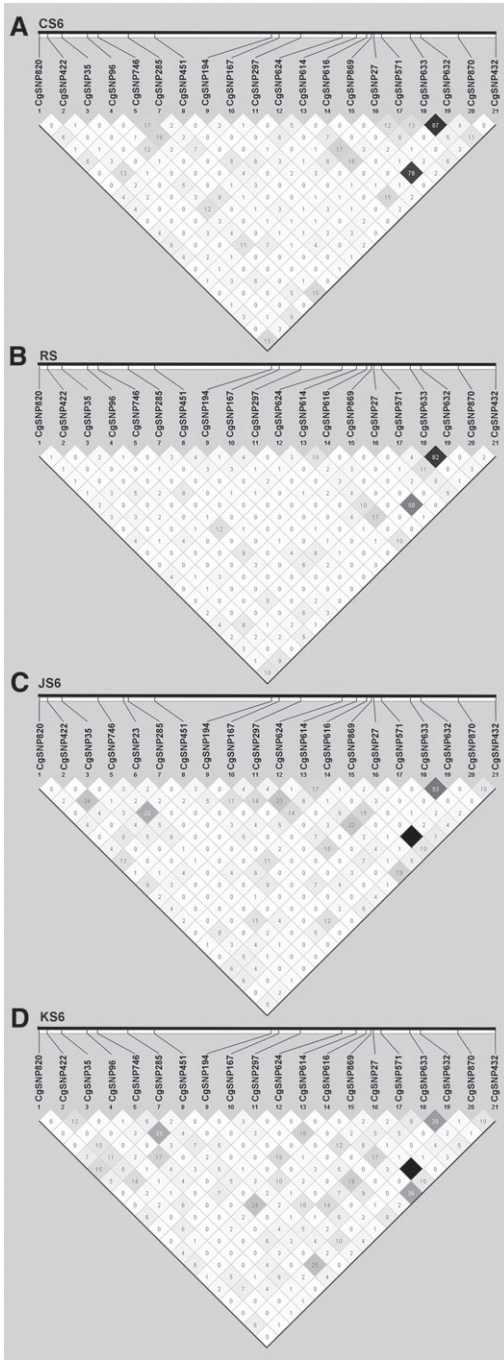

FIGURE 3. *Haploview plot of linkage disequilibrium (r^2) between single-nucleotide polymorphisms on linkage group 7 in the four populations: (A) CS6, (B) RS, (C) JS6, and (D) KS6. Black squares, $r^2 = 1$; white squares, $r^2 = 0$; squares in shades of gray, $0 < r^2 < 1$ (the intensity of the gray shading was proportional to r^2). Values of r^2 (×100) are shown as diamonds.*

significant differences were detected in the average r^2 values between the three selected lines and the wild population for both syntenic and nonsyntenic marker pairs with LD (r^2) beyond population-specific critical values ($P < 0.05$). Taken together, these findings indicated that the LD level of the selected lines was higher than that of the wild population. The similar LD analysis result was also detected in *C. gigas* using 53 microsatellites (Guo et al. 2016). Two factors, genetic drift and selection, may account for the increased levels of LD in selected lines. The large amount of within-population variance (94.65–97.96%) was detected in the LD variance analysis, suggesting that genetic drift might be a significant cause of LD in these populations. This is consistent with the small *Ne* values observed in the selected lines ($Ne \leq 58.5$), which may lead to genetic drift. This may result in loss of rare allelic combinations, thus increasing LD levels in the selected populations. Moreover, the small amount of among-population variance suggested that epistatic selection might play a small role in influencing the LD level in these populations. Selection may reduce the number of alleles in selected lines, thus improving LD of neighboring loci. Moreover, selection may result in LD between unlinked loci (epistasis) despite the fact that the loci are not physically linked (Flint-Garcia et al. 2003). In addition, the result showed that most syntenic marker pairs with high LD levels ($r^2 \geq 0.25$) were detected on linkage group 7 in all populations. This suggested that this linkage group may undergo strong natural selection. The natural selection may increase the frequencies of favorable combinations of alleles in a population, and stable LD will be expected. Nevertheless, this finding may be easily explained by random genetic drift because more than one third of the SNP markers are on linkage group 7.

It is noteworthy that both *Ne* estimation methods (the single-sample method and temporal method) suggested that small *Ne* values ($Ne \leq 58.5$) were detected in the three selected lines. Moreover, Zhong et al. (2016) found that the three selected lines (CS6, JS6, and KS6) exhibited a significant reduction in the observed

TABLE 3. *Estimates of effective population size (*Ne*) of the four oyster populations using LDNE (single-sample method) and GONE (temporal method) with 95% confidence intervals.*[a]

	LDNE			GONE		
Population	≥0.05 *Ne* (LCI–UCI)	≥0.02 *Ne* (LCI–UCI)	≥0.01 *Ne* (LCI–UCI)	Ne^{Fs} (LCI–UCI)	Ne^{Fc} (LCI–UCI)	Ne^{Fk} (LCI–UCI)
CS6	51.4 (36.5–79.3)	53.7 (38.0–83.4)	55.5 (39.2–86.4)	54.3 (39.9–79.6)	56.7 (41.5–83.6)	58.5 (42.8–86.7)
RS	542.1 (137.9–∞)	709.6 (150.0–∞)	629.1 (142.7–∞)	527.9 (146.7–∞)	708.3 (161.2–∞)	623.7 (158.0–∞)
JS6	47.6 (34.2–71.7)	53.8 (38.7–81.5)	53.9 (38.7–81.7)	49.3 (36.8–70.0)	54.4 (40.5–78.0)	54.4 (40.5–78.0)
KS6	50.1 (34.5–80.9)	49.9 (34.8–79.3)	51.3 (35.8–81.3)	56.1 (40.3–85.3)	54.9 (39.8–82.0)	56.3 (40.9–84.3)

LCI = lower 95% confidence intervals; UCI = upper 95% confidence intervals.

[a]Estimates were made using LDNE including SNPs with minor allele frequencies more than 0.01, 0.02, and 0.05; temporal *Ne* was calculated using three correction factor methods, *Fs*, *Fc*, and *Fk*.

TABLE 4. *Ohta's variance components of linkage disequilibrium for the oyster populations.*

	Within-subpopulation components		Percentages of within-subpopulation components of total variance (%)	Between-subpopulation components		Percentages of between-subpopulation components of total variance (%)	Total population component
	D^2_{IS}	Dr^2_{IS}		D^2_{ST}	Dr^2_{ST}		D^2_{IT}
All populations	0.0011	0.02112	97.96	0.02037	0.00044	2.04	0.02156
CS6 and RS	0.00105	0.01255	94.65	0.01215	0.00071	5.35	0.01326
JS6 and RS	0.00111	0.0144	95.11	0.01393	0.00074	4.89	0.01514
KS6 and RS	0.00099	0.01356	95.76	0.01313	0.0006	4.24	0.01416

heterozygosity and observed number of alleles per locus compared with the wild populations ($P \leq 0.05$), which may be another sign of a reduced *Ne* values in the selected lines. The reduced *Ne* value may lead to inbreeding, and thus may not support sustained genetic gains in the selected lines. Li et al. (2009) found that the reduced *Ne* may be affected by several factors, such as a small number of breeders, unequal sex ratios, and contribution of parents for next generations in the Pacific oyster. Therefore, in order to avoid deleterious effects of inbreeding depression and maximize genetic gains of oyster cultivation, the *Ne* per generation in the selected lines should be increased not only by adding more parents but also by equalizing the sex ratio and the contribution of each parent.

It is extremely useful to analyze the extent of LD for determining the appropriate marker density for LD mapping as LD should be observed between markers and QTLs in this method. Our study demonstrated that the extent of LD was about 0.23 cM in the three selected lines, which indicated that a very dense map of SNP markers will be needed for LD mapping in this species. As the sex-averaged map is roughly 514.1 cM in *C. gigas* (Zhong et al. 2014), 2235 SNPs may be required to identify loci of interest if markers are spaced evenly throughout the genome in the selected populations. Therefore, the approach using candidate QTL mapping first and followed by the exploitation of LD for fine mapping may be an effective strategy in *C. gigas*. However, the extent of LD reported here may be underestimated due to a small number of SNP markers and linkage groups used. Therefore, further efforts are needed to significantly increase the number of SNP markers for LD analysis and association mapping in oysters.

Acknowledgments

This study was supported by the grants from National Natural Science Foundation of China (31372524), Shandong Seed Project and Shandong Province (2016ZDJS06A06), and Laboratory for Marine Fisheries Science and Food Production Processes of Qingdao National Laboratory for Marine Science and Technology.

Literature Cited

Ardlie, K. G., L. Kruglyak, and M. Seielstad. 2002. Patterns of linkage disequilibrium in the human genome. Nature Reviews Genetics 3:299–309.

Barnaud, A., V. Laucou, P. This, T. Lacombe, and A. Doligez. 2010. Linkage disequilibrium in natural French grapevine, *Vitis vinifera* L. subsp. *silvestris*. Heredity 104:431–437.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.

Breseghello, F. and M. E. Sorrells. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177.

Conrad, D. F., M. Jakobsson, G. Coop, X. Wen, J. D. Wall, N. A. Rosenberg, and J. K. Pritchard. 2006. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nature Genetics 38:1251–1260.

Coombs, J. A., B. H. Letcher, and K. H. Nislow. 2012. GONe: software for estimating effective population size in species with generational overlap. Molecular Ecology Resources 12:160–163.

Dégremont, L., E. Bédier, and P. Boudry. 2010. Summer mortality of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). II. Response to selection for survival and its influence on growth and yield. Aquaculture 299:21–29.

DOF (Department of Fisheries). 2015. China fishery statistical yearbook. China Agriculture Press, Beijing, China.

Du, F. X., A. C. Clutter, and M. M. Lohuis. 2007. Characterizing linkage disequilibrium in pig populations. International Journal of Biological Sciences 3:166–178.

Ersoz, E. S., J. Yu, and E. S. Buckler. 2007. Applications of linkage disequilibrium and association mapping in crop plants. Pages 97–120 *in* R. K. Varshney and R. Tuberosa, editors. Genomic-assisted crop improvement: Vol. I: Genomics approaches and platforms. Springer Verlag, Berlin, Germany.

Farnir, F., W. Coppieters, J. J. Arranz, P. Berzi, N. Cambisano, B. Grisart, and C. Nezer. 2000. Extensive genome-wide linkage disequilibrium in cattle. Genome Research 10:220–227.

Flint-Garcia, S. A., J. M. Thornsberry, and S. E. Buckler. 2003. Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54:357–374.

Guo, X., Q. Li, Q. Z. Wang, and L. F. Kong. 2012. Genetic mapping and QTL analysis of growth-related traits in the Pacific oyster. Marine Biotechnology 14:218–226.

Guo, X., Q. Li, L. F. Kong, and H. Yu. 2016. Linkage disequilibrium in wild and cultured populations of Pacific oyster (*Crassostrea gigas*). Journal of Ocean University of China 15:327–333.

Hedgecock, D., G. Li, and M. L. Voigt. 2007. Mapping heterosis QTL in the Pacific oyster *Crassostrea gigas*. Aquaculture 272:268.

Hedgecock, D., G. Shin, A. Y. Gracey, D. Van Den Berg, and M. P. Samanta. 2015. Second-generation linkage maps for the pacific oyster *Crassostrea gigas* reveal errors in assembly of genome scaffolds. G3: Genes, Genomes, Genetics 5:2007–2019.

Hill, W. G. 1977. Correlation of gene frequencies between neutral linked genes in finite populations. Theoretical Population Biology 11:239–248.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genetical Research 38:209–216.

Hill, W. G. and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38:226–231.

Hubert, S. and D. Hedgecock. 2004. Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. Genetics 168:351–362.

Jiang, Q., Q. Li, H. Yu, and L. F. Kong. 2013. Genetic and epigenetic variation in mass selection populations of Pacific oyster *Crassostrea gigas*. Genes & Genomics 35:641–647.

Jorde, P. E. and N. Ryman. 2007. Unbiased estimator for genetic drift and effective population size. Genetics 177:927–935.

Kimura, M. 1956. A model of a genetic system which leads to closer linkage by natural selection. Evolution 10:278–287.

Li, L. and X. M. Guo. 2004. AFLP-based genetic linkage maps of the Pacific oyster *Crassostrea gigas* Thunberg. Marine Biotechnology 6:26–36.

Li, M. H. and J. Merilä. 2011. Population differences in levels of linkage disequilibrium in the wild. Molecular Ecology 20:2916–2928.

Li, Q., C. Park, and A. Kijima. 2002. Isolation and characterization of microsatellite loci in the Pacific abalone, *Haliotis discus hannai*. Journal of Shellfish Research 21:811–815.

Li, Q., H. Yu, and R. H. Yu. 2006. Genetic variability assessed by microsatellites in cultured populations of the Pacific oyster (*Crassostrea gigas*) in China. Aquaculture 259:95–102.

Li, R. H., Q. Li, and R. H. Yu. 2009. Parentage determination and effective population size estimation in mass spawning Pacific oyster, *Crassostrea gigas*, based on microsatellite analysis. Journal of the World Aquaculture Society 40:667–677.

Li, Q., Q. Z. Wang, S. K. Liu, and L. F. Kong. 2011. Selection response and realized heritability for growth in three stocks of the Pacific oyster *Crassostrea gigas*. Fisheries Science 77:643–648.

Moen, T., B. Hayes, M. Baranski, P. R. Berg, S. Kjøglum, B. F. Koop, W. S. Davidson, S. W. Omholt, and S. Lien. 2008. A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers. BMC Genomics 9:1.

Nei, M. and F. Tajima. 1981. Genetic drift estimation of effective population size. Genetics 98:625–640.

Nell, J. A. 2001. The history of oyster farming in Australia. Marine Fisheries Review 63:14–25.

Ohta, T. 1982a. Linkage disequilibrium with the island model. Genetics 101:139–155.

Ohta, T. 1982b. Linkage disequilibrium due to random drift in finite subdivided populations. Proceedings of the National Academy of Sciences 79:1940–1944.

Pollak, E. 1983. A new method for estimation the effective population size from allele frequency changes. Genetics 104:531–548.

Rexroad, C. E. and R. L. Vallejo. 2009. Estimates of linkage disequilibrium and effective population size in rainbow trout. BMC Genetics 10:83.

Rieger, R., A. Michaelis, and M. M. Green. 2012. Glossary of genetics: classical and molecular, Volume 332. Springer Science & Business Media, New York, New York, USA.

Sauvage, C., P. Boudry, D. J. De Koning, C. S. Haley, S. Heurtebise, and S. Lapègue. 2010. QTL for resistance to summer mortality and OsHV-1 load in the Pacific oyster (*Crassostrea gigas*). Animal Genetics 41:390–399.

Sokal, R. R. and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research, 3rd edition. W.H. Freeman, New York, New York, USA.

Wang, Q. Z., Q. Li, L. F. Kong, and R. H. Yu. 2012. Response to selection for fast growth in the second generation of Pacific Oyster (*Crassostrea gigas*). Journal of Ocean University of China 11:413–418.

Wang, J. P., L. Li, and G. F. Zhang. 2016. A high-density SNP genetic linkage map and QTL analysis of growth-related traits in a hybrid family of oysters (*Crassostrea gigas* × *Crassostrea angulata*) using genotyping-by-sequencing. G3: Genes, Genomes, Genetics 6:1417–1426.

Waples, R. S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics 7:167–184.

Waples, R. S. and C. Do. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753–756.

Waples, R. S. and C. Do. 2010. Linkage disequilibrium estimates of contemporary *Ne* using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications 3:244–262.

Ward, R. D., L. J. English, D. J. McGoldrick, G. B. Maguire, J. A. Nell, and P. A. Thompson. 2000. Genetic improvement of the Pacific oyster *Crassostrea gigas* (Thunberg) in Australia. Aquaculture Research 31:35–44.

Yeh, F. C., R. C. Yang, T. Boyle, Z. H. Ye, and J. X. Mao. 1999. POPGENE version 1.32: the user friendly software for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Canada.

Zhao, Y., H. Wang, W. Chen, and Y. Li. 2014. Genetic Structure, linkage disequilibrium and association mapping of verticillium wilt resistance in Elite Cotton (*Gossypium hirsutum* L.) germplasm population. Plos One 9:e86308.

Zhong, X. X., Q. Li, H. Yu, and L. F. Kong. 2013. Development and validation of single nucleotide polymorphism markers in the Pacific oyster, *Crassostrea gigas*, using high-resolution melting analysis. Journal of the World Aquaculture Society 44:455–465.

Zhong, X. X., Q. Li, X. Guo, H. Yu, and L. F. Kong. 2014. QTL mapping for glycogen content and shell pigmentation in the Pacific oyster *Crassostrea gigas* using microsatellites and SNPs. Aquaculture International 22:1877–1889.

Zhong, X. X., D. D. Feng, H. Yu, L. F. Kong, and Q. Li. 2016. Genetic variation and breeding signature in mass selection lines of the Pacific oyster (*Crassostrea gigas*) assessed by SNP markers. Plos One 11:e0150868.

Zhu, C., M. Gore, E. S. Buckler, and J. Yu. 2008. Status and prospects of association mapping in plants. The Plant Genome 1:5–20.

Supporting Information

Additional supporting information may be found online in the supporting information tab for this article.