Aquaculture 561 (2022) 738737

Genetic variability of mass-selected and wild populations of Iwagaki oyster (*Crassostrea nippona*) revealed by microsatellites and mitochondrial COI sequences

Yiming Hu [a],

Available online 17 August 2022

populations.

The potential for genetic improvement of desired traits through mass selection is well documented in shellfish species, particularly due to the short generation intervals and the possibility of stringent selection as a consequence of their great fecundity (Gjedrem and Baranski, 2010). Although the risk of unintentional inbreeding may be further increased in the selection scheme due to non-random mating and the absence of pedigree records (Bentsen and Olesen, 2002; Boudry et al., 2002), mass selection with a high selection intensity has been widely used in shellfish species to enhance growth or disease resistance (Li et al., 2011; Dégremont et al., 2015). What must be highlighted, however, is that minimization of inbreeding and control of the genetic structure are key issues for the selected lines to obtain sustained improvement. To preserve genetic diversity during mass selection, various effective measures to prevent inbreeding accumulation in mass-selected populations have been proposed. For instance, some cost-effective and easy-to-operate preventive measures have been applied to the oyster breeding process, including the implementation of abundant broodstock and a balanced sex ratio, and successfully maintained the genetic diversity in two shell color lines of the Pacific oyster for growth (Han et al., 2019; Xu et al., 2019b). The other breeding strategies such as subdividing the breeding nucleus and keeping independent sublines to store rare alleles, cross-breeding with the wild animals, and alleviating selection intensity were demonstrated to be valid as precautionary methods (In et al., 2016; Chen et al., 2017). Genetic tools can be used to enable a better understanding of how efficient current breeding practices are at maintaining genetic variation without direct information about the parents, which can open the possibility for timely modification of pertinent schemes in subsequent selective breeding (Hillen et al., 2017).

The Iwagaki oyster (*Crassostrea nippona*) is naturally distributed in coastal areas of East Asia such as China, Japan and South Korea (Itoh et al., 2004; Lu et al., 2017). The biological properties and its market price have recently prompted widespread interest in initializing the aquaculture industry of *C nippona*, because *C nippona* can maintain a high glycogen content and delicate flavor and thus has the potential of substitution for the edibility-restricted Pacific oyster (*C i as*) during the summer (Okumura et al., 2005; Masahiro et al., 2018). However, due to unimproved growth performance, the industry of *C nippona* has never matured into a large-scale aquaculture producer during long-term domestication (Xu et al., 2019a). Therefore, we initiated the successive three-generation mass selection to improve the growth performance of *C nippona*. A sustained genetic gain of approximately 10% per generation for shell height had been obtained at harvest, yet the genetic impact of intense artificial selection on the genetic variability of *C nippona* was not fully understood and rarely monitored. This study is the first to report genetic variation in *C nippona* using both 15 microsatellite loci and mitochondrial COI sequences (mtCOI), and aimed at estimating the

(Do et al., 2014), which implements the bias-correction and reads genotypic data in standard formats, and $_{e-lin}$ values with 95% confidence intervals (95% CI) were given. Pairwise values and analysis of molecular variance (AMOVA) were utilized to assess population differentiation and partition the genetic variance within/among populations using ARLEQUIN v.3.5 (Excoffier and Lischer, 2010). Genetic relationship among populations was estimated in a bidimensional space by performing principal coordinates analysis (PCoA) implemented in GenAlEx 6.5 (Peakall and Smouse, 2012). A pairwise matrix assessing allele frequency heterogeneity with six populations was computed using GenAlEx 6.5 based on the method of *ei's* unbiased genetic distance (*ei's*) (Hedrick, 2009), and a neighbor-joining tree was displayed in MEGA v.5.0 (Tamura et al., 2011).

Mitoc on rial se uencin an analysis

The primers originally developed by Liu and Li (2018) were used to amplify the mtCOI of *C nippona*, and the amplified fragment was sequenced by Personal Company (Shanghai, China). Sequencing results of each population were first edited and assembled by SeqMan software from DNASTAR and then the aligned sequences in MEGA v.5.0 were used to calculate the number of haplotypes ( ), haplotype diversity ( ) and nucleotide diversity ($_i$) by DNASP v.5.10 (Librado and Rozas, 2009).

3. Results

enetic i ersity

No evidence of genotypic errors due to large allele dropout and stuttering was discovered in all loci by Micro-Checker. The frequencies of null alleles above 0.2 were detected at only nine locus-population combinations: MK078347 and MK078341 in the G1 and the remaining combinations in three wild populations. Remarkably, neither inclusion nor deletion of these loci had a qualitative impact on the experimental outcome. Meanwhile, the *C* of each marker was above

subsequent analysis was performed based on all 15 microsatellite loci in six populations.

No significant loss of genetic diversity in $_a$ (7.20–10.07 across six populations), (1.34–1.83), $_o$ (0.46–0.64), $_e$ (0.63–0.75) and $_r$ (6.35–10.07) was found from three wild populations to the selected lines (Table 2). Meanwhile, the abundant genetic diversity of selected lines was successfully maintained during three generations of mass selection due to no detectable loss and the mean values of $_a$ (7.35–9.87 across three selected lines), $_o$ (0.58–0.64), $_e$ (0.67–0.71) and $_r$ (6.35–7.72) were at relatively high levels. The reduction in $_a$ (23.71%), (11.11%), $_e$ (5.63%), $_r$ (17.75%) and C (6.63%) observed from G1 to G3 might indicate the potential influence of an intense mass selection on the genetic variability of the selected lines. The C ranging from 0.5914 in JN to 0.7265 in CZ were all >0.5, indicating high polymorphism in six populations, and the positive $_{is}$ value suggested heterozygote deficiency could be prevalent in all analyzed populations. More locus-population combinations deviate from the Hardy-Weinberg equilibrium ( ) after Bonferroni correction was observed in the mass-selected lines (9–13) than in the wild populations (4–8).

A total of ten mtCOI haplotypes were found by analyzing the sequencing results of COI fragments from 120 oysters (Table 3). Besides the conspicuous haplotype (1), which was detected in all populations and had a frequency of 68.33% (82/120) in all analyzed samples, remaining haplotypes were not shared between the selected lines and wild populations. Three of the ten haplotypes were shown in the selected lines and the other one to three private haplotypes (haplotypes present in a single population) were distributed in three wild populations. The mean of six populations ranged from 0.100 in KG to 0.721 in G1, while the mean $_i$ of six populations ranged from 0.018% in KG to 0.197% in G1.

ffecti e population si e

The effective population size and 95% confidence intervals of G1-G3 were measured based on linkage equilibrium methods (Table 1). Estimated values in each selected line were as follows: G1 ($_{e-lin}$: 64.1, 95% CI: 54.4–77.1), G2 ($_{e-lin}$: 25.3, 95% CI: 22.9–28.0) and G3 ($_{e-lin}$: 47.4, 95% CI: 39.5–59.0). Except for G1, the $_{e-lin}$ value in the other two generations was lower than the actual number of broodstock. The lowest $_{e-lin}$ in G2 could be related to the unbalanced parental contributions and insufficient broodstock and then a higher $_{e-lin}$ value was found in G3 (47.4) compared with G2.

enetic ifferentiation an population structure

Analysis of molecular variance (AMOVA) revealed that most of the variation in the selected and wild populations was observed within population (84.93–99.49%) (Table 4). The global of 0.151 in the wild populations was recorded (< 0.01), suggesting large genetic differentiation among populations, while that of the selective breeding lines was 0.005. Similarly, the values of pairwise (0.006–0.010) and *ei's* (0.003–0.023) in G1-G3 were relatively low and gradually increased between adjacent breeding generations (Table 5). In this study, the highest pairwise (0.211) and *ei's* (0.730) were found in a pairwise comparison between JN and KG. All pairwise values were

Table 3
Genetic diversity of the selected and wild populations of *C nippona* at mtDNA COI region.

Haplotype	Selected lines			Wild populations			Total
	G1	G2	G3	CZ	JN	KG	
1	7	5	17	16	18	19	82
2				2			2
3				1			1
4				1			1
5	4	2	1				7
6	9	13	1				23
7					1		1
8			1				1
9					1		1
10						1	1
	3	3	4	4	3	2	
	0.721 ± 0.065	0.532 ± 0.10	0.284 ± 0.128	0.363 ± 0.131	0.195 ± 0.115	0.100 ± 0.088	
$_i$ (%)	0.197 ± 0.026	0.126 ± 0.029	0.057 ± 0.027	0.141 ± 0.080	0.124 ± 0.107	0.018 ± 0.016	

, number of haplotypes; , haplotype diversity; $_i$, percent nucleotide diversity.

Table 4
Analysis of molecular variance (AMOVA) for the selected and wild populations based on 15 microsatellite loci.

Source of variance	d *f*	Variance components	Percentage of variation	-statistic
Among selected lines				
Among populations	2	0.02708	0.50778	= 0.00508
Among individuals/ within population	141	0.70534	13.22569	= 0.13293*
Within individuals	144	4.60069	86.26653	= 0.13733*
Total	287	5.33312		
Among wild populations				
Among populations	2	0.91433	15.07175	= 0.15072*
Among individuals/ within population	93	1.43867	23.71484	= 0.27923*
Within individuals	96	3.71354	61.21342	= 0.38787*
Total	191	6.06655		

* Significant at < 0.01.

Table 2
Genetic parameters within the selected and wild populations based on 15 microsatellite loci.

Population	$_a$		$_o$	$_e$	$_r$	C	$_{is}$	
Selected lines								
G1	9.87 ± 5.99[a]	1.62 ± 0.60[ab]	0.58 ± 0.25[a]	0.71 ± 0.14[a]	7.72 ± 4.17[ab]	0.6745	0.18	12
G2	9.60 ± 5.59[a]	1.60 ± 0.58[ab]	0.64 ± 0.21[a]	0.71 ± 0.14[a]	7.51 ± 4.01[ab]	0.6735	0.10	13
G3	7.53 ± 4.90[a]	1.44 ± 0.58[ab]	0.62 ± 0.21[a]	0.67 ± 0.15[a]	6.35 ± 3.68[a]	0.6298	0.07	9
Wild populations								
CZ	10.07 ± 5.06[a]	1.83 ± 0.63[b]	0.58 ± 0.21[a]	0.75 ± 0.15[a]	10.07 ± 5.06[b]	0.7265	0.23	4
JN	7.20 ± 2.60[a]	1.34 ± 0.40[a]	0.46 ± 0.22[a]	0.63 ± 0.13[a]	6.48 ± 2.23[a]	0.5914	0.27	8
KG	8.93 ± 5.16[a]	1.52 ± 0.60[ab]	0.48 ± 0.30[a]	0.67 ± 0.17[a]	7.28 ± 3.74[a]	0.6355	0.28	7

$_a$: average number of alleles; , Shannon's information index; $_o$, observed heterozygosity; $_e$, expected heterozygosity; $_r$, alleles richness; C, polymorphic information content; $_{is}$, inbreeding coefficient; , number of loci deviating from Hardy-Weinberg equilibrium. Different letters in the same column indicate significant differences among populations (< 0.05).

Table 5
Estimated pairwise (lower diagonal) and *ei's* (upper diagonal) values of *C nippona* based on 15 microsatellite loci.

	G1	G2	G3	ZC	JN	KG
G1		0.003	0.017	0.315	0.580	0.707
G2	0.001		0.023	0.281	0.528	0.720
G3	0.008*	0.010*		0.251	0.527	0.616
ZC	0.087*	0.080*	0.083*		0.260	0.332
JN	0.170*	0.161*	0.172*	0.095*		0.730
KG	0.182*	0.184*	0.180*	0.105*	0.211*	

The significance of population pairwise tested by 1000 permutations.
* Significantly at < 0.05.

significantly different at the < 0.05 level except between G1 and G2.

To further understand the relationship among different populations, the neighbor-joining tree was constructed based on *ei's* unbiased genetic distance (Fig. 2). The tree topology showed that all experimental populations were clustered into two main branches. Wild population KG was placed on a separate branch of a cluster dendrogram and the other populations clustered on another branch which was further classified into three subgroups according to JN, CZ and the selected lines. Meanwhile, there was increasing differentiation between adjacent generations in the subcluster of G1-G3. Similar conclusions were confirmed by the results presented in PCoA that three principal components accounted for 30.51% of the total molecular variation, and the Coordinate axis 1 and 2 explained 16.73% and 8.90%, respectively (Fig. 3).

4. Discussion

In selective breeding programs, a fundamental knowledge of how to maintain the maximum level of genetic variability in the successive breeding process is requisite to obtain a sustained response from long-term selection (Lind et al., 2009). It is generally believed that the genetic diversity of mass-reared populations is inadvertently eroded in highly fecund shellfish due to genetic drift and unintentional inbreeding (Xu et al., 2019b; Rhode et al., 2020). In addition, the intense selection for best-performing individuals to achieve faster gains might further decrease the effective population size. It therefore should be necessary systematically to evaluate the genetic diversity and population structure of currently selected lines to inform decision-making for stock management and slow down consanguineous-derived adverse effects.

The complementary combination of microsatellites with mitochondrial COI sequences has been extensively applied to investigate the genetic variability, population genetic structure and demographic history in many bivalves (In et al., 2016; Cordero et al., 2017; Xu et al., 2019b). Here, this investigation revealed that no significant differences in genetic diversity, such as $_a$, , $_o$, $_e$ and $_r$, was observed in the mass-selected lines when compared to wild populations, demonstrating that current breeding procedures for *C nippona* appeared efficient at maintaining available genetic variability. In the analysis results of mtCOI, relatively few haplotypes (2–4) were observed in analyzed populations

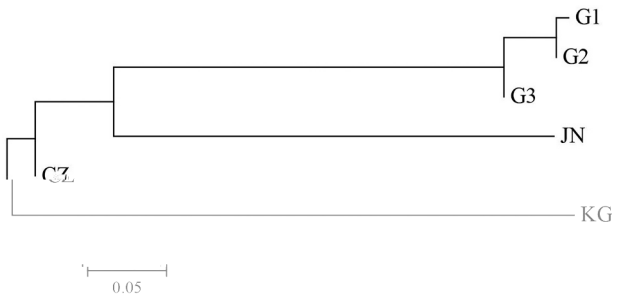

Fig. 2. Neighbor-joining tree of the selected and wild populations using *ei's* unbiased genetic distance based on 15 microsatellite loci.

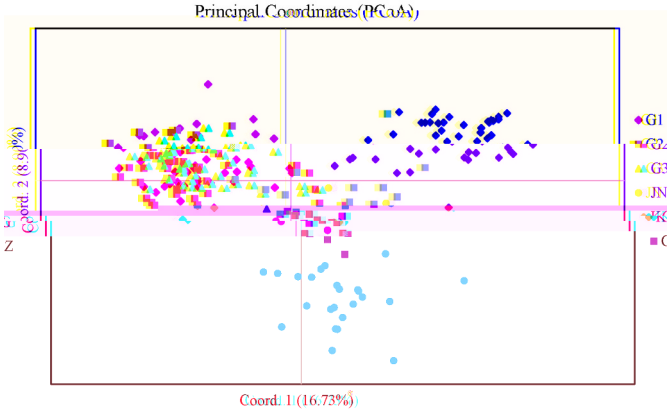

Fig. 3. Principal coordinates analysis (PCoA) of the selected and wild populations of *C nippona* based on 15 microsatellite loci. Coordinate axis 1 explains 16.73% of the variation, coordinate axis 2 explains 8.90% of the variation, and Coordinate axis 3 (not shown) explains 4.88%.

including three wild populations compared to other oyster species, and the inference was that the low number of haplotypes cannot simply be considered evidence of inbreeding but might reflect the overall level of haplotype diversity in *C nippona*, as the similar result was also been shown in other studies (In et al., 2016; Xu et al., 2019b). Nevertheless, the number of private haplotypes and haplotype diversity were reduced in the selected lines implying the potential influence of mass selection on the genetic variability of the fast-growing lines, which might be consistent with the observed slight reductions in $_a$, , $_o$, $_e$, $_r$ and *C*. Similar conclusions were reported in the study of Sydney rock oysters (*accosstrea lo erata*), and In et al. (2016) found that the line with the fewest DNA microsatellite alleles also had the fewest haplotype. Losses of microsatellite alleles variation at the population level could reduce the potentially important functional genetic variation in the genome and affect the fitness of subsequent generations with the loss of rare alleles (Lind et al., 2009). In addition, the potential loss of genetic diversity during farming seems unavoidable which has been demonstrated both theoretically and empirically. In this study, a decreasing trend was found in $_a$ (23.71%) during three generations of mass selection suggesting that the breeding strategy might require prompt intervention to avoid the inadvertent loss of $_a$ in subsequent selective breeding. A multi-line breeding program may be feasible to store the rare alleles by subdividing the breeding nucleus into multiple independent sublines, and In et al. (2016) suggested that interline crossing can help restore the genetic diversity among selected lines even after many generations of selection. Although the expected heterozygosity higher than the observed heterozygosity was found in each of the selected lines which might be indicative of some level of inbreeding, the closed line of *C nippona* showed limited evidence for an increase in a homozygous state. Analogous pattern was shown in hatchery-cultured *incta a a i a* (Lind et al., 2009) and genetically improved *C i as* (Han et al., 2019), further confirming the conclusion that alleles are more susceptible to change than heterozygosity in the immediate term. The reason for this phenomenon might be that the low-frequency alleles contribute little to overall heterozygosity (Lundrigan et al., 2005). What must be highlighted, finally, is that indications of a slight decline in overall genetic diversity are detected in the selected lines, suggesting that regular genetic monitoring is indispensable for avoiding potential problems associated with inbreeding depression in long-term selection.

In this study, 53 of 90 population-locus combinations that deviated from Hardy-Weinberg equilibrium were detected in total, and only 9 combinations showed heterozygote excess, 6 of which were in selective breeding lines. The high prevalence of heterozygote deficiency has been reported previously in studies of mollusk species (Li et al., 2007; Chen et al., 2017), which could be caused by null alleles, non-random mating,

a commixture of independent populations, and artificial and natural selection during seed production and cultivation (Chen et al., 2017). As expected, the of the selected lines (44) was much higher than those of the wild populations (19), exhibiting obvious signatures of artificial selection, and the maximum value of in G2 (13) might be associated with the lowest effective population size which was also recorded in the mass-selected lines of Pacific abalone (*aliotis iscus annai*) (Chen et al., 2017). In addition, both the relatively low genetic diversity and high observed in JN indicated non-random mating in the wild population, probably because the continuous input of low-variability oysters from hatcheries in the location had diluted the natural genetic resources of *C nippona* or sampling effects (Zhang et al., 2010; Cordero et al., 2017). This might also be the reason why the inbreeding coefficient of the selected lines in this study is lower than that of the wild populations.

The fluctuation of effective population size ($_e$) is generally conditioned by farming constraints limiting the contribution of selected parents to future generations (Lallias et al., 2010). In aquaculture practice, $_e$ can be diminished by insufficient broodstocks, biased sex ratio, the unequal contribution of gametes, and different viability of gametes (Li et al., 2007). Simultaneously, selected parents based on the best performance of important commercial traits can lead to further reduction of $_e$ in mass-farmed populations due to broodstocks just from a few outstanding families (Bentsen and Olesen, 2002). In the present study, the linkage disequilibrium method was applied to predict the effective population size ($_{e-lin}$) and the results showed that $_{e-lin}$ for G1, G2 and G3 were 64.1, 25.3 and 27.4, respectively. The lowest $_{e-lin}$ in G2 might be related to unbalanced parental contributions, insufficient broodstocks and high

Bentsen, H.B., Olesen, I., 2002. Designing aquaculture mass selection programs to avoid high inbreeding rates. Aquaculture 204, 349–359. https://doi.org/10.1016/S0044-8486(01)00846-8.

Boudry, P., Collet, B., Cornette, F., Hervouet, V., Bonhomme, F., 2002. High variance in reproductive success of the Pacific oyster (*Crassostrea i as*, Thunberg) revealed by microsatellite-based parentage analysis of multifactorial crosses. Aquaculture 204, 283–296. https://doi.org/10.1016/S0044-8486(01)00841-9.

Chapuis, M.-P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24, 621–631. https://doi.org/10.1093/molbev/.

Chen, N., Luo, X., Lu, C., Ke, C., You, W., 2017. Effects of artificial selection practices on loss of genetic diversity in the Pacific abalone, *aliotis iscus annai*. Aquac. Res. 48, 4923–4933. https://doi.org/10.1111/are.13311.

Chen, Y.L., Xu, C.X., Li, Q., 2022. Genetic diversity in a genetically improved line of the Pacific oyster *Crassostrea i as* with orange shell based on microsatellites and mtDNA data. Aquaculture 549, 737791. https://doi.org/10.1016/j.aquaculture.2021.737791.

Cordero, D., Delgado, M., Liu, B., Ruesink, J., Saavedra, C., 2017. Population genetics of the Manila clam (*u itapes p ilippinaru*) introduced in North America and Europe. Sci. Rep. 7, 1–13. https://doi.org/10.1038/srep39745.

Dégremont, L., Nourry, M., Maurouard, E., 2015. Mass selection for survival and resistance to OsHV-1 infection in *Crassostrea i as* spat in field conditions: response to selection after four generations. Aquaculture 446, 111–121. https://doi.org/10.1016/j.aquaculture.2015.04.029.

Diyie, R.L., Agyarkwa, S.K., Armah, E., Amonoo, N.A., Owusu-Frimpong, I., Osei-Atweneboana, M., 2021. Genetic variations among different generations and cultured populations of Nile Tilapia (*reoc ro is niloticus*) in Ghana: application of microsatellite markers. Aquaculture 544, 737070. https://doi.org/10.1016/j.aquaculture.2021.737070.

Do, C., Waples, R.S., Peel, D., Macbeth, G., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (*e*) from genetic data. Mol. Ecol. Resour. 14, 209–214. https://doi.org/10.1111/1755-0998.12157.

Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. Mol. Ecol. Resour. 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Gjedrem, T., Baranski, M., 2010. Selective Breeding in Aquaculture: An Introduction. Springer Science & Business Media. http://refhub.elsevier.com/S0044-8486(20)30684-0/rf0045.

Gjedrem, T., Robinson, N., Rye, M., 2012. The importance of selective breeding in aquaculture to meet future demands for animal protein: a review. Aquaculture 350, 117–129. https://doi.org/10.1016/j.aquaculture.2012.04.008.

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. J. Hered. 86, 485–486. https://doi.org/10.1093/oxfordjournals.jhered.a111627.

Han, Z.Q., Li, Q., Liu, S.K., Yu, H., Kong, L.F., 2019. Genetic variability of an orange-shell line of the Pacific oyster *Crassostrea i as* during artificial selection inferred from microsatellites and mitochondrial COI sequences. Aquaculture 508, 159–166. https://doi.org/10.1016/j.aquaculture.2019.04.074.

Hedrick, P.W., 2009. Genetics of Populations. Jones