Aquacult Int
DOI 10.1007/s10499-017-0127-0

Comparison of microsatellites and SNPs for pedigree analysis in the Pacific oyster *Crassostrea gigas*

Ting Liu [1] · **Qi Li** [1,2] · **Lingfeng Kong** [1] · **Hong Yu** [1]

Received: 3 August 2016 / Accepted: 9 February 2017

Abstract Knowledge of the pedigree relationships between individuals is a prerequisite in genetics research, and the application of molecular markers for pedigree analysis has been a booming science for over a decade. Owing to the high variability, microsatellites are considered as the marker of choice for studies on pedigree analysis. Nevertheless, single nucleotide polymorphisms (SNPs) have been increasingly used for this purpose in recent years due to the low mutation rate and genotyping error rate. To compare the utility of microsatellites and SNPs in assigning parentage in the Pacific oyster (*Crassostrea gigas*), we genotyped 384 parental and offspring individuals using 12 multiplexed microsatellites and 50 SNPs. In this study, all microsatellite loci showed high informative (*PIC* >0.5), while most SNPs were middle informative (0.25 <*PIC* <0.5). CERVUS simulations revealed that using nine microsatellites or 38 SNPs, the power of parental assignment could reach 100%. Pedigree analysis of real offspring demonstrated that 100% of the offspring were unambiguously assigned to a pair of parents when nine microsatellites or 50 SNPs were used. For microsatellites, the combined exclusion power with one parent known (EXCL2) could reach one when three microsatellites multiplex PCRs or more were used, whereas EXCL2 was 0.9999 for the 50 SNPs. In general, six SNPs were needed to obtain an equivalent exclusion power for pedigree analysis with a microsatellite locus in *C. gigas*. The information obtained in this study will be useful for assigning parentage in *C. gigas* using both marker systems.

Keywords *Crassostrea gigas* · Microsatellite multiplex PCR · Single nucleotide polymorphisms · Pedigree analysis

✉ Qi Li
qili66@ouc.edu.cn

[1] Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Yushan Road 5, Qingdao 266003, China

[2] Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao, Shandong, China

Published online: 09 March 2017

Introduction

Aquaculture is now the fastest growing animal production worldwide. In Asia Pacific region, it significantly contributes to local food/nutritional security (FAO 2014). Such an important industry is expected to be based on genetically improved stocks. As a result, genetic improvement of aquaculture stocks is urgently needed. However, genetic improvement is still limited due to the complexity and cost of maintaining adequate pedigree information in aquaculture stocks. A pedigree is one of the simplest concepts in biology (Pemberton 2008), but it is of paramount importance for breeding program as well as genetic management to obtain sustainable genetic improvement in aquaculture industry. Correct pedigree information is necessary for the accurate estimation of breeding values, whereas inaccurate and incomplete genealogies lead to deviations and errors in breeding program and genetic management. Therefore, the provision of a credible and cost-effective method to obtain pedigrees in any species and rearing system was supposed to be of great interest for a successful breeding program in aquaculture.

DNA markers are becoming increasingly important in animal breeding. Molecular parentage analysis using DNA markers has made the pedigree tracing possible in aquaculture species. This method is expected to be promising as it can avoid the initial investment in separate family rearing units and limits associated biases, even more in species with high larval mortality, small larval size, and initial live feeding (Vandeputte and Haffray 2014).

The principle of molecular pedigree analysis is very simple. It is based on the simple concept that parents pass on one of two alleles at each locus to the offspring, which therefore carries one allele from each parent (Herbinger et al. 1995). Microsatellites are of high variability and wide availability. Thus, they have been the preferred molecular markers to trace the genealogical relationships in numerous aquaculture species (Castro et al. 2007; Lallias et al. 2010; An et al. 2011; Vandeputte et al. 2011; Fu et al. 2013; Zhang et al. 2016). However, considering the presence of null alleles and higher mutation rate, which would interfere with accurate pedigree analysis with microsatellites, application of single nucleotide polymorphisms (SNPs) is growing exponentially (Guichoux et al. 2011) and they have been applied for large-scale parentage studies (Hauser et al. 2011; Pino-Querido et al. 2015) because of the speed of high-through-put screening, low genotyping error and easy transfer-ability between laboratories (Anderson and Garza 2006). In addition, SNPs have lower mutation rates (10^{-9}) per locus per generation than microsatellites (10^{-3}–10^{-4}) (Ellegren 2000), and they occur more frequently in the genome than microsatellites (Bester et al. 2008). Nevertheless, the resolving power of SNPs for pedigree assignment needs be compensated by large marker sets because each SNP locus typically has only two alleles.

As one of the most commercially important species in aquaculture, the Pacific oyster (*Crassostrea gigas*) has the highest worldwide production of cultured aquatic species. In China, *C. gigas* is still in an early stage of domestication. The marker-assisted selection could accelerate the development of the oyster industry and maximize efficiency of aquaculture production. A large number of microsatellite and SNP markers have been developed in *C. gigas* (Sekino et al. 2003; Yamtich et al. 2005; Bai et al. 2009; Zhong et al. 2013; Lapegue et al. 2014). Initially, microsatellite was amplified as single locus for pedigree analysis. To reduce the cost and time required for microsatellite genotyping so that shed the burden of laboratory works, multiplex PCR panels of microsatellites of *C. gigas* have been developed (Taris et al. 2005; Li et al. 2010; An et al. 2013; Liu et al. 2016).

In the present study, in order to compare the utility of microsatellites and SNPs in assigning parentage in *C. gigas*, we genotyped more than 300 individual oysters using 12 multiplexed microsatellites and 50 SNPs and examined the results of parentage analyses using both marker systems and the resolving power of analyses employing different numbers of microsatellites and SNPs.

Materials and methods

Experimental samples and DNA extraction

In 2007, 2-year-old Pacific oysters from one base population in Rushan, Shandong province, were used to establish the first-generation selection line for fast growth. The oysters derived from natural seed which were collected on local coast, and were cultured on ropes suspended from rafts along the coastal regions. Samples of one-year-old *C. gigas* were collected in October 2015 from one cultured population in Rongcheng, Shandong province, China. They were the eighth-generation offspring produced by successive selection. In July 2014, 100 individuals (50 males and 50 females) were selected from the seventh-generation strains to serve as parents for the eighth-generation selected strain. Tissues from all individuals were saved in pure ethanol until DNA extraction.

Genomic DNA was extracted from adductor muscle tissue as previously described by Li et al. (2006). The concentration and quality of DNA were assessed using a NanoDrop 2000 spectrophotometer and by running a small amount on a 1% agarose gel. The extracted genomic DNA was stored at −20 °C until genotyping.

DNA marker systems

To provide a benchmark for comparison, 100 parents and 284 progenies were genotyped by 12 microsatellites and 50 SNP markers, respectively. The microsatellite system consisted in a total of 12 microsatellite DNA markers multiplexed. In brief, forward primers were modified with M13-tail with different fluorescent dyes (FAM, PIT, VIC and NED). PCR reactions were carried out in 10 μl volumes containing 1× PCR buffer, 0.2 mM dNTP mix, 2.0 mM $MgCl_2$, 0.15 μM forward primer, 0.06 μM reverse primer, 0.15 μM universal primer, 0.25 U *Taq* DNA polymerase, and about 50 ng template DNA. Thermal cycling was as follows: first denaturation at 94 °C for 3 min; then 35 cycles of 30 s at 94 °C, 60 s at the optimal annealing temperature, and 75 s at 72 °C; 8 cycles of 30 s at 94 °C, 60 s at 53 °C, 75 s at 72 °C, with a final extension of 10 min at 72 °C. PCR products were genotyped on ABI-3130 with LIZ500 as internal size standard. Data were processed with the GeneMapper v4.0 software.

For the 50 previously developed SNPs (Zhong et al. 2013), PCR was performed in a 10 μl reaction mixture on a LightCycler® 480 real-time PCR instrument (Roche Diagnostics). The mixture contained 10× PCR buffer, 0.2 mM dNTP mix, 1.5 mM $MgCl_2$, 0.2 μM of each primer set, 5 μM SYTO®9 (Invitrogen Foster City, CA, USA), 0.25 U *Taq* DNA polymerase, and about 10 ng template DNA. The amplification procedure was achieved as follows: initial denaturation at 95 °C for 5 min, 45–50 cycles of denaturation at 95 °C for 10 s, annealing temperatures at 60 °C for 10 s, then extension at 72 °C for 10 s. Following amplification, melting curves were generated by collecting fluorescence data between 60 °C to 90 °C. Data were analyzed using the LightCycler® 480 Gene Scanning Software 1.5 (Roche Diagnostics).

Genetic diversity and parentage assignment

We calculated the number of alleles (*Na*), polymorphic information content (*PIC*), the observed heterozygosity (*Ho*), and expected heterozygosity (*He*) using CERVUS 3.0 (Kalinowski et al. 2007) for each microsatellite locus. All microsatellites were tested for Hardy–Weinberg equilibrium with a chi-square test (using Yates' correction) in CERVUS 3.0, and those loci showed with "not done" were tested again using exact test implemented in the program GENEPOP v 4.5.1 (Rousset 2008). The frequency of null alleles at each locus from population data was obtained following the method by Van Oosterhout et al. (2004) as implemented in the Micro-Checker 2.2.3 software. Basic genetic parameters of SNPs including *Na*, *PIC*, *Ho*, *He*, HWE, and null allele frequency were computed as for microsatellites. Allele frequencies for microsatellites and SNPs in 384 individuals were calculated directly.

Pedigree analysis of microsatellite and SNP genetic data was consequently performed with the likelihood-based approach in CERVUS 3.0. Both simulation and real parentage analysis were conducted. The exclusion probability of each locus based on the genotype of no parent known and on the genotype of one parent known was named as EXCL1 and EXCL2, respectively. Both EXCL1 and EXCL2 were calculated for each locus using program CERVUS 3.0. And the combined exclusion probabilities of different marker sets were also calculated. The exclusion power (P_u) of microsatellites multiplex PCRs and SNPs was also computed using Eqs. 1 and 7 as described in Vandeputte (2012).

Results

Genetic diversity of microsatellites and SNPs

Genetic diversity parameters were estimated in the 384 individuals. Twelve microsatellite loci were arranged into four multiplex PCRs (Table 1). The allele number (*Na*) of these loci ranged

Table 1 Information of 4 microsatellite multiplex PCRs

Multiplex sets	Locus	Ta (°C)	Concentration of forward primer (μM)	Concentration of reverse primer (μM)
Panel 1	ucdCg-120	58	0.06	0.15
	ucdCg-198	58	0.06	0.15
	ucdCg-117	58	0.06	0.15
	FAM-M13 (-21)	53	–	0.15
Panel 2	Crgi3	58	0.06	0.15
	ucdCg-146	58	0.06	0.15
	uscCgi-210	58	0.06	0.15
	PIT-M13 (-21)	53	–	0.15
Panel 3	ucdCg-170	58	0.06	0.15
	ucdCg-156	58	0.06	0.15
	ucdCg-199	58	0.06	0.15
	VIC-M13 (-21)	53	–	0.15
Panel 6	otgfa0_408293	54	0.06	0.15
	otgfa0_0139_G12	54	0.06	0.15
	ucdCg-200	54	0.06	0.15
	NED-M13 (-21)	53	–	0.15

Ta annealing temperature

from a minimum of nine for Crgi3 and otgfa0_0139_G12 to a maximum of 16 for ucdCg-146 and ucdCg-156 with an average of 11.8 (Table 2). The polymorphism information content (*PIC*) and expected heterozygosity (*He*) ranged between 0.619 (uscCgi-210) and 0.917 (ucdCg-156), between 0.647 (uscCgi-210) and 0.924 (ucdCg-156), respectively. All loci showed high informative (*PIC* > 0.5) (Botstein et al. 1980). Allele frequency distributions were highly uneven. Some loci displayed rather homogeneous allelic frequencies (such as ucdcg-146 and ucdcg-156), while some others distributed unevenly (such as ucdcg-199 and uscCgi-210). These markers were an assortment of di-, tri-, and tetranucleotide repeats. Most of them exhibited nearly regular allelic series in conformity with the reported repetition motifs except ucdcg-117, ucdcg-199 and uscCgi-210 (Fig. 1).

The 50 SNPs were divided into four groups according to the *PIC* values from high to low, including 12, 13, 13, and 12 SNPs (Table 3). There were two alleles in each SNP locus. The *PIC* ranged from 0.202 to 0.375 and the *He* ranged from 0.229 to 0.501. Only three loci (CgSNP35, CgSNP149, and CgSNP176) exhibited slightly informative (*PIC* < 0.25), the remaining were middle informative (0.25 < *PIC* < 0.5) (Botstein et al. 1980).

These two types of markers showed different allele frequency distributions (Fig. 2). For the microsatellites, most of the alleles were found to be less than 10% frequency, while all of the SNPs showed an overall frequency of greater than 10%.

Significant departures from Hardy–Weinberg (HW) proportions due to heterozygote deficiency were detected both in the microsatellites (ucdCg-117, ucdCg-156, ucdCg-199, otgfa0_0139_G12 and ucdCg-200) and the SNPs (CgSNP188, CgSNP180, CgSNP65, CgSNP52, CgSNP27, CgSNP203, CgSNP34, CgSNP147, CgSNP187, CgSNP197, CgSNP148 and CgSNP162).

Parentage analysis

The CERVUS simulations showed that the power of parental assignment with 9 microsatellites (three multiplex panels) or 38 SNPs (three groups) could be 100% (Table 4). The actual parentage analysis demonstrated that all of the offspring were correctly allocated to a pair of

Table 2 Four multiplex PCR sets and genetic parameters of 12 microsatellite loci in *C. gigas*

Multiplex sets	Locus	*N*	*Na*	*PIC*	*Ho*	*He*	EXCL1	EXCL2	HWE	F (null)
Panel 1	ucdCg-120	384	10	0.722	0.880	0.758	0.366	0.546	NS	−0.092
	ucdCg-198	384	11	0.821	0.820	0.841	0.518	0.685	NS	0.010
	ucdCg-117	384	15	0.897	0.779	0.906	0.674	0.806	*	0.069
Panel 2	Crgi3	384	9	0.655	0.875	0.683	0.294	0.481	NS	−0.254
	ucdCg-146	382	16	0.888	0.872	0.898	0.657	0.794	NS	0.013
	uscCgi-210	383	10	0.619	0.603	0.647	0.259	0.445	NS	0.010
Panel 3	ucdCg-170	383	12	0.829	0.825	0.845	0.539	0.704	NS	0.002
	ucdCg-156	383	16	0.917	0.574	0.924	0.728	0.843	**	0.183
	ucdCg-199	383	12	0.655	0.535	0.670	0.301	0.496	**	0.100
Panel 6	otgfa0_408293	383	10	0.820	0.903	0.840	0.515	0.683	NS	−0.044
	otgfa0_0139_G12	384	9	0.745	0.633	0.772	0.403	0.584	*	0.075
	ucdCg-200	383	12	0.840	0.799	0.856	0.553	0.715	*	0.026

N number of individuals, *Na* number of alleles, *PIC* polymorphic information content, *Ho* observed heterozygosity, *He* expected heterozygosity, EXCL1 and EXCL2, exclusion probabilities when no parent is known and when one parent is known, respectively, *HWE* significance of Hardy–Weinberg equilibrium test with Bonferroni correction, *NS* non-significant deviation from Hardy–Weinberg equilibrium, * and **, significant deviation from Hardy–Weinberg equilibrium, *F* (null) frequency of null allele calculated by Micro-checker 2.2.3

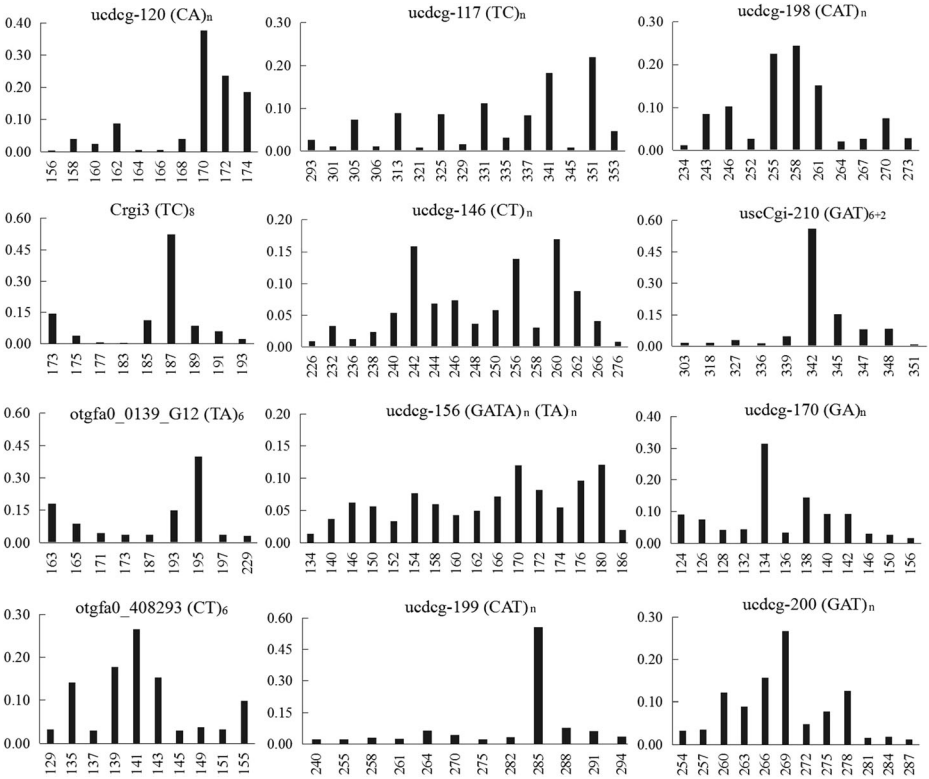

Fig. 1 Allele frequency distributions of the 12 microsatellites in the 384 individuals of *Crassostrea gigas*

parents based on three microsatellites multiplex panels or 50 SNPs (four groups) (Table 4). For microsatellites, EXCL1 values ranged from 0.259 (uscCgi-210) to 0.728 (ucdcg-156) and EXCL2 from 0.445 (uscCgi-210) to 0.843 (ucdcg-156) (Table 2). The combined exclusion probability values, both EXCL1 and EXCL2, were higher than 0.98 when two multiplex PCR panels were used (Fig. 3). Moreover, the combined exclusion probability could reach 1 for EXCL2 when three or four panels were applied. For SNPs, EXCL1 ranged from 0.026 (CgSNP176) to 0.125 (CgSNP96, CgSNP144 and CgSNP192) and EXCL2 from 0.101 (CgSNP176) to 0.278 (CgSNP224) (Table 3). The combined exclusion potentials with no parent known (EXCL1) and one parent known (EXCL2) for the 50 SNP loci were 0.9909 and 0.9999, respectively (Fig. 4). The exclusion power (P_u) was 0.9922 when three microsatellites multiplex panels or 50 SNPs were applied, while P_u reached 0.9995 when all four microsatellite multiplex panels were used (Fig. 5).

Discussion

A variety of SNP genotyping methods have been developed in aquaculture animals, including tetra-primer ARMS-PCR, TaqMan, melting temperature (Tm)-shift, and Illumina Golden Gate technology (Bai et al. 2009; Hansen et al. 2011; Liu et al. 2011; Lapegue et al. 2014). High-resolution melting (HRM) is a very attractive, fast, and cost-effective SNP genotyping technology because it is simple, nondestructive, and amendable to high-throughput on 96- or 384-well plates (Wittwer 2009). For this technique, only polymerase chain reaction (PCR), a

Table 3 Four SNP groups and genetic parameters of 50 SNP loci in *C. gigas*

Groups	SNP Name	*N*	*Na*	*PIC*	*Ho*	*He*	EXCL1	EXCL2	HWE	F (Null)
Group 1	CgSNP96	380	2	0.375	0.692	0.500	0.125	0.187	NS	−0.161
	CgSNP144	378	2	0.375	0.712	0.501	0.125	0.187	NS	−0.174
	CgSNP192	384	2	0.375	0.706	0.501	0.125	0.187	NS	−0.170
	CgSNP33	380	2	0.374	0.063	0.498	0.124	0.187	NS	−0.007
	CgSNP85	380	2	0.374	0.845	0.498	0.124	0.187	NS	−0.259
	CgSNP186	384	2	0.374	0.466	0.499	0.124	0.187	NS	0.003
	CgSNP230	384	2	0.373	0.430	0.496	0.123	0.186	NS	0.007
	CgSNP131	382	2	0.371	0.654	0.492	0.121	0.185	NS	−0.142
	CgSNP224	384	2	0.370	0.352	0.491	0.120	0.278	NS	−0.165
	CgSNP158	380	2	0.369	0.511	0.489	0.119	0.185	NS	−0.022
	CgSNP188	384	2	0.367	0.352	0.484	0.117	0.183	**	0.158
	CgSNP04	373	2	0.365	0.558	0.481	0.116	0.183	NS	−0.073
Group 2	CgSNP14	378	2	0.365	0.429	0.480	0.115	0.182	NS	0.016
	CgSNP157	381	2	0.364	0.459	0.480	0.115	0.182	NS	0.002
	CgSNP180	379	2	0.364	0.375	0.480	0.115	0.182	*	0.122
	CgSNP150	381	2	0.363	0.496	0.476	0.113	0.181	NS	−0.021
	CgSNP91	367	2	0.362	0.695	0.476	0.113	0.181	NS	−0.187
	CgSNP225	383	2	0.361	0.514	0.474	0.112	0.181	NS	−0.041
	CgSNP65	382	2	0.356	0.382	0.464	0.107	0.178	*	0.095
	CgSNP126	368	2	0.356	0.554	0.464	0.107	0.178	NS	−0.089
	CgSNP52	383	2	0.355	0.334	0.462	0.106	0.177	*	0.059
	CgSNP28	383	2	0.352	0.418	0.457	0.104	0.176	NS	0.044
	CgSNP27	371	2	0.345	0.305	0.443	0.098	0.172	**	0.184
	CgSNP203	384	2	0.342	0.208	0.438	0.096	0.171	*	0.054
	CgSNP171	383	2	0.334	0.405	0.424	0.090	0.167	NS	0.012
Group 3	CgSNP164	384	2	0.332	0.396	0.420	0.088	0.166	NS	0.009
	CgSNP34	382	2	0.327	0.120	0.413	0.085	0.164	*	0.047
	CgSNP206	381	2	0.322	0.346	0.404	0.082	0.161	NS	0.006
	CgSNP147	379	2	0.321	0.224	0.402	0.081	0.161	*	0.083
	CgSNP167	382	2	0.321	0.254	0.403	0.081	0.161	NS	0.012
	CgSNP155	379	2	0.320	0.050	0.400	0.080	0.160	NS	0.005
	CgSNP222	384	2	0.319	0.352	0.399	0.079	0.160	NS	0.002
	CgSNP187	376	2	0.317	0.250	0.396	0.078	0.159	*	0.125
	CgSNP209	377	2	0.312	0.475	0.387	0.075	0.156	NS	−0.103
	CgSNP197	380	2	0.308	0.132	0.381	0.072	0.154	*	0.085
	CgSNP148	370	2	0.307	0.232	0.378	0.071	0.154	*	0.037
	CgSNP07	380	2	0.300	0.089	0.367	0.067	0.150	NS	0.007
	CgSNP23	382	2	0.297	0.435	0.363	0.066	0.149	NS	−0.089
Group 4	CgSNP223	377	2	0.295	0.279	0.360	0.065	0.147	NS	−0.126
	CgSNP220	382	2	0.294	0.421	0.359	0.064	0.147	NS	−0.080
	CgSNP111	373	2	0.283	0.324	0.342	0.058	0.142	NS	−0.025
	CgSNP183	380	2	0.283	0.400	0.342	0.058	0.142	NS	−0.079
	CgSNP232	383	2	0.283	0.420	0.341	0.058	0.141	NS	−0.104
	CgSNP162	382	2	0.274	0.199	0.329	0.054	0.137	*	0.045
	CgSNP194	383	2	0.271	0.405	0.323	0.052	0.135	NS	−0.111
	CgSNP36	353	2	0.266	0.348	0.317	0.050	0.133	NS	−0.048
	CgSNP41	379	2	0.264	0.361	0.313	0.049	0.132	NS	−0.072
	CgSNP35	383	2	0.249	0.355	0.292	0.043	0.125	NS	−0.095
	CgSNP149	375	2	0.227	0.277	0.262	0.034	0.114	NS	−0.029
	CgSNP176	380	2	0.202	0.216	0.229	0.026	0.101	NS	0.028

All SNPs displayed in decreasing order of *PIC*

N number of individuals, *Na* number of alleles, *PIC* polymorphic information content, *Ho* observed heterozygosity, *He* expected heterozygosity, EXCL1 and EXCL2 exclusion probabilities when no parent is known and when one parent is known, respectively, *HWE* significance of Hardy–Weinberg equilibrium test with Bonferroni correction, *NS* non-significant deviation from Hardy–Weinberg equilibrium, * and **, significant deviation from Hardy–Weinberg equilibrium, *F* (null) frequency of null allele calculated by Micro-checker 2.2.3

DNA dye, and melting instrumentation are required (Zhong et al. 2013). In this study, low genetic diversity of SNPs was revealed as the polymorphism information content (*PIC*) per locus ranged from 0.202 to 0.375 with an average of 0.328. In contrast, high genetic diversity in *C. gigas* was presented using microsatellites with an average *PIC* value of 0.784. This discrepancy of *PIC* may be primarily due to the traits of markers. SNPs are bi-allelic while microsatellites are multiple alleles, thus SNPs shows lower information content than microsatellites (Aitken et al. 2004). Because of lower *PIC*, more SNPs are required than microsatellites to obtain the same power of exclusion.

All SNPs examined here were found to have an average *He* value of 0.417, which was significantly lower than the comparable value for the microsatellites (0.803). Vignal et al. (2002) predicted that SNP markers are mainly bi-allelic, such that a maximum expected heterozygosity value of 0.50 can be expected for a given SNP locus.

We found that most of the SNP alleles are at an intermediate or high frequency (Fig. 2). The differences in the frequency spectrum between microsatellites and SNPs may lead to the difference in the information content of the two types of marker (Kong et al. 2014).

The presence of null alleles was detected for pedigree analysis using microsatellite markers, which was in accordance with previous results in *C. gigas* (McGoldrick et al. 2000; Hedgecock et al. 2004; Li et al. 2009). This phenomenon could be attributed to indels or mutations on one or both of the primer-binding sites (Dakin and Avise 2004; Lemer et al. 2011; Mcinerney et al. 2011). High null allele frequency was also detected at four SNPs (CgSNP188, CgSNP180, CgSNP27, and CgSNP187). This uncommon problem was reported in Mediterranean mussel (*Mytilus galloprovincialis*) (Pino-Querido et al. 2015) as well. The same explanations presented above for high allele frequency at microsatellites may also be responsible for this observation at SNPs (Pino-Querido et al. 2015). In this study, some loci significantly deviated from Hardy–Weinberg (HW) proportions. These deviations resulted from difference between expected and observed heterozygosity. Null alleles are likely causes for heterozygote deficiency in HWE tests.

The level of unique assignment primarily depends on the assignment power of the marker set used, which is decided by the exclusion probabilities of the markers. Microsatellites were

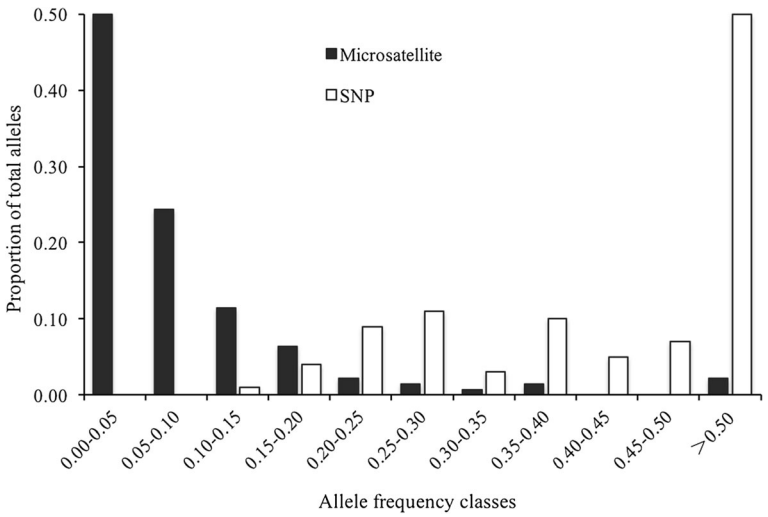

Fig. 2 Allele frequency distributions for microsatellites and SNPs in *C. gigas*

Table 4 Cumulative assignment success rates of simulated and real genotype data at the 95% confidence level. Each marker set was added in decreasing order of average polymorphic information content (*PIC*)

Type of markers	Cumulative assignment success (%)	
	Simulation data	Real data
Microsatellite multiplex panels		
Panel 1	0	0
Panel (1 + 6)	96	13
Panel (1 + 6 + 3)	100	100
Panel (1 + 6 + 3 + 2)	100	100
SNP groups		
Group 1	0	0
Group (1 + 2)	34	1
Group (1 + 2 + 3)	100	98
Group (1 + 2 + 3 + 4)	100	100

individually more powerful to exclude a false parent than SNPs since EXCL1 and EXCL2 for each microsatellite were higher. The combined exclusion probabilities for both EXCL1 and EXCL2 as well as P_u were calculated by adding marker sets from highest average *PIC* value to lowest sequentially. The combined exclusion power for EXCL2 reached 1 when 3 microsatellites multiplex PCRs or more than 50 SNPs were used. Meanwhile, P_u was 0.9922 when three microsatellites multiplex panels were applied. With 50 SNPs, the same exclusion power could also be reached. In our study, the cost of microsatellites multiplex PCR analysis was estimated to be $0.18 per sample per locus, while the cost of SNP genotyping was $0.30. With increasing loci required to obtain a higher exclusion power, the genotyping costs would also increase. In practice, optimal investment in parentage assignment is a balance between the reduction of investment and operational costs needed for the separate family rearing and the cost of genotyping (Vandeputte and Haffray 2014). An increase in the number of selected loci resulted in an initial rapid rise in the exclusion probability value followed by slower relative growth until the maximum (100%) (Yu et al. 2015). To reduce the cost of pedigree analysis, we should explore the minimum and optimal criteria for microsatellites and SNPs required to attain a prerequisite statistical power in estimating pedigrees.

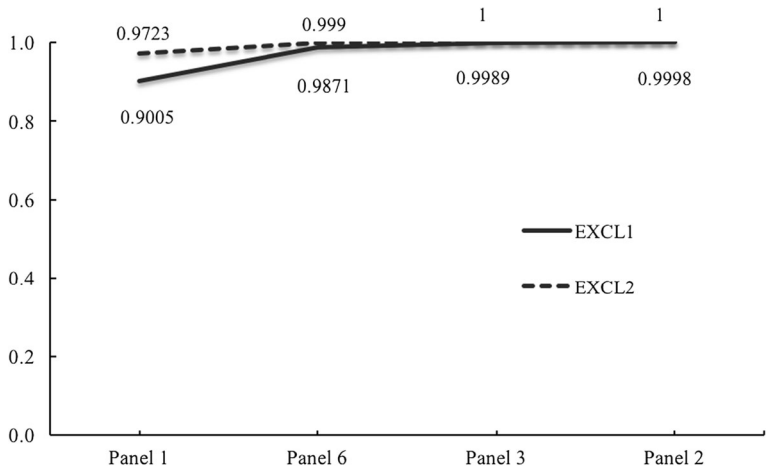

Fig. 3 Combined exclusion probabilities of the four microsatellite multiplex panels analyzed for EXCL1 and EXCL2. Each panel was added in decreasing order of average polymorphic information content (*PIC*)

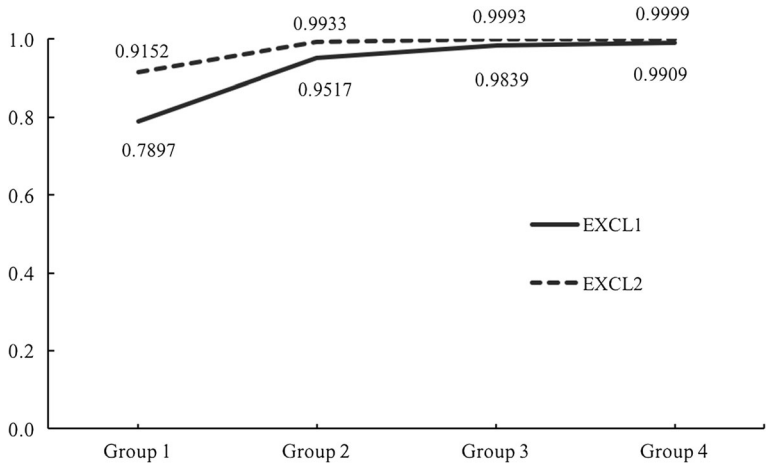

Fig. 4 Combined exclusion probabilities of the four SNP groups analyzed for EXCL1 and EXCL2. Each group was added in decreasing order of average polymorphic information content (*PIC*)

The combined exclusion probabilities of 38 SNPs for EXCL2 came to 0.9993, which is nearly equal to the index of two microsatellites multiplex PCR panels. That was to say generally about six SNPs were needed to obtain an equivalent exclusion power for pedigree analysis with a microsatellite in *C. gigas*. Similarly, Glaubitz et al. (2003) estimated about six SNPs gave the same assignment power as 1 microsatellite. Simulations showed that about 16 bi-allelic SNPs would provide the same assignment power as three 10 allelic microsatellites (Wang 2006). In an inbred Augus herd, generally 2–3 SNPs per microsatellite were needed to obtain an equivalent exclusion power value (Fernández et al. 2013). Sellars et al. (2014) found the custom-made SNP panels were significantly cheaper to run while also providing faster turnaround time on genotyping results than multiplexed microsatellite panel. In sockeye salmon, Hauser et al. (2011) showed generally higher and more accurate parentage assignment success with 80 SNPs than 11 microsatellites. However, it is difficult to decide which type of marker is most cost-

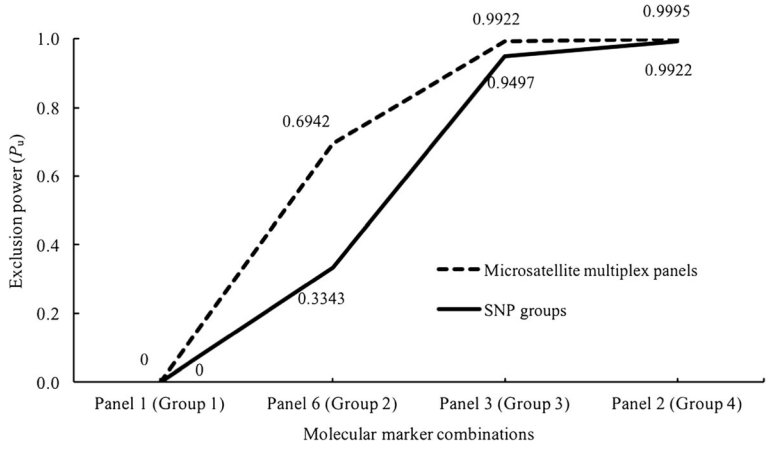

Fig. 5 Exclusion power (P_u) of four microsatellite multiplex panels and four SNP groups. Each marker combination was added in decreasing order of average polymorphic information content (*PIC*)

effective. Because SNPs remain more expensive due to the number required, but technology is rapidly evolving for SNPs and not for microsatellites for the moment (Vandeputte and Haffray 2014). The bright future of DNA markers applied for pedigree analysis will depend on the development of techniques. In fact, a few recent microsatellite studies relied on very large multiplexes (Guichoux et al. 2011). In *C. gigas*, the largest multiplex panel simply contained five loci (Miller et al. 2012). Therefore, it is necessary to further develop multiplex PCRs to coamplify at least nine microsatellites in one reaction in *C. gigas*. Simultaneous PCR amplification of several PCR products has been achieved in the case of SNPs (Gabriel et al. 2009; Vera et al. 2010; Pino-Querido et al. 2015). Helyar et al. (2011) also concluded multiple SNPs would be desirable to create minimum panels with maximum power for individual assignment. Therefore, development of SNP multiplex PCRs is worthy of consideration in *C. gigas* to make the pedigree analysis more cost-effective.

Acknowledgements This study was supported by the grants from the National Natural Science Foundation of China (31372524), Shandong Seed Project, Shandong Province (2014GHY115002, 2016ZDJS06A06), and Qingdao National Laboratory for Marine Science and Technology (2015ASKJ02).

References

Aitken N, Smith S, Schwarz C, Morin PA (2004) Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. Mol Ecol 13:1423–1431

An HS, Lee JW, Kim HC, Myeong JI (2011) Genetic characterization of five hatchery populations of the Pacific abalone (*Haliotis discus hannai*) using microsatellite markers. Int J Mol Sci 12:4836–4849

An HS, Lee JW, Kim WJ, Lim HJ, Kim EM, Byun SG, Hur YB, Park JY, Myeong JI, An CM (2013) Comparative genetic diversity of wild and hatchery-produced Pacific oyster (*Crassostrea gigas*) populations in Korea using multiplex PCR assays with nine polymorphic microsatellite markers. Genes Genom 35:805–815

Anderson EC, Garza JC (2006) The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172:2567–2582

Bai J, Li Q, Kong LF, Li RH (2009) Characterization of 20 single nucleotide polymorphism markers in the Pacific oyster (*Crassostrea gigas*). Anim Genet 40:1004

Bester AE, Roodt-Wilding R, Whitaker HA (2008) Discovery and evaluation of single nucleotide polymorphisms (SNPs) for *Haliotis midae*: a targeted EST approach. Anim Genet 39:321–324

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314

Castro J, Pino A, Hermida M, Bouza C, Chavarrías D, Merino P, Sánchez L, Martínez P (2007) A microsatellite marker tool for parentage assessment in gilthead seabream (*Sparus aurata*). Aquaculture 272:S210–S216

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. Heredity 93:504–509

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet 16:551–558

FAO (2014) The state of the world fisheries and aquaculture 2014. FAO, Rome

Fernández ME, Goszczynski DE, Lirón JP, Villegas-Castagnasso EE, Carino MH, Ripoli MV, Rogberg-Muñoz A, Posik DM, Peral-García P, Giovambattista G (2013) Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. Genet Mol Biol 36:185–191

Fu J, Shen Y, Xu X, Chen Y, Li D, Li J (2013) Multiplex microsatellite PCR sets for parentage assignment of grass carp (*Ctenopharyngodon idella*). Aquacult Int 21:1195–1207

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. Curr Protoc Hum Genet 2:1–18

Glaubitz JC, Rhodes OE, DeWoody JA (2003) Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Mol Ecol 12:1039–1047

Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591–611

Hansen MHH, Young S, Jørgensen HBH, Pascal C, Henryon M, Seeb J (2011) Assembling a dual purpose TaqMan-based panel of single-nucleotide polymorphism markers in rainbow trout and steelhead (*Oncorhynchus mykiss*) for association mapping and population genetics analysis. Mol Ecol Resour 11: 67–70

Hauser L, Baird M, Hilborn RAY, Seeb LW, Seeb JE (2011) An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. Mol Ecol Resour 11:150–161

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 23:379–386

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol Ecol Resour 11:123–136

Herbinger CM, Doyle RW, Pitman ER, Paquet D, Mesa KA, Morris DB, Wright JM, Cook D (1995) DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. Aquaculture 137:245–256

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106

Kong L, Bai J, Li Q (2014) Comparative assessment of genomic SSR, EST–SSR and EST–SNP markers for evaluation of the genetic diversity of wild and cultured Pacific oyster, *Crassostrea gigas* Thunberg. Aquaculture 420:S85–S91

Lallias D, Taris N, Boudry P, Bonhomme F, Lapegue S (2010) Variance in the reproductive success of flat oyster *Ostrea edulis* L. assessed by parentage analyses in natural and experimental conditions. Genet Res 92:175–187

Lapegue S, Harrang E, Heurtebise S, Flahauw E, Donnadieu C, Gayral P, Ballenghien M, Genestout L, Barbotte L, Mahla R, Haffray P, Klopp C (2014) Development of SNP-genotyping arrays in two shellfish species. Mol Ecol Resour 14:820–830

Lemer S, Rochel E, Planes S (2011) Correction method for null alleles in species with variable microsatellite flanking regions, a case study of the black-lipped pearl oyster *Pinctada margaritifera*. J Hered 102:243–246

Li L, Guo X, Zhang G (2009) Inheritance of 15 microsatellites in the Pacific oyster *Crassostrea gigas*: segregation and null allele identification for linkage analysis. Chin J Oceanol Limnol 27:74–79

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

Li R, Li Q, Cornette F, Dégremont L, Lapègue S (2010) Development of four EST-SSR multiplex PCRs in the Pacific oyster (*Crassostrea gigas*) and their validation in parentage assignment. Aquaculture 310:234–239

Liu T, Li Q, Song J, Yu H (2016) Development of genomic microsatellite multiplex PCR using dye-labeled universal primer and its validation in pedigree analysis of Pacific oyster (*Crassostrea gigas*). J Ocean U China, in press.

Liu WD, Li HJ, Bao XB, He CB, Li WJ, Shan ZG (2011) The first set of EST-derived single nucleotide polymorphism markers for Japanese scallop, *Patinopecten yessoensis*. J World Aquacult Soc 42:456–461

McGoldrick DJ, 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

McInerney CE, Allcock AL, Johnson MP, Bailie DA, Prodöhl PA (2011) Comparative genomic analysis reveals species-dependent complexities that explain difficulties with microsatellite marker development in molluscs. Heredity 106:78–87

Miller PA, Elliott NG, Koutoulis A, Kube PD, Vaillancourt RE (2012) Genetic diversity of cultured, naturalized, and native Pacific oysters, *Crassostrea gigas*, determined from multiplexed microsatellite markers. J Shellfish Res 31:611–617

Pemberton JM (2008) Wild pedigrees: the way forward. P Roy Soc B-Biol Sci 275:613–621

Pino-Querido A, Álvarez-Castro JM, Vera M, Pardo BG, Fuentes J, Martínez P (2015) A molecular tool for parentage analysis in the Mediterranean mussel (*Mytilus galloprovincialis*). Aquac Res 46:1721–1735

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8:103–106

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

Sellars MJ, Dierens L, McWilliam S, Little B, Murphy B, Coman GJ, Barendse W, Henshall J (2014) Comparison of microsatellite and SNP DNA markers for pedigree assignment in black tiger shrimp, *Penaeus monodon*. Aquac Res 45:417–426

Taris N, Baron S, Sharbel T, Sauvage C, Boudry P (2005) A combined microsatellite multiplexing and boiling DNA extraction method for high-throughput parentage analyses in the Pacific oyster (*Crassostrea gigas*). Aquac Res 36:516–518

Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Vandeputte M (2012) An accurate formula to calculate exclusion power of marker sets in parentage assignment. Genet Sel Evol 44:36

Vandeputte M, Haffray P (2014) Parentage assignment with genomic markers: a major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. Front Genet 5:432

Vandeputte M, Rossignol MN, Pincent C (2011) From theory to practice: empirical evaluation of the assignment power of marker sets for pedigree analysis in fish breeding. Aquaculture 314:80–86

Vera M, Pardo BG, Pino-Querido A, Álvarez-Dios JA, Fuentes J, Martínez P (2010) Characterization of single-nucleotide polymorphism markers in the Mediterranean mussel, *Mytilus galloprovincialis*. Aquac Res 41: e568–e575

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 34:275–305

Wang J (2006) Informativeness of genetic markers for pairwise relationship and relatedness inference. Theor Popul Biol 70:300–321

Wittwer CT (2009) High-resolution DNA melting analysis: advancements and limitations. Hum Mutat 30:857–859

Yamtich J, Voigt ML, Li G, Hedgecock D (2005) Eight microsatellite loci for the Pacific oyster *Crassostrea gigas*. Anim Genet 36:524–526

Yu GC, Tang QZ, Long KR, Che TD, Li MZ, Shuai SR (2015) Effectiveness of microsatellite and single nucleotide polymorphism markers for parentage analysis in European domestic pigs. Genet Mol Res 14: 1362–1370

Zhang J, Ma W, Wang W, Gui JF, Mei J (2016) Parentage determination of yellow catfish (*Pelteobagrus fulvidraco*) based on microsatellite DNA markers. Aquacult Int 24:567–576

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