RESEARCH NOTE

Parentage identification of *Odontobutis potamophlia* based on microsatellite DNA markers

PEIPEI WANG[a,b#], CHENG ZHAO[a#],SHUQIAO CHEN[b], SHAOWU YIN[a,*], QINTAO WANG[a], ZECHENG LI[a], GUOQIN ZHOU[b]

[a]College of Life Sciences, Key Laboratory of Biodiversity and Biotechnology of Jiangsu Province, Nanjing Normal University, Nanjing, Jiangsu 210023, China

[b]Institute of Fisheries Science of Nanjing city, Nanjing, Jiangsu 210036, China

[#]These authors contributed equally to this work.

*Corresponding author: yinshaowu@163.com

Short tittle: Parentage identification of *Odontobutis potamophlia* using microsatellite

Keywords: *Odontobutis potamophila*; microsatellite DNA; parentage identification;

Microsatellite loci were used for parentage identification of *Odontobutis potamophila* in five full-sib families. The combined exclusion probability of the first (E-1P) and the second parent (E-2P) revealed an obvious increased with the increase of number of microsatellite loci. The combined exclusion probability based on allele frequency suggested that at least eight microsatellite loci were needed for the identification of the 150 individuals from five families supported by the genetic distance analysis of individuals of these families. The double-blind test results indicated that the candidate individuals could find their correct parents through these loci thus eight microsatellite markers can be used for pedigree analysis of *O. potamophlia* in breeding industry as well as future selection studies.

Introduction

Odontobutis potamophila belongs to the family Odontobutidae, an important commercial fish species widely distributed in Southeast China deeply favored by people for its delicious taste, rich nutrition and therefore, high economic value (Li and

Liu 2014). However, environmental pollution, overfishing, and growing market demand, *O. potamophila* has undergone a severe reduction in population size and distribution (Zhu et al. 2014). Therefore, it is urgent to conduct genetic-based selective breeding of *O. potamophila.* The selective breeding is an effective strategy to improve variety selection in aquaculture. In breeding practices, to save space and capital, different lineages have frequently been bred together. As a consequence, it may result in inbreeding, thus reducing genetic diversity of offspring, therefore, limiting the potential for genetic gain from artificial selection and ultimately reduce the production performance. For the sake of avoiding the phenomenon of inbreeding and improving variety selection, it is extremely important to establish reliable pedigree information and identify the genetic relationship of different lineages to design breeding programs to enhance desirable traits.

In the field of ecology, microsatellite loci were commonly used for the measurement of genetic variability within and between populations (Hartl and Clark 1989; Slatkin 1995). The relative ease of development and high allelic diversity make it ideal for individual genotyping in assisting pedigree identification (Guichoux et al. 2011b), identifying unique individuals (Lukacs and Burnham 2005), and inferring phylogenetic or genealogical relationships in different groups of species (Blouin 2003). Microsatellite-based parentage identification technology has been widely used in determining genetic relationships in aquaculture, such as yellow catfish (*Pelteobagrus fulvidraco*) (Zhang et al. 2015), large yellow croaker (*Larimichthys crocea*) (Liu et al. 2013) and coral reef fish (*Lutjanus carponotatus*) (Hugo et al. 2014). All the offspring from different groups and families can be stocked together and assigned to their original families and their potential parental genotypes retrospectively by using DNA markers such as microsatellites. In this way, there is no need to mark the progeny before stocking into ponds and known and unknown parent progeny relationships can be used to evaluate the identification capability of microsatellite DNA markers.

Material and methods

Sample collection and pedigree establishment

The broodstock (1 year old) collected randomly from Xinan River in Dangtu of Anhui Province, Jiande of Zhejiang Province and Taihu of Suzhou of Jiangsu Province were well developed, healthy and free from injury. The mean length and body weight of each individual was 12.79 ± 0.33 cm, and 32.25 ± 2.48 g, respectively. Twenty-seven full-sib families (three replicates of each combination) were established and offspring from each family was recorded. Five full-sib families were selected randomly and denoted as the undetermined family. Progenies from each of these families were grown separately till they reach 5cm in length. From each family 30 individuals were selected and tail fin collected and stored in ethanol for subsequent experiments. Forty individuals from the rest offspring of five selected families (eight samples of each family) were selected, and their tail fin tissues were collected and stored in ethanol. All experiments were performed according to the Guidelines for the Care and Use of Laboratory Animals in China. This study was also approved by the Ethics Committee of Experimental Animals at Nanjing Normal University.

DNA extraction and genotyping of microsatellite DNA markers

Genomic DNA was extracted from tail fin tissue using cell/tissue genomic DNA extraction kit (centrifugal column type, Generay Biotech, Shanghai). Fifty-six markers of *O. potamophila* have been developed in our laboratory (Wang et al. 2017). We selected eight SSRs among these fifty-six developed microsatellite loci, which show relative higher polymorphism (OPRM 87, OPRM 114, OPRM 115, OPRM 117, OPRM 123, OPRM 125, OPRM 197, and OPRM 200) to genotype the offspring from five families and their original parents (Table S1). A total of 10 μL of PCR mixture was prepared by 50 ng template DNA, 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 10× PCR buffer, 1 U Taq DNA polymerase, 0.04 μM M13-tailed forward primer, 0.16 μM reverse primer and 0.16 μM fluorescence-labeled M13 primer. PCR was conducted under

following conditions: initial denaturation at 95 °C for 5 min, 30 cycles with the denaturation at 94 °C for 30 s, annealing at various temperatures (Table S1) for 30 s and extension at 72 °C for 45 s, followed by 8 cycles with the denaturation at 94 °C for 30 s, annealing at 53 °C for 45 s and extension at 72 °C for 45 s, and a final extension at 72 °C for 10 min. Subsequently, ABI 3500xl sequencer was used to analyze the PCR products and sizes were determined with GENEMAPPER version 4.1.

Statistical analysis

The genetic diversity parameters of eight SSR loci including the number of alleles (Na), observed heterozygosity (Ho), expected heterozygosity (He), polymorphism information content (PIC), frequency of null alleles (Null freq) and the exclusion probability without known parents (E-1P), and the exclusion probability with genotype of one parent known (E-2P) were estimated by CERVUS 3.0 (Kalinowski et al. 2007). According to the genetic distance between each individual in the family, the NTSYS software was used for clustering analysis, which was used for analyzing the relationship of different pedigrees. The selected five families were denoted as A, B, C, D and E, respectively. Transforming genotype data of each family into POPGENE format through POPGENE1.32 software is required, and the genetic distance analysis with POPGENE1.32 software for the selected individuals was conducted for the evaluation of genetic consistency. The EXCEL software (2013) was used to transpose the genetic consistency tables into the TXT format through NTSYS software, and then used for clustering analysis for individuals by using NTSYS software according to the degree of genetic consistency.

Results

Genetic diversity of microsatellite DNA markers

190 individuals in total from five full-sib families of *O. potamophlia* were genotyped by eight SSR markers for genetic diversity parameters. The genotype of individuals, allele range, and null allele frequency (Null freq) was shown in Table S2.

The number of alleles (Na), observed heterozygosity (Ho), expected heterozygosity (He) and the polymorphism information content (PIC) ranged from 4 to 9, 1.000 to 1.000, 0.602 to 0.833 and 0.521 to 0.824, respectively (Table S3). All eight loci were highly polymorphic, with the PIC larger than 0.50 (Botstein et al. 1980).

Computed simulation

The E-1P and E-2P of each locus ranged from 0.187 to 0.554 and from 0.320 to 0.716 (Table 1). The locus OPRM 87 exhibited the highest value of PIC, E-1P and E-2P (0.824, 0.554 and 0.716, respectively), and revealed the lowest value of PIC, E-1P and E-2P (0.85, 0.575 and 0.732, respectively). As shown in Figure 1, the combined exclusion probability values, the E-1P was 97 % and the E-2P was higher than 99 %, when eight SSRs were selected and used. All results suggested that the eight SSRs we selected were sufficient for parentage analysis. What is more, in parentage analysis the more loci were selected, the higher accuracy was achieved.

Pedigree identification and double blind test

All experimental samples were clustered into the same branch, when parents and offspring belonged to the same family. Our study indicated that all parents and offspring from five families could distribute into the same branch appropriately (Figure S1). As a result, the selected microsatellite DNA markers can distinguish the different families. For the double-blind test, a total of 40 offspring from five families (eight individuals of each family) were randomly selected. All progenies were allocated to five branches separately which indicated by UPAGE dendrogram (Figure S2). The accuracy of clustering reached up to 100% and all individuals from the same family were clustered together.

Discussion

In aquaculture, selective breeding program is important and it is necessary to select suitable markers for maintaining correct pedigree relationships. Many conventional methods, such as, physical tags and electronic tags have several

shortcomings in breeding program (Jerry et al. 2004). Biological tags have a potential power to resolve the limitations of conventional methods. In present study, we demonstrated that both parents and offsprings could be determined in the absence of physical tags and pedigree information in a mixed population. The application of microsatellite DNA markers to identify and determine the genetic relationship has gained recognition and attention in aquaculture because microsatellite DNA marker with rich polymorphism is a simple way to determine the heredity. For example, four microsatellite DNA markers are used for parentage analysis in the 100 breeding individuals of rainbow trout, and 91% of the offspring is assigned to their true parents (Herbinger et al. 1995). In mandarin fish, ten microsatellite DNA markers have been used for parentage determination (Yang et al. 2014), and a total of 180 individuals of yellow catfish in eight full-sib families have also been successfully identified with nine microsatellite DNA markers (Zhang et al. 2015). The study of parentage identification in *O. potamophlia* has not been reported yet. Therefore, we identified and analyzed the genetic relationship of five *O. potamophlia* families through eight microsatellite DNA markers. The development of parentage identification through microsatellite loci depends on the microsatellite polymorphism information, the ability of exclusion and the size of the candidate parents and offspring (Norris et al. 2000). The more candidate parents and offspring could require the more number of microsatellite loci (Vandeputte et al. 2011). As the result, before the markers were employed, it is essential to conduct simulation evaluation, which will benefit for estimating the actual number of loci and the potential performance of these markers. In the present study, all results showed that the increasing number of the selected microsatellite DNA markers and the polymorphism could result in the higher cumulative exclusion rate. When the selected markers reached up to eight, each of progeny was assigned to their correct parents (Figures S1, S2), and the exclusion probability was also increased to 99.8%. In this study, the observed heterozygosity was higher than expected heterozygosity in five families. This might be caused by the

following reasons: (1) the population of *O. potamophila* has a number of homozygous individuals; (2) The families were not mating randomly, which result in excess of heterozygotes and deviate from the Hardy Weinberg equilibrium; (3) The loci in this study have a high mutation rate, which produced many different alleles. (4) Errors might cause by inadequate sampling individuals. (Dakin and Avise 2004; Ellegren 2004; Jerry et al. 2004). The null alleles had slight impact on the accuracy of assignment, because of the frequency of null alleles was relatively low. If enough markers were selected, all offspring can be explicitly allocated to their own parents in practical study. What is more, the assignment success rate was high in our study, which also indicates the validation of our selected microsatellite loci. These eight microsatellite markers in this study will facilitate maintaining pedigree information and family-selected breeding of *O. Potamophlia* in the future.

Acknowledgments

This study was supported by the Jiangsu Province Science and Technology Support Program (BE2013441), Nanjing Science and Technology Plan Projects (201505059), Jiangsu Province Six Talent Peaks of High Level Talents Project (NY-032), Project Foundation of the Academic Program Development of Jiangsu Higher Education Institution (PAPD), and The Innovation of Graduate Student Training Project of Jiangsu Province (SJLX15_0300).

References

Blouin M. S. 2003 DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Tends Ecol Evol.* **18**, 503-511.

Botstein D., White R. L., Skolnick M., Davis R. W. 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*. **32**, 314-331.

Dakin E. E., Avise J. C. 2004 Microsatellite null alleles in parentage analysis. *Heredity*.

93, 504-509.

Ellegren, H. 2004 Microsatellites: simple sequences with complex evolution. *Nat Rev Genet*. **5**, 435-445.

Guichoux E., Lagache L., Wagner S., Chaumeil P., Léger P., Lepais O. C. et al. 2011 Current trends in microsatellite genotyping. *Mol Ecol Resour*. **11**, 591-611.

Hartl D. L., Clark A. G. 1989 Principles of population genetics. Sinauer Associates, Sunderland, Massachussets.

Herbinger C. M., Doyle R. W., Pitman E. R., Paquet D., Mesa K. A., Morris D. B. et al. 1995 DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. *Aquaculture*. **137**, 245-256.

Hugo B. H., Kevin A. F., Geoffrey P. J., Kayan M., Hicham M., Sadhasivam P. et al. 2014 Validation of microsatellite multiplexes for parentage analysis in a coral reef fish (*Lutjanus carponotatus*, Lutjanidae). *Conserv Genet Resour.* **6**, 803-806.

Jerry D. R., Preston N. P., Crocos P. J., Keys S., Meadows J. R. S., Li Y. T. 2004 Parentage determination of Kuruma shrimp *Penaeus* (Marsupenaeus) *japonicus* using microsatellite markers (Bate). *Aquaculture*. **235**, 237-247.

Kalinowski S. T., Taper M. L., Marshall T. C. 2007 Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*. **16**, 1099-1106.

Liu X. D., Zhao G. T., Cai M. Y., Wang Z. Y. 2013 Estimated genetic parameters for growth-related traits in large yellow croaker *Larimichthys crocea* using microsatellites to assign parentage. *J Fish Biol*. **82**, 34-41.

Li Q., Liu Z. 2014 New complete mitochondrial genome of the *Odontobutis potamophila* (Perciformes, Odontobutidae): genome description and phylogenetic performance. *Mitochondr. DNA*. **102**, 915-920.

Lukacs P. M., Burnham K. P. 2005 Review of capture-recapture methods applicable to noninvasive genetic sampling. *Mol Ecol*. **14**, 3909-3919.

Norris A. T., Bradley D. G., Cunningham E. P. 2000 Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers.

Aquaculture. **182**, 73-83.

Slatkin M. 1995 A measure of population subdivision based on microsatellite allele frequencies. *Genetics.* **139**, 457-462.

Vandeputte M., Rossignol M. N., 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.

Wang P. P., Ren M., Chen S. Q., Yin S. W., Zhao C., Zhang H. Y. et.al. 2017 Characterization and development of 56 EST-SSR markers derived from the transcriptome of *Odontobutis potamophila.*. *Genet. Mol. Res*. **16**, 2.

Yang M., Tian C. X., Liang X. F., Lv L. Y., Zheng H. Z., Yuan Y. C. 2014 Parentage determination of mandarin fish *(Siniperca chuatsi*) based on microsatellite DNA markers. *Biochem. Syst. Ecol*. **54**, 285-291.

Zhang J., Ma W. G., Wang W. M., Gui J. F., Mei J. 2015. Parentage determination of yellow catfish (*Pelteobagrus Fulvidraco*) based on microsatellite DNA markers. *Aquacult Int*. **24**, 1-10.

Zhu F., Zhang L. J., Yin S. W., Zhang H. W., Hou X. Y., Hu Y. L. 2014. Genetic diversity and variation in wild populations of dark sleeper (*Odontobutis potamophila*) in China inferred with microsatellite markers. *Biochem. Syst. Ecol*. **57**, 40-47.

Table 1 The exclusion probability of microsatellite DNA markers from computed simulation

Locus	E-1P	E-2P
OPRM 87	0.554	0.716
OPRM 114	0.325	0.510
OPRM 115	0.224	0.375
OPRM 117	0.187	0.320
OPRM 123	0.332	0.509
OPRM 125	0.472	0.646
OPRM 197	0.311	0.486
OPRM 200	0.355	0.533

E-1P, exclusion probability without known parents; E-2P, exclusion probability with genotype of one parent known

Table S1 Characteristics of eight polymorphic microsatellite DNA markers in *O. potamophlia*

Locus	GenBank	Primer sequences (5'-3')	Repeat motif	*Tm*/°C
OPRM 87	KR861493	F: GACGTCACGGTCAATCAATG R: CGACACCTGGACAAACAATG	$(ATA)_5$	50
OPRM 114	KR861497	F: ATCCCGTGTAACTGTACGGC R: GTGAGACCACGAGCACTTCA	$(GAA)_6$	50
OPRM 115	KR861498	F:CTGACAAAAGGCACCAGACA R: AGGCTGGCCTTGTGACTAAA	$(TTG)_6$	50
OPRM 117	KR861499	F: CACAACGCAACCAAAACATC R: GAGACGCTCCGCATAAACTC	$(GAA)_6$	50
OPRM 123	KR861501	F: ACGACAGGCAGCAGAGAAGT R: ATGGGCTTTGGTGTTGTTTC	$(CAG)_6$	50
OPRM 125	KR861503	F: TTGGGAATACGTCAGGCTTC R: CACGAGGAGCAAGTGAATGA	$(TCT)_6$	50
OPRM 197	KT805147	F: TCTAGCAACATTGTGTGCCC R: CCCCTGACCTTTGATATGGA	$(CA)_9$	50
OPRM 200	KT805148	F: GTGTGTGGAGACAGGACACG R: AAAAGTATGCACCAGGGACG	$(TG)_9$	50

Table S2 The genotype of partial individuals in eight microsatellite loci

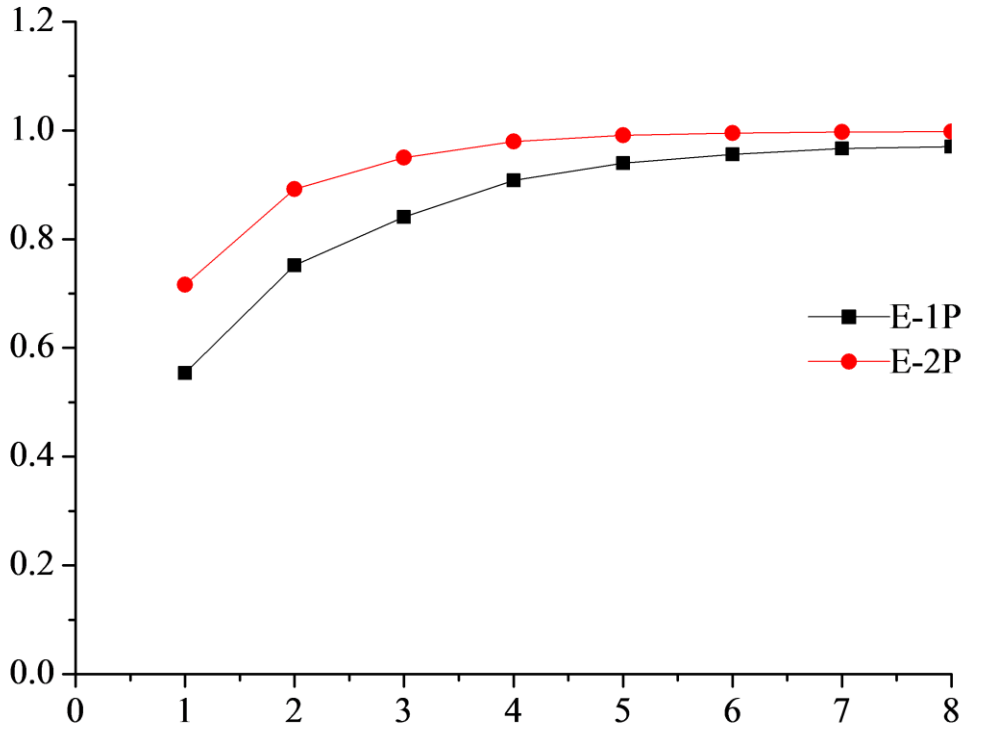
Loci / Families	S87		S114		S115	
ParentA1	277	277	234	234	201	204
ParentA2	280	280	240	240	204	204
A1	277	280	234	240	204	204
A2	277	280	234	240	201	204
A3	277	280	234	240	204	204
A4	277	280	234	240	201	204
A5	277	280	234	240	201	204
A6	277	280	234	240	201	204
A7	277	280	234	240	201	204
A8	277	280	234	240	204	204
A9	277	280	234	240	204	204
A10	277	280	234	240	201	204
ParentB1	266	266	234	234	201	201
ParentB2	277	277	240	240	201	201
B1	266	277	234	240	201	201
B2	266	277	234	240	201	201
B3	266	277	234	240	201	201
B4	266	277	234	240	201	201
B5	266	277	234	240	201	201
B6	266	277	234	240	201	201
B7	266	277	234	240	201	201
B8	266	277	234	240	201	201
B9	266	277	234	240	201	201
B10	266	277	234	240	201	201

Table S3 Genetic diversity of eight microsatellite DNA markers in 150 *O. potamophlia* individuals

Locus	Na	H(o)	H(e)	PIC	Null freq
OPRM 87	9	1.000	0.833	0.824	-0.1328
OPRM 114	7	1.000	0.722	0.687	-0.2128
OPRM 115	5	1.000	0.644	0.574	-0.2453
OPRM 117	4	1.000	0.602	0.521	-0.2822
OPRM 123	5	1.000	0.752	0.703	-0.1485
OPRM 125	8	1.000	0.823	0.795	-0.1076
OPRM 197	5	1.000	0.724	0.676	-0.1819
OPRM 200	6	1.000	0.763	0.718	-0.1437
Mean	6	1.000	0.733	0.687	-

Note: Na, the number of allele; Ho, observed heterozygosity; He, expected heterozygosity; Null freq, frequency of null allele; PIC, polymorphic information content

Figure 1 Combined exclusion probability of the eight polymorphic loci analyzed for no parent known (E-1P) and one parent known (E-2P). The 1-8 loci were the most PIC to the least PIC.



	0	1	2	3	4	5	6	7	8
E-1P		0.554	0.752	0.841	0.908	0.940	0.956	0.967	0.970
E-2P		0.716	0.892	0.950	0.980	0.991	0.995	0.997	0.998

Figure S1 The clustering analysis of the parents and offspring from five families

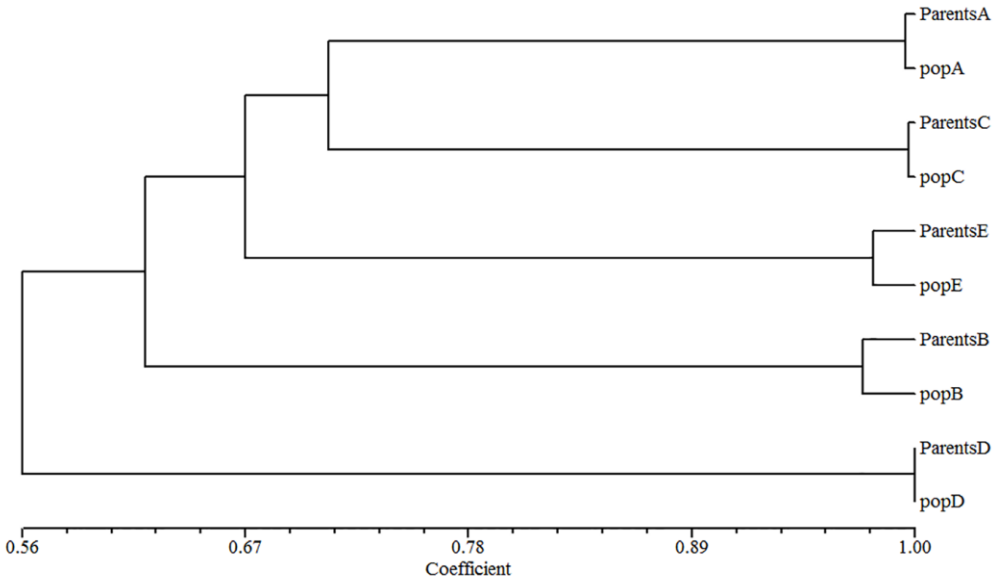


Figure S2 The clustering analysis of five polyculture offspring and parents.