GMR

Characterization and development of 56 EST-SSR markers derived from the transcriptome of *Odontobutis potamophila*

P.P. Wang[1,2], M. Ren[1], S.Q. Chen[2], S.W. Yin[1,3], C. Zhao[1], H.Y. Zhang[1], X.R. Li[1], Q.Q. Cao[1] and G.Q. Zhou[2]

[1]College of Life Sciences, Key Laboratory of Biodiversity and Biotechnology of Jiangsu Province, Nanjing Normal University, Nanjing, Jiangsu, China
[2]Institute of Fisheries Science of Nanjing city, Nanjing, Jiangsu, China
[3]Co-Innovation Center for Marine Bio-Industry Technology of Jiangsu Province, Lianyungang, Jiangsu, China

Corresponding author: S.W. Yin
E-mail: yinshaowu@163.com

Genet. Mol. Res. 16 (2): gmr16029129
Received September 2, 2016
Accepted March 21, 2017
Published May 10, 2017
DOI http://dx.doi.org/10.4238/gmr16029129

ABSTRACT. Expressed sequence tags (ETSs) are the sources of microsatellite development. In this study, we isolated and characterized microsatellite markers for *Odontobutis potamophila* by using Illumina RNA-sequencing. We sequenced a large number of ESTs and screened 200 potential microsatellites. Consequently, a total of 56 novel polymorphic microsatellite repeat markers were identified in thirty-two individuals from a wild population area (Jiande, Zhejiang Province, China). The number of alleles per locus varied from two to eight, the observed heterozygosity (H_O) ranged from 0.03571 to 0.9375, and the expected heterozygosity (H_E) ranged from 0.14326 to 0.81549. The average number of alleles, H_O, and H_E were 5.0, 0.4467, and 0.5518, respectively. By the calculation, the range of polymorphism information content (PIC) was 0.1177-0.8492. Most of the loci showed moderate or high polymorphism. These newly developed EST-simple

sequence repeat (EST-SSR) markers would serve as an efficient tool for analyzing population connectivity and provide sufficient information for genetic diversity research, parentage, and molecular breeding of *O. potamophila* and other fishes with similar genetic relationship.

Key words: *Odontobutis potamophila*; Expressed sequence tags; Microsatellites; RNA sequencing; polymorphism; simple sequence repeat

INTRODUCTION

Odontobutis potamophila is an important commercial fish species mainly inhabiting China's southeastern rivers, such as a part of the Yangtze River, the Qiantang River, and the Minjiang River systems. This fish is deeply favored by people because of its delicious taste, rich nutrition, and high economic value (Li and Liu, 2016). It has become a very promising aquaculture species in China. However, because of the rapidly declining wild fish resources, environmental pollution, overfishing, and growing market demand, *O. potamophila* has undergone a severe reduction in both size and distribution (Zhu et al., 2014a). In 2012, the International Union for the Conservation of Nature (IUCN) listed *O. potamophila* as a near-threatened species (Huckstorf, 2012). Therefore, the aim of our study was to identify and describe novel microsatellite loci in *O. potamophila* that will be useful for studying the genetic diversity of its wild populations and for exploring genetic markers associated with commercially valuable traits for the aquaculture of this species.

Previous studies have demonstrated that the microsatellite markers derived from expressed sequence tag (ESTs) have achieved high efficiency in gene mapping by EST-simple sequence repeats (EST-SSRs), which are correlated with the genes of known functions and as a useful tool for studying the genetic structure of a fish species (Gao et al., 2012; Hasselman et al., 2013). To this day, several polymorphic microsatellite markers of *O. potamophila* have been developed. However, the molecular markers for *O. potamophila* are still not enough. We cannot fully evaluate the germplasm genetics and molecular-assisted breeding system of this fish. Therefore, we urgently need to develop more molecular markers for this fish.

The denaturing polyacrylamide gel electrophoresis (PAGE) and the capillary electrophoresis (CE) are the two frequently used techniques to estimate the SSR products. Denaturing PAGE does not require expensive instruments and fluorescent labels (Pagel et al., 2016). Comparatively, CE can increase the speed of electrophoresis and provide more accurate data. Therefore, in this study, we developed 56 EST-SSR markers of *O. potamophila* by using the CE technique, and used these markers to obtain detailed genetic background information about the wild populations of this fish species.

MATERIAL AND METHODS

Sixty-four healthy experimental fish were collected randomly from Jiande, Zhejiang province, China. We extracted the genomic DNA from their tail fin tissue by using a centrifugal column-type cell/tissue genomic DNA extraction kit (Shanghai Generay Biotech Co., Ltd, Shanghai, China). Thirty-two individuals were randomly selected to screen the EST-SSR markers and the same number of individuals were used for polymorphism analysis. The PRIMER3 software was used to design the primers for polymerase chain reaction (PCR). The microsatellite repeats were screened using the MISA software (MicroSAtellite, http://pgrc.ipk-gatersleben.de/

misa). Their polymorphism was evaluated by conducting 8% (w/v) PAGE. The PCR was carried out in a 10-μL reaction mixture containing 50 ng template DNA, 200 μM dNTPs, 1500 μM $MgCl_2$, 10X PCR buffer, 0.5-1 U Taq DNA polymerase (Shanghai Generay Biotech Co., Ltd, Shanghai, China), 0.04 μM M13-tailed forward primer, 0.16 μM reverse primer, and 0.16 μM fluorescence-labeled M13 primer. The PCR amplification was performed using the following cycling conditions: initial denaturation at 95°C for 5 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at various primer-specific temperatures (Table 1) for 30 s, and extension at 72°C for 45 s, followed by 8 cycles of amplification at 94°C for 30 s, and a final extension at 72°C for 2 min. Subsequently, ABI 3500 XL sequencer (Life Technologies, Foster City, CA, USA) was used to analyze the PCR products, and their sizes were determined using GENEMAPPER, version 4.1. The POPGENE software, version 1.32, was used to calculate the number of alleles per locus, allele size range, observed heterozygosity (H_O), expected heterozygosity (H_E), genetic distances, and genetic similarity for each sample. The Hardy-Weinberg equilibrium (*HWE*) and PIC were estimated by using the programs Arlequin 3.1 and PIC_CALC, respectively. All experiments were accomplished in accordance with the guidelines of the Care and Use of Laboratory Animals protocol in China, and the fish experiment was approved by the Ethics Committee of Experimental Animals at Nanjing Normal University, China.

Table 1. Characterization and genetic parameters of 56 novel EST-SSR markers derived from *Odontobutis potamophila*.

Loci	GenBank accession No.	Primer sequences (5'→3')	Repeat motif	Tm (°C)	Size (bp)	N_A	H_O	H_E	P value HWE	PIC
OPRM1	KR861467	F: GGTGGACGGCAAAGTAAAAA R: TTGCATTTGTAATAAAATTTCTCCC	$(TTGAA)_5$	50	138-143	2	0.60000	0.48814	0.26501	0.3968
OPRM2	KR861468	F: GGCCCCCTACAGGACATTAG R: GGTTGTTGCACTTGACCAAA	$(GATT)_5$	50	220-254	5	0.37500	0.68403	0.00054*	0.5504
OPRM8	KR861469	F: GGGTCCTTCTCACGGTAACA R: GAAAAGTTGTGCTGGCTCGT	$(T)_{11}...(GAG)_5$	50	164-176	4	0.70370	0.72886	0.58076	0.7693
OPRM10	KR861470	F: ATGTCGAGCAGTAACCAGGG R: CCGCTCCTTCTGAAAAGAAA	$(AT)_{6}...(CT)_7$	50	292-294	2	0.48000	0.48980	1.00000	0.2824
OPRM11	KR861471	F: AAACAGGGAAAAGCAAAGCA R: CCAGGCTTAGGAGAGCAGAA	$(TCT)_5...(T)_{11}$	50	218-222	3	0.44828	0.60012	0.06964	0.7512
OPRM15	KR861472	F: GCTCCAGGTGATGTTGTCCTA R: CAGGAAGACAGCACTTCAAAAA	$(AT)_7...(TGA)_5$	50	186-196	4	0.26667	0.27401	0.50341	0.2322
OPRM17	KR861473	F: CGCCGAACAAACTACTGTCC R: CTGAAAACAAACCATGGGAAG	$(TTA)_{6}...(TTA)_5$	50	244-260	5	0.59375	0.73462	0.15441	0.5714
OPRM19	KR861474	F: ATGTCATCGTGGTGGTTTGA R: AGGGTTTGGAAGCAGAGGTT	$(TGG)_5(TGA)_5$	50	266-278	4	0.14815	0.14326	1.00000	0.2970
OPRM23	KR861475	F: ACAAAGCCAACCAACAAAGG R: GCTAAGGAATTCACCCCACA	$(AC)_7...(CA)_7$	50	280-284	3	0.25000	0.52179	0.01551	0.7032
OPRM24	KR861476	F: AACGCCGAACCACAAATACT R: CCACCTACAGCTCGTGGTTT	$(TG)_6(T)_{10}$	50	236-238	2	0.40741	0.50664	0.43922	0.4059
OPRM26	KR861477	F: GATTCCTGGAGAGGGAGGAG R: ATACCAGCCCAAAATGCGTA	$(GT)_9...(TG)_8$	50	244-264	5	0.66667	0.79322	0.14125	0.7273
OPRM30	KR861478	F: AGAGAACAGACTCCAACAGCAA R: CACCGAACCTGACGATTGAT	$(TC)_9...(CTC)_5$	50	224-244	6	0.50000	0.69026	0.05474	0.5961
OPRM31	KR861479	F: GGCCACATTCTTTTGTCGT R: GAGCCTGCCTCTCATTCAAG	$(TA)_6(T)_{10}$	50	248-252	3	0.53333	0.59492	0.12408	0.6303
OPRM32	KR861480	F: GGACTCCTGCGACACTGTTT R: ACGGCCTCCTGTAAATGTTG	$(T)_{10}...(GT)_6$	50	256-266	3	0.58065	0.64992	0.00080*	0.3945
OPRM38	KR861481	F: CATGCTACTGAGGGCACAGA R: CGCTTCAAGTTGTCAAATATCG	$(T)_{10}...(TGA)_5$	50	282-292	6	0.58621	0.72232	0.02815	0.7456
OPRM41	KR861482	F: CACACAATGACAAAGCCCAC R: TTTCCCCCTCATTTTATCCC	$(A)_{14}...(GT)_7$	50	290-300	4	0.50000	0.65844	0.26196	0.4081
OPRM43	KR861483	F: CGTTTATGTTAGGGGGAGGG R: TCCAGCCTTTATTTTTCCCC	$(GT)_{6}...(GT)_6$	50	284-290	4	0.48276	0.75257	0.01024	0.6916
OPRM44	KR861484	F: CCTGTGACATTGTGGCTGTT R: CATCACAGGCAGGCGTATTA	$(A)_{11}...(TG)_6$	50	236-258	6	0.78125	0.76736	0.00256	0.6156
OPRM45	KR861485	F: TGCCTTGCTCATGTCTGAAG R: GAAGAGGATCAGTGTTGTGACCT	$(T)_{16}...(T)_{14}...(AC)_6$	50	264-287	7	0.03571	0.78506	0.00000*	0.7411
OPRM49	KT805136	F: GTGCCTTGCAGGCTATAGGT R: CAGAGAATTCCAGGCCATTG	$(CA)_{8}...(CA)_6$	50	186-192	3	0.16129	0.15283	1.00000	0.5877
OPRM50	KR861486	F: GTGCCTTGCAGGCTATAGGT R: CAGAGAATTCCAGGCCATTG	$(CA)_{8}...(CA)_6$	50	185-191	3	0.15625	0.14831	1.00000	0.3925
OPRM54	KR861487	F: TTTGCAGAAACGCTCTTGAA R: CGTTTGATTCATTGAAAGCC	$(T)_{11}...(TCA)_6$	50	254-260	4	0.39286	0.66429	0.01409	0.6094
OPRM55	KR861488	F: TTGCCTTATAAATGCCTTCTTGA R: GCCACAGTGACACCCTGAAT	$(T)_{10}(GA)_6$	50	280-284	2	0.34375	0.32887	1.00000	0.4018
OPRM57	KR861489	F: GGGTGTCTACACGACTGCAA R: GACTCTGATTGTGACGTGGC	$(T)_{11}...(GA)_9$	50	242-250	3	0.26087	0.52077	0.00976	0.4610
OPRM58	KR861490	F: CCCCCTTTTCTCCACATTT R: GCACGAGGGGAGTTAGCATA	$(TC)_9...(CT)_6$	50	222-230	5	0.51852	0.69811	0.01109	0.6463
OPRM60	KR861491	F: GGACAAAGGGCATTAGCAA R: ACACAACCCTAGCATTTGCC	$(TG)_6...(T)_{10}$	50	206-212	4	0.42857	0.54091	0.16912	0.4693
OPRM61	KR861492	F: GGACAAAAGGGCATTAGCAA R: ACACAACCCTAGCATTTGCC	$(TG)_6...(T)_{10}$	50	185-187	3	0.58065	0.52459	0.00014*	0.4339
OPRM62	KT805137	F: GAGAGGCAGGGAGACAACTG R: GGTGCCTGTTTTCTGAGAGC	$(TG)_6...(A)_{10}$	50	182-198	3	0.93750	0.56548	0.00000*	0.5645

Continued on next page

Table 1. Continued.

Loci	GenBank accession No.	Primer sequences (5'→3')	Repeat motif	Tm (°C)	Size (bp)	N_A	H_O	H_E	P value HWE	PIC
OPRM76	KT805138	F: GCTTTTGAGCCTGCTGTTTT R: AGGAAGGTGAGGCTGGAAAT	$(TA)_{7...}(GT)_7$	50	272-296	3	0.66667	0.58156	0.01484	0.4451
OPRM87	KR861493	F: GACGTCACGGTCAATCAATG R: CGACACCTGGACAAACAATG	$(ATA)_5$	50	270-282	4	0.56000	0.54531	0.51515	0.7065
OPRM99	KT805139	F: AGAGGGTTGGGAGGACTGTT R: AGACCTTAACCCCAGCCATC	$(TAT)_7$	50	202-206	2	0.33333	0.40881	0.36430	0.6220
OPRM104	KR861494	F: GCACTATGTGTTGCTGAGGC R: CCCTCGTGCTCGTGTCTATT	$(ATT)_6$	50	276-285	4	0.61290	0.73559	0.04637	0.4042
OPRM109	KR861495	F: TTCTGCCCTCCTTGATGTCT R: TTGGTTCAAAAGCAGCTGTG	$(CTT)_6$	50	288-292	2	0.56667	0.46271	0.25397	0.5666
OPRM110	KT805140	F: GTTTCAGAATAATGCGCGGT R: CGTACATGAAGGGAAGGTGG	$(GCA)_6$	50	202-206	2	0.38462	0.50679	0.25694	0.3740
OPRM112	KR861496	F: GGATTTTGGACAACTGGCAC R: GGCACCTTCTCCAGTCGATA	$(ATT)_6$	50	226-232	3	0.26087	0.24058	1.00000	0.3770
OPRM114	KR861497	F: ATCCCGTGTAACTGTACGGC R: GTGAGACCACGAGCACTTCA	$(GAA)_6$	50	231-248	7	0.33333	0.75611	0.00000*	0.7038
OPRM115	KR861498	F: CTGACAAAAGGCACCAGACA R: AGGCTGGCCTTGTGACTAAA	$(TTG)_6$	50	202-206	2	0.48276	0.47913	1.00000	0.4061
OPRM117	KR861499	F: CACAACGCAACCAAAACATC R: GAGACGCTCCGCATAAACTC	$(GAA)_6$	50	272-276	2	0.43333	0.50339	0.48033	0.3956
OPRM118	KT805141	F: ACGATGAGGCTCTTTCCAGA R: CTTGGTGTTGGATTGTCGTG	$(GCT)_6$	50	230-236	3	0.53125	0.59524	0.63681	0.5813
OPRM120	KR861500	F: AAAGTGAGTGCCAGTCAGCA R: GCCCTCGCCCTAGTCTTTAC	$(TAA)_6$	50	127-172	8	0.34483	0.81549	0.00000*	0.7748
OPRM123	KR861501	F: ACGACAGGCAGCAGAGAAGT R: ATGGGCTTTGGTGTTGTTTC	$(CAG)_6$	50	218-227	4	0.48276	0.67453	0.03301	0.7685
OPRM124	KR861502	F: TGGAACACACAAGCTCCAAA R: ACGCGCAGCTATCAATTTCT	$(AAT)_6$	50	254-266	3	0.12500	0.22569	0.05065	0.1177
OPRM125	KR861503	F: TTGGGAATACGTCAGGCTTC R: CACGAGGAGCAAGTGAATGA	$(TCT)_6$	50	160-169	4	0.53333	0.61864	0.11246	0.7045
OPRM126	KR861504	F: ACCTACCTCTGCATCCCCTT R: GATCCCTGAGTGACCTGAGC	$(CTG)_6$	50	118-124	3	0.44828	0.62674	0.03903	0.3732
OPRM129	KR861505	F: CTCCCAAAACAAACCCAGAA R: AAAGGCCGTCTGAGCTGTTA	$(CCT)_6$	50	254-266	3	0.09375	0.14831	0.15576	0.1719
OPRM151	KR861506	F: CAGTTAGAGGCAACATGGCA R: CAGTCGGGTTAAGTTGGGAA	$(AG)_{10}$	50	240-268	5	0.06250	0.28423	0.00003*	0.2678
OPRM154	KR861507	F: TTACAGCCTCCATCAGGGTC R: TTGTTTAAGACCGGGTCCAG	$(GT)_{10}$	50	222-230	3	0.61538	0.67798	0.08565	0.8492
OPRM170	KR861508	F: AGCATTCCTGCTGCTGTCTT R: AACTCCTCTGACTGCCTGGA	$(GA)_9$	50	256-272	5	0.25000	0.53312	0.00000*	0.4880
OPRM171	KR861509	F: AGCATTCCTGCTGCTGTCTT R: AACTCCTCTGACTGCCTGGA	$(GA)_9$	50	258-272	4	0.44000	0.56245	0.00001*	0.5001
OPRM176	KT805142	F: AGTCCGACGGAGTTTTTGTG R: CAATTGACTAGCTGCACGGA	$(AC)_9$	50	216-226	6	0.61905	0.79907	0.08677	0.4248
OPRM180	KT805143	F: AGCTTCCCAACGTTTACCCT R: TCAGAGCCAGTGCTGCTTTA	$(AC)_9$	50	262-268	4	0.50000	0.65932	0.04475	0.5969
OPRM189	KT805144	F: TGTGCAACATCCCTGCTTAG R: CTCCTTTCTTGCTCTGTGGG	$(CA)_9$	50	260-266	4	0.90625	0.65129	0.00052*	0.5813
OPRM191	KT805145	F: CAACCGTAAACCTGTTGCCT R: CCAACGGAATCCAGCTTAAA	$(GT)_9$	50	256-264	5	0.67742	0.69223	0.21157	0.5734
OPRM195	KT805146	F: TTCTCAGCACATCAACCAGC R: ATACTGGCATTAGGCAACCG	$(AC)_9$	50	244-252	5	0.53571	0.61558	0.63076	0.4698
OPRM197	KT805147	F: TCTAGCAACATTGTGTGCCC R: CCCCTGACCTTTGATATGGA	$(CA)_9$	50	236-262	8	0.33333	0.50226	0.00049*	0.6037
OPRM200	KT805148	F: GTGTGTGGAGACAGGACACG R: AAAAGTATGCACCAGGGACG	$(TG)_9$	50	269-301	6	0.16129	0.26758	0.01903	0.3750

Tm, annealing temperature; N_A, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; P value HWE, probability values for exact tests of Hardy-Weinberg equilibrium (*statistically significant after Bonferroni correction); PIC, polymorphism information content.

RESULTS AND DISCUSSION

A total of 40,905 EST sequences from the cDNA library of *O. potamophila* were constructed in our laboratory. We identified a total of 1,321 SSR loci, of which 200 were selected for microsatellite marker optimization. All loci were successfully amplified. Finally, fifty-six microsatellites were successfully amplified, and all loci were shown to be polymorphic in the Jiande population. The allele number per locus ranged from two to eight. Expected heterozygosities (H_E) were 0.14326-0.81549 (mean 0.5518). The H_O level range was 0.03571-0.9375 (mean 0.4467), and the *PIC* value ranged from 0.1177 to 0.8492. Three microsatellite loci (OPRM15, OPRM124, and OPRM129) in the Jiande samples showed low polymorphism; all other sites presented moderate or high polymorphism. No significant linkage disequilibrium was observed for any locus in the Jiande samples. Microsatellite loci were successfully investigated for the Jiande population of *O. potamophila*. The repeat motifs, primer sequences, and polymorphic parameters are shown in Table 1.

In the present study, we detected a greater number of polymorphic microsatellite

markers than the previous studies (Zhang et al., 2014; Zhu et al., 2014b; Li et al., 2015), and most of them were in *HWE*. These newly identified and characterized polymorphic microsatellite loci will serve as a useful tool for determining commercially valuable traits, such as quantitative trait locus position, and for studying population genetic diversity, parentage assessment, and molecular ecology of *O. potamophila*.

Conflicts of interests

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the Jiangsu Province Science and Technology Support Program (#BE2013441), Jiangsu Province Six Talent Peaks of High-Level Talents Project (NY-032), Nanjing Science and Technology Plan Projects (#201505059), Jiangsu University Brand Professional Construction Project Funds (J1103507), and Project Foundation of the Academic Program Development of Jiangsu Higher Education Institution (PAPD).

REFERENCES

Gao Z, Luo W, Liu H, Zeng C, et al. (2012). Transcriptome analysis and SSR/SNP markers information of the blunt snout bream (*Megalobrama amblycephala*). *PLoS One* 7: e42637. http://dx.doi.org/10.1371/journal.pone.0042637

Hasselman DJ, Ricard D and Bentzen P (2013). Genetic diversity and differentiation in a wide ranging anadromous fish, American shad (*Alosa sapidissima*), is correlated with latitude. *Mol. Ecol.* 22: 1558-1573. http://dx.doi.org/10.1111/mec.12197

Huckstorf V (2012). *Odontobutis potamophilus*. In: IUCN 2012. IUCN red list of threatened species.

Li Q and Liu Z (2016). New complete mitochondrial genome of the *Odontobutis potamophila* (Perciformes, Odontobutidae): genome description and phylogenetic performance. *Mitochondrial DNA A DNA Mapp Seq. Anal.* 27: 163-164.

Li Q, Wang XB and Liu ZZ (2015). Isolation and characterization of polymorphic microsatellite markers for the river sleeper (*Odontobutis potamophila*). *Conserv. Genet. Resour.* 7: 251-253. http://dx.doi.org/10.1007/s12686-014-0350-1

Pagel UR, Reis RS, Carvalho VP, Santos EVW, et al. (2016). Comparative analysis of short tandem repeat data obtained by automated and gel electrophoresis techniques. *Genet. Mol. Res.* 15: 1-7. http://dx.doi.org/10.4238/gmr.15038436

Zhang LJ, Zhang HW, Zhang YP, Zhu F, et al. (2014). Development and characterization of 42 novel polymorphic microsatellite markers for *Odontobutis potamophila* from EST sequences. *Conserv. Genet. Resour.* 6: 469-472. http://dx.doi.org/10.1007/s12686-013-0130-3

Zhu F, Luo J, Yin SW, Zhang LJ, et al. (2014a). Isolation and characterization of twenty-eight polymorphic microsatellite markers in *Odontobutis potamophila* and cross-amplification in other Gobioidei. *Conserv. Genet. Resour.* 6: 601-604. http://dx.doi.org/10.1007/s12686-014-0150-7

Zhu F, Zhang LJ, Yin SW, Zhang HW, et al. (2014b). Genetic diversity and variation in wild populations of dark sleeper (*Odontobutis potamophila*) in China inferred with microsatellite markers. *Biochem. Syst. Ecol.* 57: 40-47. http://dx.doi.org/10.1016/j.bse.2014.07.002