Pakistan J. Zool., vol. 51(4), pp 1335-1341, 2019. DOI: http://dx.doi.org/10.17582/journal.pjz/2019.51.4.1335.1341

Development and Validation of a Transcriptome-Based Simple Sequence Repeats Markers in *Coilia nasus*

Fukuan Du[1,2], Yongkai Tang[1], Juhua Yu[1], Shengyan Su[1], Fan Yu[1], Jianlin Li[1], Hongxia Li[1], Meiyao Wang[1] and Pao Xu[1,2,*]

[1]*Key Laboratory of Freshwater Fisheries and Germplasm Resources Utilization, Ministry of Agriculture, Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, Wuxi 214081, China*
[2]*Wuxi Fisheries College, Nanjing Agricultural University, Wuxi, Jiangsu, 214081, China*

ABSTRACT

Estuarine tapertail anchovy (*Coilia nasus*) is a commercially important species. However, eco-environment deterioration and overfishing have almost caused extinction of the species in the Yangtze River. Therefore, it is very urgent to perform genetic research on *C. nasus* to protect the wild population. In this study, we sequenced the brain transcriptome of *C. nasus* using Illumina Hiseq 4000 platform. We obtained 123,764 unigenes with an average length of 1092 bp and a total of 115,169 putative Simple Sequence Repeats (SSRs). Among them, 37 SSRs were selected for the validation experiments. All of the loci were found to be polymorphic and showed bi-allelic in 60 individuals of *C. nasus*. These SSR markers should not only be useful for population conservation, but also for construction of genetic linkage map and economic performance improvement of *C. nasus*.

Article Information
Received 25 July 2018
Revised 29 August 2018
Accepted 15 September 2018
Available online 02 May 2019

Authors' Contribution
PX and FKD conceived the study and designed the project. FKD and YL prepared the samples. YL implemented the SSR validation. YKT, JHY, FY, SYS, JLL, HXL, MYW and FKD discussed the data. FKD wrote the manuscript. All authors contributed to data interpretation.

Key words
***Coilia nasus*, SSR markers, Transcriptome.**

INTRODUCTION

Estuarine tapertail anchovy (*Coilia nasus*) is a commercially important species due to its nutritive value and delicacy (Jiang *et al*., 2016; *Et al*., 2017a). This fish widely distributed in the Ariake Sound of Japan, the coastal waters of Korea and the Yangtze River in China (Jiang *et al*., 2012, 2016). However, eco-environment deterioration and overfishing have almost caused extinction of the species in the Yangtze River (Dou *et al*., 2012; Zhang *et al*., 2005). Therefore, it is very urgent to perform genetic research on *C. nasus* to protect the wild population. To date, few studies have been focused on this fish, especially the study of its genetic diversity using molecular markers (Fang *et al*., 2015; Han *et al*., 2015). To facilitate a study on the conservation genetics of this species, here we isolated and characterized 37 novel simple sequence repeats (SSRs) for *C. nasus* that will be as a powerful tool to study genetic diversity.

MATERIALS AND METHODS

Experimental animals and tissue sampling

C. nasus was collected from the Yangtze river and the fish was adapted to a 10.0 × 10.0 × 15 m^3 offshore cage with a dissolved oxygen concentration of 9.2 ± 3.5 mg O2/L. The fish was fed by fresh shrimp three times daily, at 7:00 AM, 12:00 noon and 5:00 PM. The fish was euthanized with 70 mg/L buffered tricaine methanesulfonate (MS-222) and submerged immediately in crushed ice to retard degradation of RNA. The brain of the fish was removed and placed in liquid nitrogen.

Animal welfare and experimental procedures were carried out in accordance with the Guide for the Care and Use of Laboratory Animals (Ministry of Science and Technology of China, 2006), and were approved by the animal ethics committee of Chinese Academy of Fishery Sciences.

RNA sequencing, assembly and annotation

RNA was extracted from brain tissues of these 60 individuals by using TRIzol reagent (Invitrogen, Shanghai, China), respectively. After extracting total RNA and treating with DNase I, Oligo(dT) was used to isolate mRNA. The mRNA is fragmented with the fragmentation buffer. Then cDNA is synthesized using the mRNA fragments as templates. Short fragments are purified and resolved with EB buffer for end reparation and single nucleotide A (adenine) addition. After that, the short fragments are connected with adapters. The suitable fragments are selected for the PCR amplification. Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System are used in quantification and qualification of

* Corresponding author: xup13806190669@163.com
0030-9923/2019/0004-1335 $ 9.00/0

the sample library. Then the library is sequenced using Illumina HiSeq 4000 that generated approximately 20.89 Gb paired-end (PE) raw reads (BGI, Wuhan, China; Table I). Raw data were deposited in the NCBI Sequence Read Archive under the accession number (PRJNA144596). After removing adaptor sequences, ambiguous 'N' nucleotides (with the ratio of 'N' greater than 5%) and low quality sequences (with quality score less than 10), the remaining clean reads were assembled using trinity software (Grabherr *et al*., 2011) as described for *de novo* transcriptome assembly without a reference genome. Blast (Lobo, 2008) was used to align Unigenes to NT, NR, COG, KEGG and SwissProt to get the annotation, use Blast2GO (Conesa *et al*., 2005) with NR annotation to get the GO annotation, and use InterProScan5 to get the InterPro annotation (Quevillon *et al*., 2005).

Unigene SSR analysis

SSR was detected with software Microsatellite (MISA, http://pgrc.ipk-gatersleben.de/misa/misa.html) using unigenes as references. Keep the SSRs which the length of both ends on the Unigene is more than 150 bp. We use these sequences to design primers (*Et al*., 2017b).

SSRs validation

Primer pairs for selected SSRs were designed using the Primer 3.0 software. For each primer pair developed, the forward primer was labeled with FAM and HEX fluorescent dye. Thirty *C. nasus* individuals collected from the Yangtze River. Genomic DNA was extracted from muscle tissues using the kit (Tiangen, Beijing, China). PCR amplifications were performed in 25 uL volumes containing 1.25 units of Taq DNA polymerase (Takara, Dalian, China), 2.5 uL PCR buffer, 0.2 mM dNTP mix, 1 mM of each primer, and about 100 ng of template DNA. The amplifications were programmed using the following conditions: 95°C for 5 min, then 35 cycles at 94°C for 0.5 min, proper temperature (Table I) for 0.5 min, 72°C for 0.5 min, and a final extension at 72°C for 10 min. The PCR products were genotyped on an ABI 3730 sequencer (Applied Biosystems, CA, USA) and analyzed with GeneMarker v1.8 software. POPGENE 32 software (Francis *et al*., 2000; Shao *et al*., 2016) was used to analyze the Observed and expected heterozygosity values, deviations from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD).

RESULTS

Generation of transcriptome data

In this study, we sequenced the brain transcriptome of 60 *C. nasus* using Illumina Hiseq 4000 platform and generated about 20.89 Gb bases in total. After removing the

Table I.- Raw data and the reads quality metrics.

Parameters	Values
Total raw reads (Mb)	208.94
Total clean reads (Mb)	199.89
Total clean bases (Gb)	19.99
Clean reads Q20 (%)	97.24
Clean reads Q30 (%)	92.89
Clean reads ratio (%)	95.67

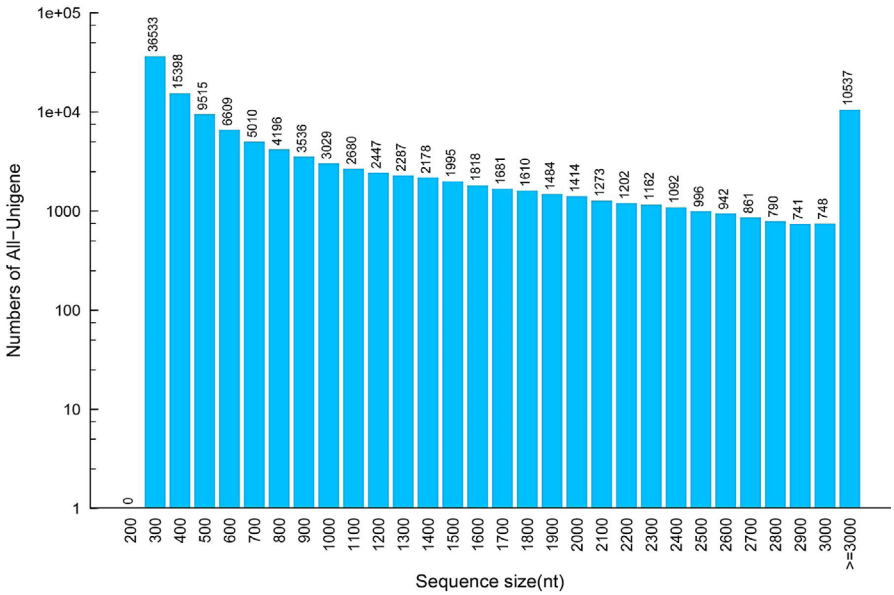

Fig. 1. The unigene length distribution.

sequencing reads which containing low-quality, adaptor-polluted and high content of unknown base(N) reads, a total of 199.89 Mb clean reads were obtained and these reads covered a total of 19.99 Gb bases. The reads quality metrics are shown in Table I.

De novo *assembly of transcriptome data*

After reads filtering, Trinity was used to perform de novo assembly with clean reads, and the Unigene length distribution is shown in Figure 1. We got 123,764 Unigenes, the total length, average length, N50, and GC content of Unigenes are 135,176,777 bp, 1,092 bp, 2,198 bp, and 48.74 %, respectively, and then annotate Unigenes with 7 functional databases, finally, 67,244 (NR: 54.33%), 78,462 (NT: 63.40%), 55,701 (Swissprot: 45.01%), 23,431(COG: 18.93%), 48,294 (KEGG: 39.02%), 13,147 (GO: 10.62%), and 39,403 (Interpro: 31.84%) Unigenes are annotated.

Unigene SSR filter and analysis

Using MISA (Beier *et al*. 2017) to mining putative SSRs, a total of 115,169 putative SSRs were detected, which including 11,569 mono-nucleotide, 85,627 di-nucleotide, 11,655 tri-nucleotide, 3,614 quad-nucleotide, 1,998 Penta-nucleotide and 706 hexa-nucleotide (Fig. 2).

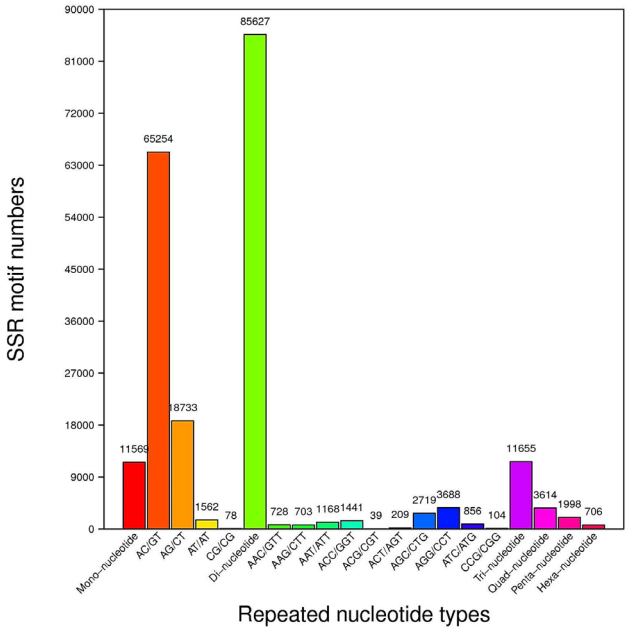

Fig. 2. SSRs type statistics.

Table II.- Characterization of 37 developed microsatellite markers in *C. nasus*.

Locus	Repeat motif	Primer (5'-3')	Ta (°C)	Na	Size range (bp)	H_o/H_E	P
9713	(TAGAA)7	F: TCCTCAGAGTATCACCTGTTCATCA R: ACCTGTGTACAACATTTCGTTACCT	58	8	181-247	0.958/0.884	0.999
1123	(GAATA)4	F: ACCATCTATCAACAGAATTTCTCCTCCT R: CGACTGATTCATTGTACAGTGCACA	59	6	120-180	0.833/0.821	0.960
1124	(AAAAT)5	F: TGAGCCTTCACTCTGTCTAATCAGA R: TTTTTGTTTTGGTTTTTCACTCACAACA	58	7	80-150	0.583/0.820	0.211
5214	(TCTTC)22	F: TGAGTGCCCCATATTTTCTTGGCA R: AAGGAAAGCTGTGAGACACCTTGG	60	10	199-283	0.750/0.800	0.985
6542	(TTCTC)21	F: AGTCCTCTTCCAAAGTCCTCTTCC R: AACAAAGAGAAAGCAAGAGAAGGGA	58	8	125-184	0.792/0.862	0.593
9715	(ATTCT)13	F: TGTGTATAGTGGCAATGAAAGGCA R: TTTTGACACCACACTCGCTTTCAT	58	7	105-168	0.833/0.836	0.868
5208	(TATAT)12	F: CTGAGACTGAGCAGAATGTTGCCA R: GTCATGCTCACCACTGCTCCTATC	61	9	93-206	0.832/0.856	0.984
6524	(ATAAT)11	F: TGCCTCAGTATCACTTTATCACGCT R: GAAGAGAGGGCATTTGGGCAGTTA	60	6	113-178	0.833/0.809	0.897
11429	(AGAAT)9	F: GCTGGCATCACCGAACCTTACTTT R: TTCAGCGGCAACAAAGGTTCATTG	61	6	143-169	0.792/0.788	0.966
7651	(GTGCT)19	F: CTCTGTTGAGTGCTGTGCTGTGC R:CAGCAGAGGAATCAGTGTGCACAC	62	8	137-187	0.958/0.889	0.998
2280	(AAAAC)4	F: TAGATGGCTGTTCTGTTCCAGTT R: TGCATTAACGGTTTTAACTTTCA	56	6	104-156	0.667/0.809	0.033
3174	(AAAAC)4	F: ATGACCCTCAACTCAACTCTCAA R: CTTGTACTGCACAAAGACGCATA	58	8	105-165	0.667/0.853	0.225

Locus	Repeat motif	Primer (5'-3')	Ta (°C)	Na	Size range (bp)	H_o/H_E	P
27050	(AAGCC)4	F: GAGAGAGAGAAAGAGAGAGGGGA R: AGGCTACGAATCCATCTGAGAG	58	7	105-155	0.917/0.858	0.857
9212	(AAGG)5	F: GCTTGATGAGAGAACCTCACACT R: TATGTACACGCTGTTTGCTCAGT	58	6	110-160	0.708/0.818	0.322
3121	(ACAG)5	F: TGTATGTGTGTATGGGAGAGTGG R: AGAGTTCTCCGTCTGATTCACAT	58	5	128-178	0.292/0.753	0.000*
16915	(AATC)6	F: CAAAGCTTGGGATCAATAAAGTG R: AGTTTGAGACCCAAAACAAACAC	58	8	105-176	0.875/0.880	0.997
2424	(AATA)5	F: CCAGTCAGCTCTTTCGACTAAAA R: AAAGCAGGTGTCAGTCATTTGAT	58	7	146-187	0.750/0.810	0.900
6517	(AAGA)5	F: CCAAATACAGCCGAGATTACAGT R: ATGGTGTGTTTCCAAAGTGTTGT	58	8	138-186	0.792/0.856	0.893
920	(AAAT)6	F: AAAGAAGGAAGTAAGGAAGCGAA R: ACTGCAATCCATTTCTGTTGTTT	58	7	135-176	0.625/0.804	0.782
16253	(AACT)5	F: CTGACTGTCCAAACAAACAAACA R: GTCTGAGCAGATGGTCAGTGG	58	9	106-165	0.750/0.864	0.440
24982	(AACC)6	F: AGAAGTAACGGAGCTGAACACAC R: TTCTAGCCCAGGTTTGTTGATAG	58	10	103-176	0.792/0.874	0.715
8681	(AACA)5	F: AAAATTCAAGTGAGGCAGAATTG R: GTTGATGTCATTTTGGGTTGTTT	58	8	135-168	0.875/0.883	0.993
8895	(AACA)5	F: GTGGAGAGGATGGACTGAACTTA R: GTGTCCTGCGGTGTAGTATTTGT	58	7	125-186	0.667/0.857	0.033
395	(AAT)5	F: TCGAGTGAGATTTACACATTTGG R: TGGCAACTATCCACTTGACACTA	58	10	106-168	0.792/0.871	0.805
1204	(ACA)5	F: ACAGCCTGTTCAATCACAACC R: CTGCCAGTTCTCACTCTCTCTGT	58	9	105-180	0.791/0.831	0.966
4993	(CCA)5	F: GATTCCAGACAGATGGGAGTACA R: GACTAGTCCTGGTGGTTTGCAC	58	8	120-185	0.917/0.881	0.998
19568	(CAT)8	F: GGAGGAAGAGAAGGAGATGAAGA R: GTTCTGTAGCTCATGGATATGGC	58	11	128-150	0.625/0.747	0.965
3293	(CAG)5	F: GCTAATACAGTACGTGGCCAATG R: GTCTGGTAGCAAAGGAGTTTGG	58	8	134-168	0.875/0.879	0.987
1030	(AAC)5	F: GTTACTAGATGCCCAGCTGTTGT R: ACATGCTTGCATCACATGTTTAC	58	9	104-156	0.542/0.807	0.031
622	(AAG)5	F: CGAACGTAAAATGAATCCAGAAG R: CAGGAACAAACAGAGTTTCCAGT	58	10	124-165	0.583/0.782	0.640
7543	(CA)9	F: ACTTGTGAAGGACACACACACAC R: TACTCTGACATTGAAAGCTGCTC	56	9	135-180	0.625/0.819	0.126
34568	(CA)6	F: TCTTTAAGGGTGGTGTACATGGT R: AAGTAACTTGACGTTCACGGG	58	10	120-154	0.625/0.838	0.177
11827	(AT)11	F: AATCTGAGACCAAACCAGTTCAA R: TCTCTTACGCTTCAGTAAATGCC	58	5	134-148	0.292/0.653	0.018
19989	(AG)20	F: GCAATTAGCTTTTTCCAATGATG R: TTTCTCTCTTCCTCGCTCTCTCT	58	9	134-166	0.708/0.864	0.494
12125	(AG)18	F: GCAGTGTTTATGCCTTTCTGTGT R: GAAAGACTTGCTGGAAACACATT	58	10	134-168	0.958/0.899	0.993
1009	(AC)44	F: AGTCTCTCTACGTGGCCAACC R: GCGTGTGTGTGTATGTGTGTGT	58	8	134-156	0.917/0.885	0.997
85	(AC)6	F: CACACACACACACAGAGAAAGAGA R: CTGTTATAGCAGCAACATGCAGA	58	7	134-170	0.917/0.870	0.985

*T*a, annealing temperature of primer pairs; *N*a, number of alleles; *H*o, observed heterozygosity; H_E, expected heterozygosity. *Indicates significant deviation from Hardy-Weinberg equilibrium after Chi-square test ($P < 0.01$).

SSRs validation in Coilia nasus

In order to validate the polymorphism of development SSR loci, 37 SSRs were selected for the validation experiments. In these 37 SSRs, all markers showed polymorphic in *C. nasus* and all the 37 pairs of primers were tested in the 60 *C. nasus* individuals (Table II). The observed numbers of alleles per locus ranged from 5 to 11 (mean = 8) (Table II). Observed heterozygosity ranged from 0.292 to 0.958 and expected heterozygosity ranged from 0.653 to 0.889. One locus (3121) deviated significantly from Hardy–Weinberg equilibrium after the Chi-square test ($P < 0.01$). No significant linkage disequilibrium was observed in pairwise loci. These SSR loci will be helpful in studies of genetic diversity and population structure of *C. nasus*.

DISCUSSION

115,169 SSR loci were found in 123764 Unigene sequences of the transcriptome in *C. nasus* and these SSRs distributed in 57,204 Unigenes, which the frequency was 46.22%. This frequency was much higher than that of *Fugu rubripes* (11.50%) (Edwards *et al*., 1998), *Ictalurus punctatus* (11.20%) (Serapion *et al*., 2004), *Carassius auratus* (8. 13%) (Zheng *et al*., 2014), *Cyprinus carpio* (5. 55%) (Wang *et al*., 2007), *Chrysophrys major* (4.00%) (Chen *et al*., 2005) and *Paralichthys olivaceus* (7. 95%) (Chen *et al*., 2013). There could be two reasons for these differences. The first one, different parameters were set when screening SSRs. The types of single base repeats were count in our study, so there were more SSRs. The second one, the transcriptome sequencing depth was more than 30 folds, and Unigene assembly was relatively complete. Therefore, more SSR markers could be identified. The above data indicated that increasing the depth of sequencing could significantly increase the detection rate of SSRs. This also suggested that the depths of transcriptional sequencing that normally used to detect SSR markers were not saturation, and needed to be increased.

The types of SSRs were abundant in the transcriptome of the *C. nasus*, mono- to hexa-nucleotide repeats were found (Fig. 2). The main repeat types were mono-nucleotide, di-nucleotide and tri-nucleotide repeats, accounting for 94.51% of the total SSRs. Quad-nucleotide, pentanucleotide and hexanucleotide repeats accounted for 5.30%. This result was consistent with the distribution of other plants (Cardle *et al*., 2000; Chen *et al*., 2006; Deng *et al*., 2006; Dreisigacker *et al*., 2004) and animals (Teneva *et al*., 2013). It was reported that the proportion of AC/GT in fish was very high, while CG/GC repeats do not exist or the ratio was very low. For example, AC/GT repeat SSRs accounted for 75.4% of the total number of dinucleotide repeat sites, but no CG/GC repeat SSRs were found in *Ictalurus punctatus* (Serapion *et al*., 2004). AC/GT repeat sites accounted for 29.50% of the total number of dinucleotide repeat sites, but no CG/GC repeat sites were found in *Paralichthys olivaceus* (Chen *et al*., 2013). In this study, AC/GT (76.21%) is also dominant, but we also found 78 CG/GC (0.09%), although the proportion is very low, the absolute number is still more. The lower frequency of CG/GC may be associated with cytosine methylation, which causes it to become thymine by deamination, leading to a decrease in the frequency of CG and the increase in the frequency of AC /GT (Schorderet and Gartler, 1992; Cooper and Youssoufian, 1988; Stallings, 1992).

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (Grant No. 31502152), the National Natural Science Foundation of China (Grant No. 31672643).

Statement of conflict of interest

The authors declare no conflict of interest.

REFERENCES

Beier, S., Thiel, T., Münch, T., Scholz, U. and Mascher, M., 2017. Misa-web: A web server for microsatellite prediction. *Bioinformatics,* **33**: 2583. https://doi.org/10.1093/bioinformatics/btx198

Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D. and Waugh, R., 2000. Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics,* **156**: 847.

Chen, C., Zhou, P., Choi, Y.A., Huang, S. and Gmitter, Jr. F.G., 2006. Mining and characterizing microsatellites from citrus ESTs. *Theor. appl. Genet.,* **112**: 1248-1257. https://doi.org/10.1007/s00122-006-0226-1

Chen, S.B., Xie, X.L., Li, G. and Liu, H.J., 2013. Data mining for simple sequence repeats in expressed sequence tags from Japanese flounder (*Paralichthys olivaceus*). *Adv. Mat. Res.,* **765**: 274-277. https://doi.org/10.4028/www.scientific.net/AMR.765-767.274

Chen, S.L., Liu, Y.G., Xu, M.Y. and Li, J., 2005. Isolation and characterization of polymorphic microsatellite loci from an EST-library of red sea bream (*Chrysophrys major*) and cross-species amplification. *Mol. Ecol. Resour.,* **5**: 215-217.

https://doi.org/10.1111/j.1471-8286.2005.00880.x

Conesa, A., Terol, J. and Robles, M., 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics,* **21**: 3674-3676. https://doi.org/10.1093/bioinformatics/bti610

Cooper, D.N. and Youssoufian, H., 1988. The CpG dinucleotide and human genetic disease. *Human Genet.,* **78**: 151-155. https://doi.org/10.1007/BF00278187

Deng, J., Liu, Y. and Huang, H., 2006. Development and characterization of microsatellite markers in *Eucommia ulmoides* Oliver (Eucommiaceae). *Mol. Ecol. Resour.,* **6**: 496-498. https://doi.org/10.1111/j.1471-8286.2006.01287.x

Dou, S.Z., Yokouchi, K., Yu, X., Cao, L., Kuroki, M., Otake, T. and Tsukamoto, K., 2012. The migratory history of anadromous and non-anadromous tapertail anchovy *Coilia nasus* in the Yangtze River Estuary revealed by the otolith Sr:Ca ratio. *Environ. Biol. Fishes,* **95**: 481-490. https://doi.org/10.1007/s10641-012-0042-1

Dreisigacker, S., Zhang, P., Warburton, M.L., Van Ginkel, M., Hoisington, D., Bohn, M. and Melchinger, A.E., 2004. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different mega environments. *Crop Sci.,* **44**: 381-388. https://doi.org/10.2135/cropsci2004.3810

Du, F.K., Xu, G.C., Li, Y., Nie, Z.J. and Xu, P., 2017a. Molecular cloning and differential expression of the glucocorticoid receptorgene in the estuarine tapertail anchovy *Coilia nasus*. *Genet. mol. Res.,* **16**: 16029125. https://doi.org/10.4238/gmr16029125

Du, F.K., Li, Y., Wen, Z.X., Chen. R.G. and Xu, P., 2017b. Development of simple sequence repeats (SSR) by transcriptome in Chinese Mitten Crab (*Eriocheir sinensis* H. Milne Edwards). *Pakistan J. Zool.*, **49**: 497-503.

Edwards, Y.J., Elgar, G., Clark, M.S. and Bishop, M.J., 1998. The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, Fugu rubripes: Perspectives in functional and comparative genomic analyses. *J. mol. Biol.,* **278**: 843. https://doi.org/10.1006/jmbi.1998.1752

Fang, D.A., Zhou, Y.F., Duan, J.R., Zhang, M.Y., Xu, D.P., Liu, K., Xu, P. and Wei, Q., 2015. Screening potential SSR markers of the anadromous fish *Coilia nasus* by de novo transcriptome analysis using Illumina sequencing. *Genet. mol. Res.,* **14**: 14181-14188. https://doi.org/10.4238/2015.November.13.1

Francis, C.Y., Rong, C.Y. and Boyle, T., 2000. *Popgene Microsoft Windows based Freeware for Population Genetic Analysis.* Molecular Biology and Biotechnology Centre. University of Alberta, Edmonton.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Xian, A., Fan, L., Raychowdhury, R. and Zeng, Q., 2011. Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnol.,* **29**: 644. https://doi.org/10.1038/nbt.1883

Han, Z.Q., Han, G., Wang, Z.Y. and Gao, T.X., 2015. The possible physical barrier and coastal dispersal strategy for Japanese grenadier anchovy, *Coilia nasus* in the East China Sea and Yellow Sea: evidence from AFLP markers. *Int. J. mol. Sci.,* **16**: 3283-3297. https://doi.org/10.3390/ijms16023283

Jiang, T., Yang, J., Liu, H. and Shen, X.Q., 2012. Life history of *Coilia nasus* from the Yellow Sea inferred from otolith Sr:Ca ratios. *Environ. Biol. Fishes,* **95**: 503-508. https://doi.org/10.1007/s10641-012-0066-6

Jiang, T., Liu, H., Lu, M., Chen, T. and Yang, J., 2016. A possible connectivity among estuarine tapertail anchovy (*Coilia nasus*) populations in the Yangtze River, Yellow Sea, and Poyang Lake. *Estuar. Coasts,* **39**: 1762-1768. https://doi.org/10.1007/s12237-016-0107-z

Lobo, I., 2008. Basic local alignment search tool (BLAST). *J. mol. Biol.,* **215**: 403-410.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. and Lopez, R., 2005. InterProScan: Protein domains identifier. *Nucl. Acids Res.,* **33**: 116-120. https://doi.org/10.1093/nar/gki442

Schorderet, D.F. and Gartler, S.M., 1992. Analysis of CpG suppression in methylated and nonmethylated species. *Proc. natl. Acad. Sci. U.S.A.,* **89**: 957. https://doi.org/10.1073/pnas.89.3.957

Serapion, J., Kucuktas, H., Feng, J. and Liu, Z., 2004. Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*). *Mar. Biotechnol.,* **6**: 364-377. https://doi.org/10.1007/s10126-003-0039-z

Shao, W., Lin, Z., Fan, X., Ding, G. and Wei, L., 2016. Characterization of 25 microsatellite loci for the guenther's frog hylarana guentheri. *Conserv. Genet. Resour.,* **10**: 1-3.

Stallings, R.L., 1992. CpG suppression in vertebrate genomes does not account for the rarity of (CpG) n microsatellite repeats. *Genomics,* **13**: 890-891. https://doi.org/10.1016/0888-7543(92)90178-U

Teneva, A., Dimitrov, K., Petrovic´, V.C., Petrovic´, M.P., Dimitrova, I., Tyufekchiev, N. and Petrov, N., 2013. Molecular genetics and SSR markers as a new practice in farm animal genomic analysis for breeding and control of disease disorders. *Biotechnol. Anim. Husband.,* **29**: 405-429. https://doi.org/10.2298/BAH1303405T

Wang, D., Liao, X., Cheng, L., Yu, X. and Tong, J., 2007. Development of novel EST-SSR markers in common carp by data mining from public EST sequences. *Aquaculture,* **271**: 558-574. https://doi.org/10.1016/j.aquaculture.2007.06.001

Zhang, M., Xu, D., Liu, K. and Shi, W., 2005. Studies on biological characteristics and change of resource of *Coilia nasus* Schlegel in the lower reaches of the Yangtze River. *Resour. Environ. Yangtze Basin*, **14**: 693-698.

Zheng, X., Kuang, Y., Lu, W., Cao, D. and Sun, X., 2014. Transcriptome-derived EST–SSR markers and their correlations with growth traits in crucian carp *Carassius auratus*. *Fish. Sci.,* **80**: 977-984. https://doi.org/10.1007/s12562-014-0782-2