



Isolation and characterization of polymorphic microsatellite markers from *Coilia ectenes*

X.J. Rong^{1,2}, Y.J. Xu², Q.Y. Wang^{1,2}, M.J. Liao², X.Z. Liu², C.Y. Pan², Z. Zhang² and Y.G. Wang²

¹Ocean University of China, Qingdao, China

²Qingdao Key Laboratory for Marine Fish Breeding and Biotechnology, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, China

Corresponding author: Y.G. Wang

E-mail: wangyg@ysfri.ac.cn

Genet. Mol. Res. 12 (4): 6011-6017 (2013)

Received July 4, 2013

Accepted November 11, 2013

Published November 26, 2013

DOI <http://dx.doi.org/10.4238/2013.November.26.11>

ABSTRACT. *Coilia ectenes* (Jordan and Seale 1905) is an important anadromous species that is an important resource at risk of extinction because of over-fishing, pollution, and coastal construction. To evaluate the genetic diversity of *C. ectenes* for use in breeding programs, elite microsatellite-enriched libraries were constructed and novel microsatellite markers were developed, and applied to genetically detect wild populations. Out of 92 randomly selected and sequenced clones, 89 contained a CA or GA repeat motif. Twenty-two pairs of primers were designed to investigate the polymorphism and genetic structure of a wild population collected from the Yellow River estuary, China. It was found that 2 loci were monomorphic and 20 loci were polymorphic. The number of alleles per polymorphic loci ranged from 3 to 13, with an average of 7.9. The expected heterozygosity per locus ranged from 0.05 to 0.89, with an average of 0.68. The isolated polymorphic markers are expected to be of use in future genetic breeding programs for *C. ectenes*,

and in the assessment of genetic variation within this species.

Key words: *Coilia ectenes*; Microsatellite marker; Genetic diversity

INTRODUCTION

Coilia ectenes (Jordan and Seale 1905) is a commercially important small-to moderate-sized anchovy that inhabits a wide area, including the northwest and western Pacific Ocean, and extending southward toward Canton in southern China and northward to the Ariake Sound of southwestern Japan, including all of the Yellow Sea and the area off the western coast of Korea (Whitehead et al., 1988). *C. ectenes* is an anadromous species; whereby, during the spawning season, individuals migrate annually from coastal waters to freshwater areas (Li et al., 2007). In China, four *Coilia* species have been identified by taxonomists: *Coilia grayii*, *C. ectenes*, *Coilia mystus*, and *Coilia brachygnathus* (Feng, 1997). However, the current taxonomic status of *Coilia* fishes remains unclear, with issues over the classification of *C. ectenes* and *C. ectenes taihuensis*, despite extensive effort to establish clear distinctions using traditional meristic, ecological, and physiological methods (Yuan et al., 1976; Liu, 1995; Cheng and Han, 2004; Cheng et al., 2005a) and genetic analyses (Cheng et al., 2005b, 2008; Ma et al., 2011). Thus, it is necessary to determine the genetic background of *C. ectenes* to clarify the taxonomy of *Coilia* species.

The flesh of *C. ectenes* is delicious and nutrient-rich; thus, it is a popular table food of Chinese consumers. In China, *C. ectenes* is mainly distributed in Qiantang River, the Yangtze River, the Yellow River, the Liaohe River, and several other rivers that connect with the East China, the Yellow, and the Bohai Seas (Yuan, 1980; Yuan and Qin, 1984). This species is used to support important commercial fisheries in China. However, over-fishing, pollution, coastal construction, and a number of other factors have caused a dramatic decline in the landings of *C. ectenes* in recent decades; consequently, this species represents a fishery resource that is on the verge of exhaustion. Thus, there is an immediate requirement to protect this important fishery species in China through various strategies, including fisheries resource assessments, the environment protection of its habitats, and artificial breeding programs for aquaculture. These strategies would contribute towards the sustainable use and management of this important fisheries resource.

Microsatellite markers are valuable tools for studying the genetic composition and variation in a given population. In addition, microsatellites are one of the most polymorphic and effective types of markers for disclosing the genetic diversity and divergence within or between populations. Furthermore, these markers have been proved to be an extremely important tool for genetics studies, as well as in the conservation and management of genetic resources. To date, there are two papers reported to isolate microsatellite markers for *C. ectenes* (Ma et al., 2011; Chen et al., 2012), but the amount of microsatellite is still not enough for assessments about the population structure, molecular phylogeny, and molecular assisted selective breeding of this commercially important fish species. Thus, development of new microsatellite markers for *C. ectenes* is anticipated to facilitate the development of sustainable fishing strategies, and potential artificial breeding programs of this important fish species.

MATERIAL AND METHODS

Sample collection DNA extraction

Sixty individuals of *C. ectenes* were collected from the Yellow River estuary in China. Samples were preserved in alcohol until DNA extraction. Genomic DNA was extracted from Fin clips using an E.Z.N.A.[®] Tissue DNA Kit (Omega, Norcross, USA). The extracted genomic DNA was examined by electrophoresis on agarose gel and stored at -20°C until genotyping.

Microsatellite-enriched library construction

A microsatellite-enriched library was constructed using the FIASCO (Fast Isolation by AFLP of Sequences Containing Repeats) method, which is described in detail by Zane et al. (2002), but with minor modifications, which were described in detail by Liao et al (2007). In brief, genomic DNA was digested with *MseI* enzyme (Sangon, Shanghai, China) at 37°C for 4 h, and ligated with a synthesized *MseI* adaptors (5'-TAC TCA GGA CTC AT-3'/5'-GAC GAT GAG TCC TGA G-3') using T4 DNA ligase (Sangon, Shanghai, China). The digestion-ligation mixture was amplified using the adaptor-specific primer (5'-GAT GAG TCC TGA GTA A-3'). Microsatellite-containing fragments were selectively enriched, captured, and washed using biotinylated-(CA)₁₂ and Streptavidin Magnetic Sphere[®] Paramagnetic Particles (Promega, Madison, USA). Fragments containing microsatellites were ligated with the pMD18-T vector (TaKaRa, Dalian, China), and transferred into *E. coli* competent cell JM109 (TaKaRa, Dalian, China) by electroporation.

Isolation of microsatellite-containing DNA fragments and primer design

To check whether the microsatellite motif was located in the middle of the insert, each recombinant was subjected to 3 separate PCR screenings using 2 universal sequencing primers and the (CA)₁₂DN oligonucleotide. In the first reaction, universal forward and universal reverse sequencing primers were used. In the second reaction, a universal forward sequencing primer and (CA)₁₂DN oligonucleotide were used. In the third reaction, a universal reverse sequencing primer and (CA)₁₂DN oligonucleotide were used. Recombinant clones that produced products of obviously different lengths between the first, second, and/or third reactions were sequenced and trimmed. The sequencing data were scanned using the software SSRHunter V1.3 (Li and Wan, 2005). Sequences with microsatellite motifs and flanking regions were selected for PCR primer design by Primer Premier.

PCR amplification and genotyping

The designed microsatellite primers were used to amplify the genomic DNA of 60 *C. ectenes* individuals. The PCR mixture contained 1X buffer, 1.5 mM/L MgCl₂, 200 μM/L dNTP (each), 200 μM/L primer (each direction), and about 50 -80 ng genomic DNA. The PCR conditions were: denaturing at 94°C for 1 min, followed by 30 cycles of 1 min at 94°C, 1 min at annealing temperature, and 1 min at 72°C, with a final extension for 5 min at 72°C. The PCR product was separated on 6% denaturing polyacrylamide gel and visualized by silver staining.

Allele size was determined with the Quantity One V4.62 software (Bio-Rad, Hercules, CA, USA) by referring to a 20-bp DNA ladder marker (TaKaRa, Dalian, China).

Genetic data analysis

POPGENE version 1.32 (Yeh et al., 1999) was used to calculate the number of alleles (N_A), the number of effective alleles (N_E), observed heterozygosity (H_O), and expected heterozygosity (H_E). The polymorphism information content (PIC) of each locus was calculated according to Botstein et al. (1980). The Hardy-Weinberg equilibrium and linkage disequilibrium tests were conducted using the online version of Genepop (Rousset, 2008). Significance criteria of all multiple tests were corrected following sequential Bonferroni correction (Rice, 1989).

RESULTS AND DISCUSSION

Out of 200 randomly selected recombinant clones, 92 (46%) contained inserts with a microsatellite motif in the middle position, indicating that enrichment was highly effective. Of the 92 clones surviving PCR screening, 89 (70.6%) contained a microsatellite motif in the middle position after being sequenced (Table 1). The sequences were divided into 3 categories following the classification rules developed by Weber (1990): 1) 58 perfect repeat sequences without interruptions in the runs of CA or GT di-nucleotides (65.2% of total), 2) 27 imperfect repeat sequences with 1 or more interruptions in the run of repeats (30.3%), and 3) 4 compound repeat sequences with adjacent tandem simple repeats of a different sequence (4.5%). When classified using the repeat sequence type, all clones had di-nucleotide repeats, of which the repeat number of 41 clones was between 5 to 9, while the repeat number of 48 clones was higher than 10.

Table 1. Classification of microsatellite DNA sequences obtained in this study.

Criterion	Category	No. of sequences	Percentage (%)
Weber (1990)	Perfect	58	65.2
	Imperfect	27	30.3
	Compound	4	4.5
Repeat motif	Two bases $5 \leq n \leq 9$	41	46.1
	Two bases $n \geq 10$	48	53.9

In the first batch, 22 primer pairs were designed based on 30 sequences, and used to investigate the polymorphism of 60 *C. ectenes* individuals. Of the 22 primer pairs tested, 20 loci showed clear band patterns and polymorphism (Table 2). A total of 159 alleles were detected at 20 loci, with a total of 93 effective alleles. The N_A at each locus ranged from 3 (HD28 and HD 118) to 13 (HD25), with an average of 7.9. The N_E ranged from 1.1 (HD118) to 8.9 (HD25), with an average of 4.6. The difference between N_A and N_E was caused by the uneven frequency of each allele. The relationship between the number of microsatellite repeats and polymorphism is subject to debate. Qu et al. (2010) suggested that the polymorphism would be higher when the number of microsatellite repeats increased. Alternatively, Zheng et al. (2008) suggested that many high polymorphic loci would be missed if only loci with high numbers of microsatellite repeats were chosen in genetic research. The results of the current experiment showed that there is no relationship between the num-

ber of microsatellite repeats and polymorphism, which was consistent with Zheng et al. (2008).

Table 2. Characterization of microsatellite DNA markers developed for *Coilia ectenes*.

locus	Accession No.	Repeat motif	Primer sequence (5'→3')	Size range (bp)	T _m (°C)	N _A	N _E	H _O	H _E	PIC	P
HD4	KC506602	(CA) ₁₄	F: CGCTGAGCTGATGTTTG R: GACCCAGTAACCTCATAACCT	320-345	58	7	3.2	0.53	0.69	0.64	0.0338
HD8	KC506603	(CA) ₁₂	F: AATCACAACAAGCGCAATC R: GCAATGACCCTTTCCTTC	86-145	54	12	8.2	0.49	0.88	0.87	0.000*
HD9	KC506604	(CA) ₅	F: TGCTCTTGACATATCGAAGG R: TGGGATTTGTCAGGGAGA	155-186	54	10	3.7	0.50	0.73	0.70	0.000*
HD16	KC506605	(AC) ₇	F: CACAGCTCCCGTGGCTCAA R: AGGCGAGGAGGCGGTATGAG	88-98	63	5	1.7	0.49	0.40	0.35	0.9975
HD25	KC506606	(CA) ₈ (AC) ₁₄	F: ACTCTAACCTTCTCACCTGCTA R: GGGTGAAGGGAAAAGTGT	173-217	55	13	8.9	0.55	0.89	0.88	0.000*
HD28	KC506607	(CA) ₈	F: GACACTTGCGTATTTCCG R: CTGGGAGACACTTTGCTG	187-193	55	3	2.0	0.71	0.50	0.40	0.9882
HD30	KC506608	(CA) ₁₆	F: GCCAAGTGTCTCCAACC R: GTGCGTGCCTGTAGGTGT	165-188	59	8	2.5	0.18	0.60	0.58	0.000*
HD40	KC506609	(CA) ₇ (AC) ₈	F: TCTAAGCCCAACCAAAAAAGATAA R: TAAGGACGAGGCGGGACAGG	622-640	58	5	3.5	0.64	0.72	0.67	0.0159
HD42	KC506610	(AC) ₁₃	F: GCTGACGGCATACTTGGC R: GGGCGTTCTGATAAGGGA	460-482	59	7	4.5	0.52	0.78	0.74	0.000*
HD44	KC506611	(AC) ₂₂	F: AGCCTGCCTTTGCTACAC R: CCTACTGGTGGCTCTGGT	93-118	57	9	4.9	0.53	0.80	0.77	0.000*
HD48	KC506612	(CA) ₇ (CA) ₂₀	F: AACATGCTGGCCGTATCTG R: AGCCTGCCTGATTTAGCG	438-476	59	5	1.9	0.44	0.48	0.40	0.4584
HD55	KC506613	(CA) ₁₉	F: CTGTCCACCTTGCTCACT R: ATAGATCTGCAGGCTCGG	37-129	56	11	8.6	0.58	0.89	0.87	0.000*
HD56	KC506614	(CA) ₆	F: AATGAGAAATAAACAGCAGGAC R: CAGCCTGACCTAGAACCATC	222-258	55	6	1.8	0.28	0.45	0.42	0.0245
HD71	KC506615	(AG) ₁₃	F: CAATGAAACAGTTGAAGACAGC R: ATCCCTCACTGTCCGCTC	483-521	59	9	5.9	0.45	0.84	0.81	0.000*
HD85	KC506617	(GA) ₁₅	F: TACCAAGGGTGTAATAAGC R: ACGGAGCAATGTCAAGTGT	503-551	54	4	1.8	0.44	0.43	0.36	0.6907
HD118	KC506618	(AG) ₂₅	F: GGTTCAGTCTCCCTGCCTCT R: TGCTCCCGCTAACAAAGG	349-357	61	3	1.1	0.03	0.05	0.05	0.0408
HD127	KC506619	(AG) ₅	F: AAGAGGGTGGACTGATGG R: AGGTTGCGTCAGGTTTC	228-268	58	10	4.9	0.78	0.80	0.77	0.3265
HD143	KC506621	(AG) ₁₀ (AG) ₁₁	F: TACATTTGTGGTGGTGGC R: AAAGCAGCAAACCTCATCG	168-196	56	11	8.8	0.51	0.89	0.88	0.000*
HD154	KC506622	(GA) ₁₄	F: CATGAGGGGAACAGGAG R: TCGGACCTAATCTGTGCTG	256-302	57	9	6.8	0.45	0.86	0.84	0.000*
HD160	KC506623	(AG) ₁₈ (AG) ₁₇	F: CACTTTGTCTTGACCCCTT R: AGGGTTATTTTCAGAGCAG	124-186	53	12	8.3	0.67	0.89	0.87	0.000*
Mean							7.9	4.6	0.49	0.68	0.64

T_m = annealing temperature (°C); N_A = allele number; N_E = effective allele number; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism index content; P = P value for exact test for Hardy-Weinberg equilibrium (HWE); *departure from HWE after Bonferroni correction.

The H_O of each locus ranged from 0.03 (HD118) to 0.78 (HD127), with an average of 0.49. The H_E of each locus ranged from 0.05 (HD118) to 0.89 (HD25, HD55, HD143, and HD160), with an average of 0.68. Based on the polymorphic index content (PIC) values for each locus, 1 locus (HD118) exhibited low polymorphism (PIC < 0.25), 5 loci (HD16, HD28, HD48, HD56, and HD85) were moderately polymorphic (0.25 < PIC < 0.5), and the other 14 loci were highly polymorphic (PIC > 0.5). None of the loci exhibited significant linkage disequilibrium. After sequential Bonferroni correction for multiple tests, 11 loci were found to depart significantly from the Hardy-Weinberg equilibrium (HWE). Further tests indicated that heterozygote deficiency at these loci was responsible for this departure. Another possible

explanation for the departure from the Hardy-Weinberg equilibrium is the dramatic decline in *C. ectenes* spawning populations, and non-random mating or genetic bottlenecks.

CONCLUSIONS

In the present study, a microsatellite-enriched genomic library was constructed for *C. ectenes*, with a total of 20 novel genomic microsatellite DNA markers being developed. These microsatellite markers are expected to facilitate the management and exploration of the genetic resources of *C. ectenes*, and to assist in the genetic improvement of aquaculture bred populations.

ACKNOWLEDGMENTS

Research supported by the China Agriculture Research System (#CARS-50), the National High Technology Research and Development Program (#2012AA10A413), and the National Key Technology R&D Program (#2012BAD17B03).

REFERENCES

- Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.
- Chen F, Zeng L and Cheng Q (2012). Development of thirty-four novel polymorphic microsatellite markers in *Coilia ectenes* (Clupeiformes: Engraulidae) and cross-species amplification in two closely related taxa. *J. Genet.* 91: e37-e43.
- Cheng QQ and Han J (2004). Morphological variations and discriminant analysis of two populations of *Coilia ectenes*. *J. Lake Sci.* 16: 356-364.
- Cheng QQ, Lu DR and Ma L (2005a). Morphological differences between close populations discernible by multivariate analysis: a case study of genus *Coilia* (Teleostei: Clupeiformes). *Aquat. Living Resour.* 18: 187-192.
- Cheng QQ and Lu DR (2005b). PCR-RFLP analysis of cytochrome b gene does not support *Coilia ectenes taihuensis* being a subspecies of *Coilia ectenes*. *J. Genet.* 84: 307-310.
- Cheng Q, Cheng H, Wang L, Zhong Y, et al. (2008). A preliminary genetic distinctness of four *Coilia* fishes (Clupeiformes: Engraulidae) inferred from mitochondrial DNA sequences. *Genetika* 44: 395-400.
- Feng ZX (1997). Ichthyology. 2nd edn. China Agriculture Press, Beijing.
- Li Q and Wan JM (2005). SSRHunter: development of a local searching software for SSR sites. *Yi Chuan* 27: 808-810.
- Li YX, Xie SG, Li ZJ and Gong WB (2007). Gonad development of an anadromous fish *Coilia ectenes* (Engraulidae) in lower reach of Yangtze River, China. *Fish. Sci.* 73: 1224-1230.
- Liao MJ, Yang GP, Wang XC, Wang DQ, et al. (2007). Development of microsatellite DNA markers of silver carp (*Hypophthalmichthys molitrix*) and their cross-species application in bighead carp (*Aristichthys nobilis*). *Mol. Ecol. Notes* 7: 95-99.
- Liu W (1995). Biochemical and morphological comparison and interspecific relationships of four species of the genus *Coilia* in China. *Oceanologia et Limnologia Sinica* 26: 565-568.
- Ma CY, Cheng QQ, Zhang QY, Zhuang P, et al. (2010). Genetic variation of *Coilia ectenes* (Clupeiformes: Engraulidae) revealed by the complete cytochrome b sequences of mitochondrial DNA. *J. Exp. Mar. Biol. Ecol.* 385: 14-19.
- Ma CY, Cheng QQ and Zhang QY (2011). Development of 12 polymorphic microsatellite markers in *Coilia ectenes* Jordan and Seale, 1905 (Clupeiformes: Engraulidae) and cross-species amplification in *Coilia mystus* Linnaeus, 1758. *Environ. Biol. Fish.* 91: 243-249.
- Qu NN, Gong SY, Huang GJ, Tong JG, et al. (2010). Isolation and screening of microsatellite markers from the Chinese pearl oyster *Pinctada fucata* based on FIASCO. *J. Trop. Oceanogr.* 29: 47-54.
- Rice WR (1989). Analyzing tables of statistical tests. *Evolution* 43: 223-225.
- Rousset F (2008). Genepop'007: a complete reimplementation of the genepop software for Windows and Linux. *Mol. Ecol. Res.* 8: 103-106.
- Weber JL (1990). Informativeness of human (dC-dA)_n(dG-dT)_n polymorphisms. *Genomics* 7: 524-530.

- Whitehead PJP, Nelson GJ and Wongratana T (1988). Vol. 7. Clupeoid fishes of the world (Suborder Clupeoidei): An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies and wolf-herrings. Part 2- Engraulidae In: FAO species catalogue. FAO Fisheries Synopsis No. 125, Rome, 470-471.
- Yeh FC, Yang RC and Boyle T (1999). POPGENE version 1.31: Microsoft Windows-Based Freeware for Population Genetic Analysis. University of Alberta and Centre for International Forestry Research, Edmonton. Available at [http://www.ualberta.ca/~fyeh/popgene_download.html] Accessed April 20, 2011.
- Yuan CM, Lin JB, Qin AL and Liu HR (1976). On the classification history and status of genus *Coilia* in China. *J. Nanjing Univ. (Natural Science Edition)* 2: 1-12.
- Yuan CM (1980). On the classification of the anchovies, *Coilia*, from the lower Yangtze River and the southeast coast of China. *J. Nanjing Univ. (Mathematical Biquarterly)* 3: 67-82.
- Yuan CM and Qin AL (1984). Ecological Habits and distribution of *Coilia* along the Chinese coast and its changes of output. *Mar. Sci.* 5: 35-37.
- Zane L, Bargelloni L and Patarnello T (2002). Strategies for microsatellite isolation: a review. *Mol. Ecol.* 11: 1-16.
- Zheng JS, Liao XL, Tong JG and Wang D (2008). A pilot study on isolation of microsatellite DNA in the Yangze finless porpoise (*Neophocaena phocaenoides asiaeorientalis*). *Acta Hydrobiologica Sinica* 32: 25-31.