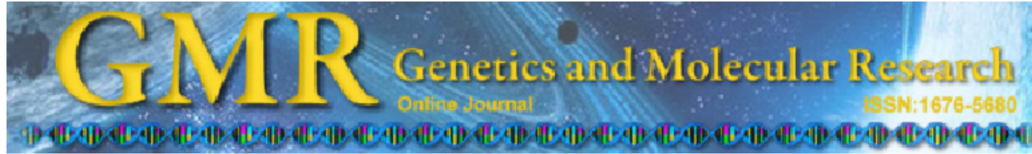

New polymorphic microsatellite loci of threadsail filefish, *Stephanolepis cirrhifer* (Teleostei, Monacanthidae), from Korean waters

C.-M. An, H.S. An, J.W. Lee and S.W. Hong

Biotechnology Research Division,
National Fisheries Research and Development Institute, Busan, Korea

Corresponding author: H.S. An
E-mail: hsan97@korea.kr

Genet. Mol. Res. 12 (2): 1679-1690 (2013)
Received November 12, 2012
Accepted April 24, 2013
Published May 14, 2013
DOI http://dx.doi.org/10.4238/2013.May.14.8

ABSTRACT. The threadsail filefish, *Stephanolepis cirrhifer* (Monacanthidae), is found mainly in the western Pacific. It is intensively caught in Korea and is a highly appreciated seafood delicacy. Consequently, the natural population of this species has drastically decreased, despite introductions from hatcheries. To provide information necessary for its conservation and management, we developed 24 polymorphic microsatellite markers using a combination of a total enriched genomic library and a small-scale 454 pyrosequencing. A total of 90,847 raw reads were obtained, and 75,128 unique sequences were generated, with an average length of 477 bp; 5350 (7.12%) sequences contained a minimum of 5 di- to tetranucleotide repeat motifs. Seventy-four sequences were used for microsatellite primer design. They all amplified successfully; 24 were polymorphic, with 8 containing trinucleotide repeats and 3 containing tetranucleotide repeats. The genetic variations based on 15 primer sets were investigated using 45 wild individuals from the coastal waters of Geomun Island. The number of alleles per locus varied from 4 to 15, with an average of

 ©FUNPEC-RP www.funpecrp.com.br

7.47. The observed and expected heterozygosities ranged from 0.333 to 0.956 and from 0.316 to 0.870, with averages of 0.692 and 0.701, respectively. No linkage disequilibrium was found between any pair of loci, indicating their independence. One locus significantly deviated from Hardy-Weinberg equilibrium after Bonferroni's correction; this may be due to the existence of a null allele. Cross-amplification was also tested for all 24 polymorphic loci in another monacanthid species, *Thamnaconus modestus*; 7 loci were effectively amplified. The high degree of polymorphism that was exhibited by the 15 newly developed microsatellites will be useful for assessing genetic variation and for conservation genetic studies of these 2 monacanthid species.

Key words: Korean threadsail filefish; *Stephanolepis cirrhifer*; Microsatellite loci; Next-generation sequencing; Genetic variability

INTRODUCTION

Stephanolepis cirrhifer (Temminck and Schlegel, 1850), which is commonly known as the threadsail filefish, is an economically important marine fish species in the family Monacanthidae. It is widely distributed in the western Pacific Ocean, including the Korean Peninsula, Japan, and the China Sea (Yoon, 2002). The annual catch of *S. cirrhifer* in the Northwest Pacific exceeded 250,000 tonnes in 1985, but it has declined continuously since then, amounting to less than 350 tonnes in 2002 (Garibaldi and Caddy, 2004). In Korea, it is intensively caught, where it is highly appreciated as seafood. Thus, the natural population of this species has drastically decreased as a result of overfishing and environmental changes (Yoon et al., 2011). To ease pressures on wild fishery stocks and to satisfy the growing levels of consumption, an early-stage effort has been made to develop the aquaculture of this species. Furthermore, the Korean government has sponsored the release of artificial juvenile threadsail filefish as predators to prevent blooms of the jellyfish *Aurelia aurita* (Miyajima et al., 2011). However, aquaculture practices are likely to reduce genetic diversity due to founder effects, inbreeding, and random genetic drift in hatchery-reared stocks. This situation has fostered a need for a more effective fishery management strategy based on solid scientific data. Information regarding the genetic variability and patterns of the stock structure is a prerequisite for developing effective fishery conservation strategies, management, and remediation efforts (Reiss et al., 2009). Molecular markers are a powerful tool for evaluating levels and patterns of genetic diversity (Liu and Cordes, 2004). Microsatellite (MS) DNA markers or simple-sequence repeats and variable mitochondrial DNA markers have long been utilized for population level studies and are of great utility in conservation and management (Christiakov et al., 2006; An et al., 2012a; Hong et al., 2012). Statistical methods for the analysis of population genetics often require highly variable molecular markers, such as MSs. Unfortunately, in spite of the commercial importance of the threadsail filefish in Korea, there are no reports on species-specific MS markers, and the literature describing its genetic background is scarce.

The genetic structure of *S. cirrhifer* in Korea has been investigated only recently, and there are just 2 studies involving the genetic characterization of *S. cirrhifer* in Korea. One study showed that the low variability of the cytochrome *b* gene of mitochondrial DNA is in-

adequate to detect genetic divergence between populations (Yoon et al., 2011). Another study reported that cross-species MS markers were able to detect significant differentiation between a wild population and a hatchery population of *S. cirrhifer* (An et al., 2011). MS markers are commonly used across species boundaries in closely related taxa (Athrey et al., 2007). However, such use is often complicated by the occurrence of null alleles due to failed amplification resulting from a polymorphism in the annealing site of one or both primers (Pompanon et al., 2005). Data sets containing null alleles will be deficient in heterozygotes, and such errors can result in incorrect parentage assignment or exclusion and biases in population genetic data (Dakin and Avise, 2004). It has also been indicated that further studies using species-specific MS markers will be necessary for a more reliable assessment of genetic diversity in *S. cirrhifer* (An et al., 2011). Thus, the genetic structure and information regarding the current exploitable potential of *S. cirrhifer* stocks have yet to be defined. Therefore, a set of species-specific polymorphic and co-dominant MS markers for *S. cirrhifer* must be developed for the administration of population genetic analyses as a means of devising conservation strategies because only a limited number of cross-species MS markers are available.

Traditional cloning-based methods for developing genomic MS markers involve significant trial and error (Queller et al., 1993) and as a result can be cost prohibitive to the development of large numbers of loci that are necessary for the detailed representation of genome-wide variation. There have been recent advances in sequencing technology and the accessibility of high-throughput genomic sequencing, and next-generation sequencing platforms, such as the 454 GS-FLX platform (Roche Applied Science, Indianapolis, IN, USA) that has been designed for whole genome sequencing, are valuable and cost-effective means of searching for genetic markers (including MS) in organisms for which adequate databases are not currently available (Abdelkrim et al., 2009). Hundreds if not thousands of MS loci can be identified from a fraction of a single next-generation sequencing run (Parchman et al., 2010). To date, this new technology has been successfully applied in the development of MS markers in many taxa, including marine organisms (An and Lee, 2012; Wang et al., 2012).

In the present study, we developed 24 novel polymorphic MS primer sets for *S. cirrhifer* using the 454 GS-FLX pyrosequencing, and 15 of the MS markers developed were characterized. Additionally, the applicability of these markers in another monacanthid species, *Thamnaconus modestus*, was evaluated via cross-species amplification experiments. The polymorphic MS markers that are described here will provide a valuable resource for future population genetic studies and fishery management of 2 monacanthid species, *S. cirrhifer* and *T. modestus*, in Korea.

MATERIAL AND METHODS

Sample collection and the 454 sequencing

A total of 45 wild threadsail filefish samples (body lengths of approximately 15 cm) were collected by direct sampling from the coastal waters of Geomun Island, Korea, in September of 2010. Caudal fin or muscle samples from the Korean threadsail filefish were clipped and immediately preserved in 99.9% ethanol until DNA extraction was performed.

For MS isolation, the TNES-urea buffer method (Asahida et al., 1996) was used to isolate high molecular weight DNA (≥2 μg) from the fin-clips of individual *S. cirrhifer* fish.

A whole-genome shotgun library was generated from 2 μg genomic DNA using the GS DNA Library Preparation kit (Roche Applied Science) according to the manufacturer protocol. The DNA library was titrated by sequencing using the Genome Sequencer FLX system (Roche Applied Science). On the basis of the results of the titration sequencing run, an appropriate amount of the DNA library was used to set up the emulsion PCR. Subsequently, the clonally amplified DNA fragments that were bound to the capture beads were enriched and sequenced on 1/8th of a plate in a Genome Sequencer FLX Titanium instrument (454 Life Sciences Corp., Roche Applied Science).

To characterize the MS loci via genotyping, total DNA from the fin-clip of each sample of *S. cirrhifer* was extracted using a MagExtractor-Genomic DNA Purification kit (Toyobo, Osaka, Japan) for an automated DNA extraction system, the MagExtractor MFX-2100 (Toyobo). The extracted genomic DNA was stored at -20°C until further use.

Microsatellite discovery and primer screening

The resulting raw sequences from *S. cirrhifer* were assembled into contigs using the Newbler software (ver. 2.3). A Perl script was run to select sequences that were longer than 300 bp with a minimum of 5 repeats of di-, tri-, or tetranucleotide repeat motifs. For these reads, a Perl script was also used to design the primers, and the following criteria were used to identify loci that could be reliably amplified: optimal primer length of 18-20 bp, optimal melting temperature of 60°C (range 58°-68°C), optimal GC content of 50% (range 30-90%), and low levels of self- or pair complementarity. Primer redundancy was tested using NCBI BLAST (http://ncbi.nlm.nih.gov/blast).

DNA amplification and genotyping

The newly designed PCR primer pairs were tested for consistency in PCR amplification and polymorphisms; these tests were performed on a sample set of 8 threadsail filefish collected from Geomun Island, Korea. PCR amplifications were performed using an ABI 9700 Thermal Cycler System (Applied Biosystems, Foster City, CA, USA) in a 25-μL reaction volume containing 12.5 μL 2X Multiplex PCR Pre-Mix (SolGent, Korea, Cat. No. SMP01-P096), 100 ng template DNA, and 10 pmol each primer. The forward primer from each pair was 5'-end-labeled with 6-FAM and HEX dyes (Applied Biosystems). PCR was run for 15 min at 95°C followed by 30 cycles of 20 s at 95°C, 40 s at 60°C, and 2 min at 72°C, with a 3-min final extension at 72°C. The annealing temperature of 60°C was 4-5°C below the Tm estimated from the nucleotide compositions of the primer pairs. The PCR amplification was considered to be successful based on the presence of a visible band after running 5 μL PCR product on a 5% denaturing agarose gel. The 1-kb Plus DNA Ladder molecular weight marker (SolGent, Cat. No. SDL54-B500) was used as a standard to assess the product size. If no amplification was detected, that primer set was excluded from further analysis by multiplex PCR. For further testing, a subset of polymorphic loci was selected based on the number of alleles detected, and genetic variation was examined in 45 samples collected. MS polymorphisms were identified using an ABI PRISM 3100 Automated DNA Sequencer (Applied Biosystems), and alleles were designated by PCR product size relative to a molecular size marker [GENESCAN 400 HD (ROX), Applied Biosystems]. Fluorescent DNA fragments

were analyzed using GENESCAN (ver. 3.7) and GENOTYPER (ver. 3.7) software packages (PE Applied Biosystems). The samples were multiplexed for genotyping by pooling samples tagged with different dyes within a well. We assessed the reliability of the primers by repeating the amplification and genotyping 16 samples.

Finally, all of the newly developed polymorphic MS loci in *S. cirrhifer* were assessed for cross-amplification in another monacanthid species, *T. modestus*, using 8 individuals collected from Jeju Island, Korea.

Genetic analysis

MICRO-CHECKER 2.2.3 (van Oosterhout et al., 2004) was used to detect genotyping errors due to null alleles, stuttering, or allele dropout using 1000 randomizations. To access genetic diversity, the number of alleles per locus (N_A), observed and expected heterozygosities (H_O and H_E, respectively), polymorphic information content (an indicator of the utility of the marker for linkage or population genetic studies), and inbreeding coefficient (F_{IS}) based on the allele frequencies pooled across all samples were determined using CERVUS version 3.03 (Kalinowski et al., 2007). Deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium were determined using Arlequin version 3.0 (Excoffier et al., 2005).

RESULTS AND DISCUSSION

Results of the 454 sequencing

The raw 454 sequence data from a 1/8th plate run included 27.9 Mbp containing 90,847 reads or sequences with an average length of 307.2 bp (maximum: 644 bp, minimum: 40 bp). Next-generation sequencing provides a wealth of genetic information, validating the potential value for both MS-related and MS-unrelated research (Greenley et al., 2012). The total coverage read length has been shown to be longer than 292 bp in the channel catfish *Ictalurus punctatus* (Jiang et al., 2011) but lower than 367 bp in the bream *Megalobrama pellegrini* (Wang et al., 2012). The different average read lengths may result from the total number of raw reads and different target materials that are subjected to sequencing. The raw sequences could be assembled into contigs. A total of 8951 reads were assembled into 769 contigs with an average length of 477.2 bp (maximum: 4835 bp, minimum: 100 bp). Most reads were unassembled as singletons, leaving 74,359 singletons (Table 1). The mean length of these 75,128 sequences (769 contigs plus 74,359 singletons) was 477.2 bp, which was longer than that of the raw sequences (477.2 bp *vs* 307.2 bp). This process eliminated repetitive sequences and created longer reads. Longer reads increase the likelihood that a single read will contain MS repeats along with suitable flanking regions of unique sequences (Lai and Sun, 2003). The lengths of the contigs that are generated depend on the depth of genome coverage (Farrer et al., 2009).

Isolation of MS loci

Of the 75,128 unique sequences, 5350 (7.12%) contained a minimum of 5 di-, tri-, or tetranucleotide repeat motifs, which were suitable for use as polymorphic MS markers. A total of 5350 MS comprised 67.36% dinucleotide repeats, 23.25% trinucleotide repeats and

9.39% tetra- to octanucleotide repeats (Figure 1). The high number of dinucleotide repeats in *S. cirrhifer* is consistent with that shown by previous studies in fish (Wang et al., 2012). Motifs containing 5 to 6 repeats were the most abundant (63.4%), followed by 7 to 9 repeats (28.5%) and those with over 10 repeats (8.1%).

Table 1. Summary of the 454 pyrosequencing.

Description	Dataset
Total number of bases (Mbp)	27.91
Average read length (bp)	307.2
Number of reads	
Total reads	90,847
Assembled	8,951
Singleton	74,359
Repeat number of contigs	167
Total contigs	769
Average contig read length (bp)	477.2
Max contig length/Min contig length (bp)	4835/100
Number of large contigs >500 bp	267

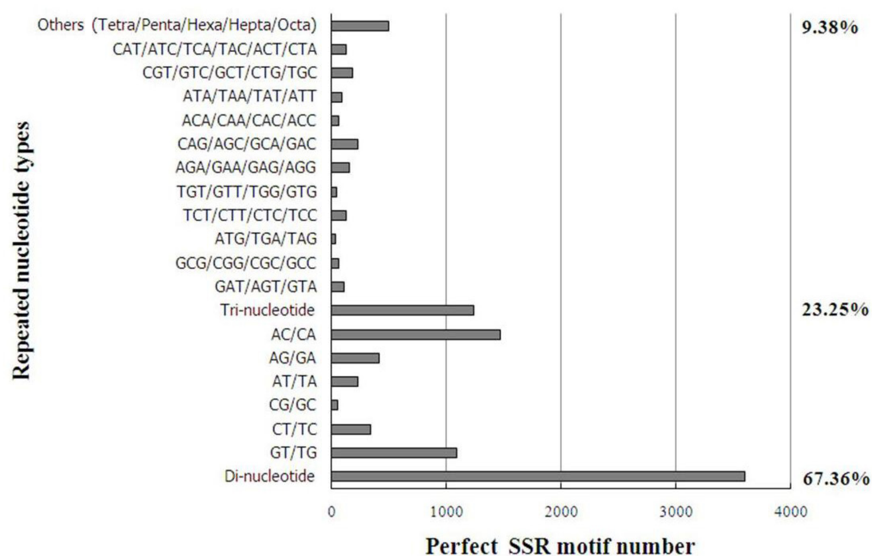

Figure 1. Distribution of simple-sequence repeat (SSR) nucleotide classes among different nucleotide types in *Stephanolepis cirrhifer*.

Among these, 74 sequences that were of adequate length (more than 300 bp) and unique sequences flanking the MS array (minimum of 100 bases) were selected for the primer design. All of the 74 (50 di-, 16 tri-, and 8 tetranucleotides) MS loci were amplified successfully (as viewed on agarose gel) during the initial evaluation of the MS primers. Of these, 32 (17 di-, 12 tri-, and 3 tetra-nucleotides) were selected based on the pattern of PCR products on an agarose gel for subsequent polymorphism screening. Further screening revealed that 24 (32.4%) loci

were polymorphic in the 8 *S. cirrhifer* samples. The primer sequences, repeat motifs, annealing temperatures, fluorescent labels, and GenBank accession numbers of the 24 new MS loci are summarized in Table 1. A homology search using the BLAST program showed that none of the 24 sequences were similar to any GenBank (Benson et al., 2011) sequence.

An increasing number of studies have used the 454 sequencing technology to identify MSs in marine organisms because it is more cost-effective, less labor-intensive, and more efficient than traditional methods (An and Lee, 2012; An et al., 2012b; Wang et al., 2012). In this study, 5350 sequences containing a minimum of 5 repeat motifs were detected for *S. cirrhifer* using the small-scale 454 pyrosequencing. We tested 74 primer pairs and characterized 24 polymorphic MS loci in *S. cirrhifer*, and 5276 additional sequences that contained MSs were also obtained. Thus, 7.12% of all of the unique reads that were obtained using the 454 sequencing contained MSs with a minimum of 5 di-, tri-, or tetranucleotide repeat motifs. Similar results were reported from other studies using an identical MS isolation method. MSs were detected in over 4.6% of the reads in the bream *M. pellegrini* (Wang et al., 2012), 2.26% in *Haliotis diversicolor supertexta* (An et al., 2012b) and 1.45% in *Mytilus coruscus* (An and Lee, 2012). Comparing the number of loci that were detected across different taxa is difficult because the genomes vary substantially in their MS frequencies. The numbers of MSs that are detected are likely to be inflated by the occurrence of multiple reads covering the same sequence. The frequency with which this occurs will be influenced by the genome coverage (Perry and Rowe, 2011).

A further advantage of the 454 sequencing is that many loci are detected, allowing the targeted selection of loci that are the most likely to amplify and exhibit polymorphisms. All of the 74 loci that we tested were successfully amplified, and a relatively high proportion (32.4%) of those that were amplified proved to be polymorphic. Of these, 13 polymorphic MS loci were dinucleotide repeats, 8 trinucleotide repeats and 3 tetranucleotide repeats. For most fish species for which MS markers have recently been developed, dinucleotide repeats such as $(CA)_n$ are still the predominant markers. Tri- and tetranucleotide MSs, however, have the advantage over dinucleotides of being highly polymorphic and more stable and showing clearer bands (Edwards et al., 1991; Lindqvist et al., 1996).

Genetic characterization

The samples of 45 natural *S. cirrhifer* that were collected from the coastal waters of Geomun Island, Korea, were screened for variation at the 15 novel polymorphic MS loci that were selected based on their polymorphisms in the 8 *S. cirrhifer* samples. The remaining 9 primers showed only 2 alleles and were thus eliminated from subsequent analyses. The 15 primer sets yielded variable profiles. Reruns were conducted for 35.6% of all individuals to ensure allele scoring reproducibility. The statistical results for the 15 novel MS loci are summarized in Table 2. Rare alleles with a frequency <5% were detected at most loci.

The 15 novel MS markers that were developed from *S. cirrhifer* vary widely in their degree of polymorphism. In total, 112 alleles were observed for the 15 loci; the number of alleles per locus varied from 4 at KSc10 to 15 at KSc5 (mean of 7.47; Table 2). The H_O ranged from 0.33 at KSc13 to 0.96 at KSc8 (mean of 0.70), whereas the H_E varied from 0.32 at KSc13 to 0.87 at KSc2 (mean of 0.69; Table 2). The polymorphic information content ranged from 0.29 to 0.85 (mean of 0.66), revealing high information content (>0.5).

Table 2. Characteristics of the 24 polymorphic microsatellite loci developed for *Stephanolepis cirrhifer* and their cross-amplification in another monacanthid species, *Thamnaconus modestus*.

Locus	Repeat motif	Primer sequence (5'-3')[a]		Ta (°C)	Population (N = 45)					GenBank accession No.	Cross-amplification *T. modestus* (N = 8)
		Forward	Reverse		N_A	H_O	H_E	PIC	F_{IS}		
KSc1	$(AT)_{11}$	Hex-CATGGTGTGTGAGGATAGAGTCAT	GCTTCAGCTAATTAAGACAATATG	60	5	0.821	0.782	0.735	-0.050	JQ678709	F
KSc2	$(CA)_{10}$	6-Fam-GAGGAAGAGCTTTATCTATCCTAC	GTGACAGAATGCAGAGAATTGACA	60	11	0.395*	0.870	0.846	0.549*	JQ678710	F
KSc3	$(TG)_{11}$	6-Fam-ACCACCACACACCGTTAGCTACAC	ACAATCGTCCGCTCATTCACTCTG	60	9	0.778	0.823	0.789	0.055	JQ678711	M
KSc4	$(CTT)_{12}$	Hex-GCAGTGATCGGCTTTTGCCAC	GCGTCTCAGCATCTGACTCAC	60	6	0.800	0.722	0.662	-0.110	JQ678712	M
KSc5	$(AAC)_{14}$	6-Fam-CGGAGGATGTGGTGGGAAGAA	CGGATCGATGCTAAAACCCTGA	60	15	0.800	0.758	0.731	-0.056	JQ678713	F
KSc6	$(ATC)_{10}$	Hex-ATCCATCCGTCCGTTCAGCAATCA	TATTGAAGCCGAGCCACGAGACT	60	7	0.644	0.691	0.642	0.068	JQ678714	F
KSc7	$(CT)_{12}$	6-Fam-GTGCAATAATCTGGCAACCTGC	GCCCACGTTGCTCGTTCGAAG	60	9	0.711	0.681	0.643	-0.043	JQ678715	P
KSc8	$(TG)_{10}$	6-Fam-GCCAACACTGCCCTCTATTGG	AGCCACTGAGGACGTGAAGTC	60	7	0.956	0.784	0.745	-0.222	JQ678716	F
KSc9	$(ACC)_9$	Hex-GAGTGGAAGCAGAGTGTAGTGGT	CTTGTCTTGCTCAATGGGTTACC	60	5	0.591	0.612	0.535	0.034	JQ678717	F
KSc10	$(CAT)_7$	6-Fam-CTGCTCTGTTACTACGGATCCT	ACACACAGGGAAGCCACACACTGG	60	4	0.429	0.371	0.346	-0.159	JQ678718	F
KSc11	$(TGC)_7$	6-Fam-ACCTGCTGCCAGTGTGTCAGCTGA	GAGCAACTTCAGGCCCTAAATCAAC	60	6	0.762	0.731	0.671	-0.043	JQ678719	P
KSc12	$(TC)_{10}$	6-Fam-GAGCAGGCAGGTTCTCACACATTG	GTGCTTCCTCTCGTTTCAGTTCAAG	60	5	0.722	0.772	0.720	0.065	JX468879	F
KSc13	$(GGCT)_6$	Hex-CGAAGACATGCTTGACATTCCACACA	AGCACTGTGGCACGCAGCATCTGC	60	5	0.333	0.316	0.289	-0.055	JX468880	F
KSc14	$(TGAG)_{12}$	Hex-GACACACTTAAAGTATGTGAGCACTT	CTCTTCACATAGATGCTAACTGCTAA	60	6	0.778	0.733	0.674	-0.062	JX468881	F
KSc15	$(TTC)_{13}$	Hex-GATTGGTTTTGGACGACGTGAG	CAGTTGCACAACCAGCAACACA	60	12	0.857	0.863	0.836	0.006	JX468882	M
KSc16	$(AT)_7$	6-Fam-TCTATTATGAGCACCAGCCAAGGA	CTCTTCCGCACAAGATATTGACAA	60	2					JX468883	F
KSc17	$(CA)_9$	Hex-ACTCTGACTTCTTCTGGAGGCTCT	ATGCAGTTGTGCACCTTAACTCTA	60	2					JX468884	F
KSc18	$(CA)_6$	6-Fam-TGCAGCTGCGGACTGACAGAAA	CAGTGTGACATTTAGACCAGTGGT	60	2					JX468885	F
KSc19	$(GT)_7$	Hex-CAATTATGAGGTCAACAGAAGTTT	GTGTTCAATAGCTGCAGAGGAAGA	60	2					JX468886	M
KSc20	$(GA)_7$	6-Fam-GTTGTGTAGCTGTAGTCGGCTAAA	GATTCAGTTGTTTGCATGACGGTC	60	2					JX468887	F
KSc21	$(CTA)_8$	6-Fam-ATCTGGCAATACTATTCTTGCTGC	TTTGCCACTGATTGTCACCATGCT	60	2					JX468888	M
KSc22	$(GTTT)_6$	6-Fam-AGCCACCAAGTTGACCTTTCTGGT	ATGATATGGTTCTTCATGTTCCTT	60	2					JX468889	F
KSc23	$(TC)_6$	6-Fam-TGTGAGCAAGGTTGCAACAGCTGA	ACAGACGAGCTGCTTTTGTTTGCA	60	2					JX468890	F
KSc24	$(CA)_5$	Hex-ACACCAGCTTGTTGTTGCAGGCGA	CCTACCTGAACCTCAGGTAAACAC	60	2					JX468891	F

[a]Primers were 5'-end-labeled with the indicated dyes. Ta = optimal annealing temperature; N_A = number of alleles per locus; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content; F_{IS} = inbreeding coefficient; M = monomorphic and F = failed to amplify or multiple non-specific amplification. Exact tests of Hardy-Weinberg equilibrium showed significant heterozygote deviation (*P < 0.003). Calculations assume that individuals with one microsatellite band are homozygous for the allele.

©FUNPEC-RP www.funpecrp.com.br

The development of powerful and efficient MS markers with a high genetic diversity is an essential step in the analysis of the genetic background of *S. cirrhifer*. In this study, the allelic polymorphism (mean N_A = 7.47) was lower than that reported for marine fish (mean N_A = 19.96 ± 6.6), and the heterozygosity values (mean H_E = 0.69 and mean H_O = 0.70) were slightly lower than those reported for other marine fish species (mean H_E = 0.77 ± 0.19) (DeWoody and Avise, 2000). The moderate H_E and low inbreeding indices that were observed in the present study indicate that the Korean threadsail filefish possesses moderate genetic diversity in large populations. Because the allele number was positively related to the sample size and the mutation rates at the polymorphic loci, the numbers of alleles that were observed at all 15 loci are certainly associated with the relatively small sample sizes (Liu et al., 2009). Nevertheless, similar genetic variabilities were reported for other marine species, including the yellow croaker (Wu et al., 2011) and Atlantic wolffish (Pampoulie et al., 2012), suggesting that these polymorphic MSs were sufficient to reveal the intraspecific diversity of the Korean threadsail filefish.

There was no evidence of genotyping errors or allele dropouts due to stuttering affecting the allele scoring. The samples that failed to amplify after the rerun were excluded, and thus, the likelihood that poor DNA quality affected the results was low. A significant departure from HWE after Bonferroni's correction ($P < 0.003$) was observed at KSc2, indicating that deviation from HWE was due to a heterozygote deficiency. Significant heterozygote deficiencies have been previously reported in marine organisms including other fishes (Sekino et al., 2002; Yue et al., 2004; Han et al., 2012). The presence of null alleles is a locus-dependent effect that is found frequently at MS DNA loci. Null alleles most likely cause heterozygote deficiencies in HWE tests (Callen et al., 1993). MICRO-CHECKER analysis revealed the existence of null alleles at 1 locus, KSc2, which had a significant heterozygote deficit. Further investigations are required to determine whether these null alleles are due to population subdivision (a Wahlund effect) or inbreeding. Regardless, the possible existence of nulls in the relatively small number of samples that were collected and tested cannot be excluded. Because this study was limited by the number of populations that were screened, the genetic diversity parameters and Hardy-Weinberg disequilibrium at KSc2 in the natural samples may be explained using data from additional populations, which would provide more precise estimates of the genetic characterization of the MS loci. Thus, our results should be interpreted with caution. The examination of linkage disequilibrium in all pairs of loci using the ARLEQUIN version 3.0 (Excoffier et al., 2005) likelihood-ratio test revealed that all 15 MS loci were in linkage equilibrium ($P > 0.05$).

The genetic diversity of *S. cirrhifer* at the 15 polymorphic MS loci that were identified in this study was moderate. Consequently, its populations are most likely declining, despite its moderate genetic diversity, and may become even smaller, given the continuously decreasing size of its wild population (Yoon et al., 2011). Thus, its genetic diversity should be protected, and there is an urgent need to create effective management strategies for the conservation of its natural populations.

Additionally, the cross-species amplification of 24 MS markers was performed in another monacanthid species, *T. modestus*. The 2 species that were studied in this experiment are the most important monacanthid species in Korea. Seven primer pairs (29%) of the 24 MSs effectively amplified the target sequences under the same PCR conditions; 4 primer pairs showed no polymorphic amplification. Although the cross-species transferability of MS

markers is limited, these results suggest the possibility that cross-species polymorphic markers can be developed in fish species using the 454 pyrosequencing technique. Generally, the number of cross-species-amplified loci tends to decrease in proportion to the increasing divergence between species (Peakall et al., 1998). One drawback of the MS markers is their high species-specificity, resulting in low cross-amplification success. Poor cross-species amplification of the MS DNA loci is due to the occurrence of null alleles due to failed amplification resulting from polymorphisms in the annealing sites of one or both primers (Pompanon et al., 2005). Furthermore, the rapid evolution of these markers renders many of them useless even in closely related species due to the absence of the repeat or low repeat number polymorphisms (Deitz et al., 2012). Thus, researchers should evaluate the need to develop species-specific MS markers for their research subjects if only a limited number of MS markers are available for the taxa.

In conclusion, the pyrosequencing method was applied to develop MS markers for *S. cirrhifer*, which is an economically important species in Korea. With this method, 74 sequences could be used for MS primer design and all were successfully amplified, of which 24 were polymorphic with 8 trinucleotide and 3 tetranucleotide repeats. The genetic diversity of *S. cirrhifer* at the 15 polymorphic MS loci that were identified was moderate in a wild population and may become even lower, given the continuously decreasing size of its wild population. Moreover, over 29% of the markers were successfully amplified in another monacanthid species, *T. modestus*. The polymorphic MS markers that were described here will be useful for future population genetic studies and fisheries management of the 2 monacanthid species in Korea.

ACKNOWLEDGMENTS

Research supported by the National Fisheries Research and Development Institute (NFRDI; contribution #RP-2013-BT-3). The views expressed here are those of the authors and do not necessarily reflect the views of NFRDI.

REFERENCES

Abdelkrim J, Robertson B, Stanton JA and Gemmell N (2009). Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* 46: 185-192.

An HS and Lee JW (2012). Development of microsatellite markers for the Korean mussel, *Mytilus coruscus* (Mytilidae) using next-generation sequencing. *Int. J. Mol. Sci.* 13: 10583-10593.

An HS, Hong SW, Kim EM and Myeong JI (2011). Comparative genetic diversity of wild and hatchery populations of Korean threadsail filefish *Stephanolepis cirrhifer* using cross-species microsatellite markers. *Genes Genomics* 33: 605-611.

An HS, Lee JW and Dong CM (2012a). Population genetic structure of Korean pen shell (*Atrina pectinata*) in Korea inferred from microsatellite marker analysis. *Genes Genomics* 34: 681-688.

An HS, Lee JW and Hong SW (2012b). Application of novel polymorphic microsatellite loci identified in the Korean Pacific abalone (*Haliotis diversicolor supertexta* (Haliotidae)) in the genetic characterization of wild and released populations. *Int. J. Mol. Sci.* 13: 10750-10764.

Asahida T, Kobayashi T, Saitoh K and Nakayama I (1996). Tissue preservation and total DNA extraction from fish stored at ambient temperature using buffers containing high concentrations of urea. *Fish. Sci.* 62: 727-730.

Athrey NR, Leberg PL and Klerks PL (2007). Laboratory culturing and selection for increased resistance to cadmium reduce genetic variation in the least killifish, *Heterandria formosa*. *Environ. Toxicol. Chem.* 26: 1916-1921.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. (2011). GenBank. Available at [http://www.ncbi.nlm.nih.gov]. Accessed 10, May 2011. *Nucleic Acids Res.* 39: D32-D37.

Callen DF, Thompson AD, Shen Y, Phillips HA, et al. (1993). Incidence and origin of “null” alleles in the (AC)n microsatellite markers. *Am. J. Hum. Genet.* 52: 922-927.

 ©FUNPEC-RP www.funpecrp.com.br

Christiakov DA, Hellemans B and Volckaert FAM (2006). Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. *Aquaculture* 255: 1-29.

Dakin EE and Avise JC (2004). Microsatellite null alleles in parentage analysis. *Heredity* 93: 504-509.

Deitz KC, Reddy VP, Reddy MR, Satyanarayanah N, et al. (2012). Limited usefulness of microsatellite markers from the malaria vector *Anopheles gambiae* when applied to the closely related species *Anopheles melas*. *J. Hered.* 103: 585-593.

DeWoody JA and Avise JC (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J. Fish Biol.* 56: 461-473.

Edwards A, Civitello A, Hammond HA and Caskey CT (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* 49: 746-756.

Excoffier L, Laval G and Schneider S (2005). ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinf.* 1: 47-50.

Farrer RA, Kemen E, Jones JD and Studholme DJ (2009). De novo assembly of the *Pseudomonas syringae* pv. *syringae* B728a genome using Illumina/Solexa short sequence reads. *FEMS Microbiol. Lett.* 291: 103-111.

Garibaldi L and Caddy JF (2004). Depleted Marine Resources: An Approach to Quantification Based on the FAO Capture Database. Fisheries and Aquaculture Organization, FAO Fisheries Circular, Rome, No. 1011: 18-19.

Greenley AP, Muguia-Vega A, Saenz-Arroyo A and Micheli F (2012). New tetranucleotide microsatellite loci in pink abalone (*Haliotis corrugata*) isolated via 454 pyrosequencing. *Conserv. Genet. Resour.* 4: 265-268.

Han HS, Nam BH, Kang JH, Kim YK, et al. (2012). Genetic variation in wild and cultured populations of the sea squirt *Halocynthia roretzi* inferred from microsatellite DNA analysis. *Fish. Aquat. Sci.* 15: 151-155.

Hong SE, Kim JG, Yu JN, Kim KY, et al. (2012). Genetic variation in the Asian shore crab *Hemigrapsus sanguineus* in Korean coastal waters as inferred from mitochondrial DNA sequences. *Fish. Aquat. Sci.* 15: 49-56.

Jiang Y, Lu J, Peatman E, Kucuktas H, et al. (2011). A pilot study for channel catfish whole genome sequencing and de novo assembly. *BMC Genomics* 12: 629.

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

Lai Y and Sun F (2003). The relationship between microsatellite slippage mutation rate and the number of repeat units. *Mol. Biol. Evol.* 20: 2123-2131.

Lindqvist AK, Magnusson PK, Balciuniene J, Wadelius C, et al. (1996). Chromosome-specific panels of tri- and tetranucleotide microsatellite markers for multiplex fluorescent detection and automated genotyping: evaluation of their utility in pathology and forensics. *Genome Res.* 6: 1170-1176.

Liu F, Xia JH, Bai ZY, Fu JJ, et al. (2009). High genetic diversity and substantial population differentiation in grass carp (*Ctenopharyngodon idella*) revealed by microsatellite analysis. *Aquaculture* 297: 51-56.

Liu ZJ and Cordes FJ (2004). DNA marker technology and their applications in aquaculture genetics. *Aquaculture* 238: 1-37.

Miyajima Y, Masuda R, Kurihara A, Kamata R, et al. (2011). Juveniles of threadsail filefish, *Stephanolepis cirrhifer*, can survive and grow by feeding on moon jellyfish *Aurelia aurita*. *Fisher Sci.* 77: 41-48.

Pampoulie C, Skirnisdóttir S, Daníelsdóttir AK and Gunnarsson Á (2012). Genetic structure of the Atlantic wolffish (*Anarhichas lupus* L.) at Icelandic fishing grounds: another evidence of panmixia in Iceland? *ICES J. Mar. Sci.* 69: 508-515.

Parchman TL, Geist KS, Grahnen JA, Benkman CW, et al. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics* 11: 180.

Peakall R, Gilmore S, Keys W, Morgante M, et al. (1998). Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Mol. Biol. Evol.* 15: 1275-1287.

Perry JC and Rowe L (2011). Rapid microsatellite development for water striders by next-generation sequencing. *J. Hered.* 102: 125-129.

Pompanon F, Bonin A, Bellemain E and Taberlet P (2005). Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* 6: 847-859.

Queller DC, Strassmann JE and Hughes CR (1993). Microsatellites and kinship. *Trends Ecol. Evol.* 8: 285-288.

Reiss H, Hoarau G, Dickey-Collas M and Wolff WJ (2009). Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish Fisher.* 10: 361-395.

Sekino M, Hara M and Taniguchi N (2002). Loss of microsatellite and mitochondrial DNA variation in hatchery strains of Japanese flounder *Paralichthys olivaceus*. *Aquaculture* 213: 101-122.

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

Wang J, Yu X, Zhao K, Zhang Y, et al. (2012). Microsatellite development for an endangered bream *Megalobrama*

pellegrini (Teleostei, Cyprinidae) using 454 sequencing. *Int. J. Mol. Sci.* 13: 3009-3021.

Wu XW, Liu XD, Cai MY, Ye XJ, et al. (2011). Genetic analysis of farmed and wild stocks of large yellow croaker *Larimichthys crocea* by using microsatellite markers. *Afr. J. Biotechnol.* 10: 5773-5784.

Yoon M, Jung J.-Y, Nam YK and Kim DS (2011). Genetic diversity of thread-sail filefish *Stephanolepis cirrhifer* populations in Korean coastal waters inferred from mitochondrial DNA sequence analysis. *Fish. Aquat. Sci.* 14: 16-21.

Yoon CH (2002). Fishes of Korea with Pictorial Key and Systematic List. Academic Seojeok, Seoul.

Yue GH, Li Y, Chen F, Lim LC, et al. (2004). Monitoring the genetic diversity of three Asian arowana (*Scleropages formosus*) captive stocks using AFLP and microsatellites. *Aquaculture* 237: 89-102.