Isolation and characterization of 21 novel polymorphic microsatellite loci in the Chinese soft-shelled turtle *Pelodiscus sinensis*

X.J. Bu, L. Liu, L. Wang and L.W. Nie

Life Science College, Anhui Normal University, The Provincial Key Lab of the Conservation and Exploitation Research of Biological Resources in Anhui, Wuhu, Anhui, China

Corresponding author: L.W. Nie
E-mail: lwnie@mail.ahnu.edu.cn

ABSTRACT. Twenty-one novel polymorphic microsatellite loci were isolated from the Chinese soft-shelled turtle, *Pelodiscus sinensis* (Trionychidae). This is a commonly consumed Asian species, especially in China, where it is reared in large numbers on farms. We screened 34 unrelated individuals from Wuhu of Anhui province in China and detected high levels of polymorphism for all 21 loci, with the number of alleles/locus ranging from 5 to 23 (mean 15.67). The expected and observed heterozygosities ranged from 0.421 to 0.946 and from 0.324 to 0.941, respectively. Hence, these microsatellites could facilitate studies on genetic diversity and population structure and marker-assisted breeding of this vulnerable species.

Key words: *Pelodiscus sinensis*; Microsatellites; Genetic diversity; Conservation

The Chinese soft-shelled turtle *Pelodiscus sinensis* Wiegmann, 1835, occurs throughout China (with the exception of the provinces of Xinjiang, Qinghai, Tibet, and Ningxia) and in northern Vietnam and Korea. It has been introduced to Thailand as well as Japan (Bonin Islands) and Hawaii (Bonin et al., 2006). *Pelodiscus sinensis* is favored for its medicinal value and has been consumed as food for a long period of time. However, in recent decades, populations of Chinese soft-shelled turtles have declined throughout this turtle's range due to unregulated harvest, habitat destruction, chemical pollution, and fragmentation of large rivers by dams in the wild (Altherr and Freyer, 2000). Consequently, the Chinese soft-shelled turtle was listed as a vulnerable species in the China Red Data Book of Endangered Animals in 1998 (Zhao, 1998) and the IUCN List category in 2000 (http://www.iucnredlist.org). To provide effective conservation and management of the once abundant wild populations, conservation managers are now looking to develop protocols for managing this species. In addition, captive populations of *P. sinensis* have been farmed since the 1980s, and it is possible that some captive populations have escaped from farms. Do these captive turtles have an effect on wild populations? However, little is known about the situation of this species in the wild environment. Highly polymorphic microsatellite loci will provide a powerful molecular toolkit and the statistical power of genetic analysis of the genetic diversity, population structure and molecular-derived breeding data in this important species. In this study, we isolated 21 novel microsatellite loci for *P. sinensis*.

Genomic DNA for marker development and subsequent surveys was extracted from *P. sinensis* muscle using the standard phenol-chloroform protocol (Sambrook and Russell, 2001). Microsatellite enrichment followed the protocols of Fischer and Bachmann (1998) and Bloor et al. (2001), with some modifications. A 400- to 1000-bp *Sau*3AI-restricted fraction was isolated from total genomic DNA, then ligated to linkers SauL-F (5'-GGC CAG AGA CCC CAA GCT TCG-3') and SauL-R (5'-PO$_4$-GAT CCG AAG CTT GGG GTC TCT GGC C-3'). The ligated fragments were hybridized to biotinylated (CA)$_{12}$, (AAT)$_8$, (AAC)$_8$, (ACAG)$_6$, and (GATA)$_6$ oligonucleotide probes, and the enriched fraction was recovered using streptavidin-coated magnetic beads (Roche). The eluted fragments were amplified using SauL-F as forward and reverse primers in a 50-µL reaction system composed of 2 µL template DNA (40-60 ng/µL), 2 U *Taq* DNA polymerase (TaKaRa), 5 µL 10X polymerase chain reaction (PCR) buffer (TaKaRa), 4 µL 25 mM MgCl$_2$, 4 µL 20 mM dNTPs, and 2 µL primer (SauL-F). PCR conditions were as follows: 95°C for 5 min, then 35 cycles of 94°C for 40 s, 60°C for 45 s, and 72°C for 60 s, with a final extension at 72°C for 10 min. The purified PCR products were ligated into pMD18-T vector (TaKaRa) overnight at 16°C, followed by transformation into *Escherichia coli* DH5α competent cells and plated onto LB/Amp/X-Gal/IPTG agar medium for blue/white selection. Recombinant clones were screened by amplification directly from bacterial colonies, using SauL-F and corresponding probe sequence as forward and reverse primers, respectively. A total of 228 positive clones were sequenced by ABI PRISM 3730 Sequencer. Based on the flanking sequences, 63 primer pairs were designed from the flanking sequences using PRIMER PREMIER 5.0 (http://www.premierbiosoft.com/). A total of 33 pairs giving consistent and specific PCR products were tested for polymorphism, and each forward primer of the microsatellite loci was labeled with a fluorescent dye (FAM, HEX or TAMRA) at its 5' end. Polymorphism was tested with 34 unrelated wild individuals of *P. sinensis* from Wuhu of Anhui province in China. Microsatellite DNA amplification reactions were performed in a PTC-200 (MJ Research) with 25-µL total volumes consisting of 1 µL template DNA (40-60 ng/µL), 1 U *Taq* DNA polymerase (TaKaRa), 2.5 µL 10X PCR buffer (TaKaRa), 2 µL 25 mM MgCl$_2$, 2 µL 20 mM dNTPs, and 0.5 µL of each 25 µM primer. The PCR cycles were as follows: an initial denaturation at 95°C for 5 min, followed by 35 cycles of 94°C for 40 s, 45 s at optimized primer-specific annealing temperature (Table 1), and

Table 1. Characteristics of the 21 Chinese soft-shelled turtle (*Pelodiscus sinensis*) microsatellite DNA loci.

Locus (Accession No.)	Repeat motif	Primer sequence (5'-3')	Size range (bp)	Ta (°C)	No. of alleles	H_O	H_E	PIC	P values (HWE)	Null freq
CST16 (HQ207543)	$(GATA)_{21}$	F:5'-TTCTGGTAAAACCGAAGTCT-3' R:5'-AAAAGTCCACGGAGTGTCTA-3'	231-307	55	16	0.882	0.922	0.901	0.0283	0.0137
CST21 (HQ207544)	$(GT)_{41}$	F:5'-TGGGGCAGTTGCTCTGAAGT-3' R:5'-CAAGGGGGATGTTACGGGA-3'	217-285	53	12	0.324	0.813	0.778	0.0000	0.4333
CST26 (HQ207545)	$(GATA)_{19}$	F:5'-TGTTGTTTTGTATGATGAGAA-3' R:5'-ATTTGCTGTAATAGTGGCT-3'	206-270	58	15	0.941	0.914	0.893	0.0161	-0.0216
CST27 (HQ207546)	$(CAAAA)_4 (CAA)_6$	F:5'-TGCCAATATAGTCTCAGTAGT-3' R:5'-GAAGGAAGTTGTTAGAACATC-3'	196-235	60	12	0.676	0.919	0.792	0.0015	0.0742
CST28 (HQ207547)	$(AG)_{11}(CATA)_{29}$	F:5'-GGAAAACACCTTTAGACCCA-3' R:5'-ACCTCAATCCACTGTGTCCA-3'	177-239	60	22	0.471	0.933	0.914	0.0000	0.3284
CST29 (HQ207548)	$(GATA)_{17}$	F:5'-GCTAAGTTTCCTATTTTTC-3' R:5'-GTTGTATGGTCCTTGTTTCT-3'	304-376	58	16	0.441	0.913	0.892	0.0000	0.3433
CST33 (HQ207549)	$(GATA)_{15}$	F:5'-ATCACCTGACCATCTGTTTC-3' R:5'-TTCTCTCTGCCATCTTACTT-3'	282-350	60	16	0.765	0.890	0.866	0.0005	0.0702
CST34 (HQ207550)	$(ATAG)_{16}$	F:5'-CCTTGAACGCACTCTGTCTA-3' R:5'-TTTTGCTCATAGTGTTTCCC-3'	225-297	58	16	0.882	0.910	0.888	0.0000	0.0057
CST35 (HQ207551)	$(GT)_{14}$	F:5'-GGAATAAGGGAAGTCAGAGG-3' R:5'-CAAAGGATACCATCACCAAA-3'	214-272	58	18	0.706	0.903	0.882	0.0034	0.1130
CST38 (HQ207552)	$(AAC)_9$	F:5'-TTAGAGGTAGTGGAAGACAG-3' R:5'-TTAGTTAGCCAATAAGGATA-3'	265-283	55	5	0.353	0.421	0.388	0.0026	0.1514
CST41 (HQ207553)	$(GATA)_{27}$	F:5'-TTACTGTCAAGTGCTGTCCC-3' R:5'-AGCGAATCCATAATCCTGTT-3'	220-288	56	16	0.647	0.908	0.886	0.0000	0.1668
CST42 (HQ207554)	$(GATA)_{20}$	F:5'-TGTCACCACACTACCTATTT-3' R:5'-GTGTTCTTAGTCCACTCCTC-3'	227-303	58	18	0.441	0.918	0.898	0.0000	0.3513
CST43 (HQ207555)	$(GT)_{14}$	F:5'-CGAAGGTGAAAAAAGGTGTT-3' R:5'-ATGTTTGTTTGCTGGGAAGG-3'	165-227	58	23	0.794	0.946	0.928	0.0000	0.0813
CST46 (HQ207557)	$(GATA)_{32}$	F:5'-GATCAATCGAGGGTGTGTG-3' R:5'-AAGCGGAATAGAGAAAACG-3'	164-268	62	21	0.647	0.924	0.904	0.0000	0.1686
CST47 (HQ207558)	$(GATA)_{20}$	F:5'-GAGGTTGTCTTATTCTCTGC-3' R:5'-TCACAAGCCTTACTAAAAAC-3'	211-287	60	15	0.853	0.918	0.897	0.3719	0.0296
CST48 (HQ207559)	$(CA)_{12}$	F:5'-CAAAGGATACCATCACCAAG-3' R:5'-GAATAAGGGAAGTCAGAGGC-3'	214-266	58	16	0.794	0.903	0.881	0.0094	0.0564
CST50 (HQ207560)	$(AGAC)_8$	F:5'-ATTTTTTTGGCATCCTTGTA-3' R:5'-CTCTTTCCGTCTGTCTTTT-3'	266-282	58	5	0.382	0.554	0.478	0.0089	0.1803
CST52 (HQ207561)	$(ATAG)_{13}$	F:5'-GGACAACATCCCAGTGAAATG-3' R:5'-CCAATCCCAAAGAAGGTAA-3'	219-283	60	15	0.824	0.918	0.897	0.0015	0.0474
CST54 (HQ207562)	$(TAGA)_{11}$	F:5'-TAGTCCTTTATTGAACCCTG-3' R:5'-CTTGAAAGAGCATCTCCATA-3'	244-340	58	18	0.735	0.907	0.886	0.0509	0.0975
CST57 (HQ207563)	$(GAAA)_{31}$	F:5'-TAAAGTGAGGCTAGAGAAG-3' R:5'-AAAAAATGAGTGAGTGATG-3'	305-389	58	19	0.912	0.923	0.903	0.0445	-0.0023
CST61 (HQ207564)	$(TAGA)_{21}$	F:5'-TAGACCAGGACCCCCTTATT-3' R:5'-TGATGGGATGTTGGTAGCAGT-3'	212-280	58	15	0.882	0.906	0.884	0.4589	0.0069

Locus designations (GenBank accession No.), repeat motif, primer sequences, size of amplified product (bp), annealing temperature (Ta), number of alleles, average observed (H_O) and expected (H_E) heterozygosities, polymorphism information content (PIC), probability (P) values for exact tests of Hardy-Weinberg equilibrium, and null allele frequency (Null freq).

72°C for 45 s, followed by a final extension at 72°C for 10 min. Amplified products were analyzed on an ABI PRISM 3730 Genetic Analyzer, with ROX 350 or LIZ 500 (Applied Biosystems) as the internal size standard. Microsatellite alleles were precisely sized using the Genemarker software (Applied Biosystems) to calculate their number, range and distribution.

We used CERVUS 3.0 (Kalinowski et al., 2007) to calculate H_O (observed heterozygosity), H_E (expected heterozygosity), mean polymorphic information content (PIC), and frequency of null alleles. Linkage disequilibrium (LD) and Hardy-Weinberg equilibrium (HWE) exact tests were conducted using web-based GENEPOP 4.0 (Raymond and Rousset, 1995). MICRO-CHECKER version 2.2.3 was selected to identify genotyping errors (Van Oosterhout et al., 2004).

In total, 21 microsatellite loci showed polymorphism. Locus designations, repeat motifs, primer sequences, range of allele sizes in base pairs, number of alleles per locus, PIC, H_O, and H_E are listed in Table 1. The number of alleles per locus ranged from 5 to 23 (average 15.67), showing a high PIC value of 0.840 (range 0.388-0.928). Values of H_E and H_O ranged from 0.421 to 0.946 (average 0.683) and from 0.324 to 0.941 (average 0.870), respectively, which indicate a relatively high level of genetic diversity in the Chinese soft-shelled turtle. Fifteen of the 21 loci (CST21, CST27, CST28, CST29, CST33, CST34, CST35, CST38, CST41, CST42, CST43, CST46, CST48, CST50, CST52) showed significant deviations from HWE (P < 0.01), which were revealed to result from null alleles by MICRO-CHECKER. Twenty-nine of the pairwise comparisons among loci (CST47-CST61, CST43-CST27, CST16-CST27, CST46-CST54, CST61-CST54, CST21-CST48, CST27-CST48, CST42-CST48, CST48-CST50, CST43-CST33, CST16-CST33, CST48-CST33, CST33-CST26, CST43-CST34, CST46-CST34, CST54-CST34, CST33-CST34, CST26-CST34, CST46-CST57, CST42-CST57, CST43-CST29, CST42-CST29, CST33-CST29, CST26-CST29, CST34-CST28, CST16-CST38, CST54-CST38, CST34-CST38, CST47-CST52) exhibited significant LD (P < 0.001). The small sample size was also the cause for the low power of the tests for LD and HWE.

In a recent study, Que et al. (2007) developed 15 polymorphic loci for the same species from Wuhan, Hubei province in China. These loci were mainly dinucleotide microsatellite loci, showing 2-7 alleles for each locus. The loci reported in the present study were more polymorphic than those of Que et al. and were mainly tetranucleotide microsatellite loci. As a result, this set of novel polymorphic microsatellite loci would provide a powerful tool for the population genetic studies of the Chinese soft-shelled turtle. Furthermore, these markers could serve as a potential source of microsatellites for marker-assisted breeding in the future.

ACKNOWLEDGMENTS

REFERENCES

Altherr S and Freyer D (2000). Asian turtles are threatened by extinction. *Turtle Tortoise Newsl.* 1: 7-11.
Bloor PA, Barker FS, Watts PC and Noyes HA (2001). Microsatellite libraries by enrichment. Available at [http://www.genomics.liv.ac.uk/animal/MICROSAT.PDF]. Accessed January 2009.
Bonin F, Devaux B and Dupré A (2006). Turtles of the World. Johns Hopkins University Press, Baltimore, 146-147.

Fischer D and Bachmann K (1998). Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *Biotechniques* 24: 796-800, 802.

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.

Que Y, Zhu B, Rosenthal H and Chang JB (2007). Isolation and characterization of microsatellites in Chinese soft-shelled turtle, *Pelodiscus sinensis*. *Mol. Ecol.* 7: 1265-1267.

Raymond M and Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity* 86: 248-249.

Sambrook J and Russell DW (2001). Molecular Cloning. 3rd edn. Cold Spring Harbor Laboratory Press, New York.

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.

Zhao EM (1998). China Red Data Book of Endangered Animals. In: Amphibia and Reptilia. Science Press, Beijing, 167-169.