Isolation and characterization of 13 new microsatellite markers in the triangle mussel (*Hyriopsis cumingii*)

Han-wen Yuan [1a]; Long-gen Guo [2 a;] Qiao-qing Xu [1 *]

Abstract

Microsatellite enriched library of *Hyriopsis cumingii* was constructed according to the strong affinity between biotin and streptavidin. One hundred clones with foreign inserts were sequenced and 65 clones were screened for usable microsatellites, of which 36 were deemed unique, of sufficient length (more than 8 repeats) and possessed adequate flanking regions for primer designment. Among 36 primer pairs designed, 25 yielded scorable amplification products. Upon testing 30 individuals were sampled using the 25 pair primers from Dongting Lake of Hunan Province, China. Then thirteen polymorphic microsatellite markers were isolated and characterized. These loci exhibited high levels of genetic polymorphism, so the observed number of alleles per locus ranged from 4 to 9. The ranges of observed and expected heterozygosity were 0.2543 to 0.8913 and 0.3629 to 0.8217, respectively, and the average polymorphic information content was 0.5198. Two microsatellite loci were significantly deviated from Hardy-Weinberg equilibrium due to the presence of null alleles, and no linkage disequilibrium found. These microsatellite loci will be useful for assessing the genetic diversity and population structure of *H. cumingii*.

Keywords: *Hyriopsis cumingii*, Microsatellite, Polymorphism, Genetic diversity

1- Engineering Research Center of Wetland Agriculture in the Middle Reaches of the Yangtze River, Ministry of Education, Yangtze University, Jingzhou, Hubei Province, 434025, China .

2- Donghu Experimental Station of Lake Ecosystems, State Key Labortory for Freshwater Ecology and Biotechnology of China, Institute of Hydrobiology, The Chinese Academy of Sciences, Wuhan 430072, China.

* Corresponding author's email: xuqiaoqing@163.com

a The two authors contributed equally to this work.

Introduction

In China, the production of freshwater pearl is high, and in 2008 it reached 1400 tons, approximately 95% of the world's total productivity. *Cristaria plicata* and *Hyriopsis cumingii* are two major species used for pear culture. *C. plicata* is widely distributed in many countries, such as China, Japan, Korea, and Vietnam (Dong and Li, 2004). *C. plicata* boasts unique advantages regarding its growth rate, disease-resistant ability, and productivity. However, due to its low pearl quality, it has been replaced by *H. cumingii* in the past 10 years for pearl production. *H. cumingii* is considered as a unique freshwater mussel for pearl production in China. Boasting many merits such as productivity, convenience of the pearl-producing operation, high efficiency of pearl production, and the high quality of the pearls yielded, *H. cumingii* is the dominant freshwater pearl-producing mussel in China. Although the industry of mussel cultivation and pearl production is developing and increasing numbers of pearls are being produced, the quality of the produced freshwater pearls continues to decline. Moreover, over-exploitation and changes in water quality decreased both the numbers and quality of freshwater *H. cumingii* (Li et al., 2006; Wang et al., 2007). In addition, the growth features of *H. cumingii* vary in different bodies of water. For instance, DNA analysis revealed that the genetic diversity of *H. cumingii* in two of China's largest freshwater lakes, Dongting Lake and Poyang Lake, is higher than that in other bodies of water (Hua et al., 2003; Li et al., 2005). To better protect and utilize the important freshwater mussel species, more polymorphic microsatellite markers are needed. Microsatellites or simple sequence repeats (SSRs) have been widely used as DNA markers in population genetic studies, parentage and kinship analysis because of their high level of polymorphism and codominant Mendelian inheritance (O'Connell and Wright 1997). In the present study, dinucleotide-enriched microsatellite genomic library of *H. cumingii* was constructed and 13 novel polymorphic microsatellite loci were developed from the library to permit the evaluation of genetic variation in future studies.

Materials and methods

Materials

Triangle mussels for construction of microsatellite genomic library were collected from a pearl culture farm in Gong'an, Hubei province of China and maintained at 18 °C in aerated tanks for a week before processing. Reagents used in the experiment were purchased from Progema (America) and streptavidin-coated beads were purchased from NEB (USA). Adapters were respectively *Mse* I A (5′-TACTCAGGACTCAT-3′), *Mse* I B (5′-GACGATGAGTCCTGAG-3′), *Mse*I-N (5′-GATGAGTCCTGAGTAAN-3′) and biotin-labeling probe (5′-biotin - (AC) 13-3′).

Construction of microsatellite enriched library

Total DNA was extracted from the feet and adductor muscles of five mussels by using the Wizard Genomic DNA Purification Kit (Promega, USA). Microsatellite isolation was performed following the process of Guo et al. (2009) with some modifications. Approximately 200 ng of total genomic

DNA was digested with *Mse* I and then ligated to the mixture of *Mse* I A and *Mse* I B. The treated DNA samples were then pooled and fragments were separated on a 1.2% agarose gel prior to size selection. The resulting fragments (400–1000 bp) were excised, purified with the E.Z.N.A™ Cycle-pure kit (Omega, USA), and amplified (23 cycles) with *Mse*I-N primers. The amplified DNA was hybridized with 250 pmol of biotin-labeling probe in a total volume of 100 μl of hybridization buffer (6× SSC and 0.1% SDS). The mixture was incubated at 95°C for 5 min, followed by annealing at 58°C for 30 min and gradual cooling to room temperature. During this hybridization, 30 μl (per treatment) of streptavidin-coated beads were resuspended in 400 μl of 1× hybridization buffer and washed 3 times. The hybridization mixture was added to the washed beads and incubated for 30 min at room temperature. The beads were washed twice with 1× hybridization buffer, thrice with high wash solution (0.2× SSC + 0.1% SDS) and twice with high wash solution at 10°C below the annealing temperature. The enriched fragments were recovered from the beads by resuspending the beads in 100 μl of 1× TE (pH 8.0) and heating to 95°C for 10 min. The enriched solution was immediately removed from the beads by using a magnetic stand. DNA-containing repeats were amplified with *Mse*I-N primers. The cycling conditions were as follows: an initial denaturation at 94°C for 5 min, followed by 23 cycles at 94°C for 30 s, 53°C for 1 min and 72°C for 1 min, with a final extension at 72°C for 10 min. The cleaned products were ligated into pMD18-T vectors (TaKaRa) and transformed into DH5a-competent cells. The recombinant clones were detected by PCR amplification with *Mse*I-N primers. Clones with foreign inserts were sequenced by an ABIPRISM 3730 automated sequencer (Invitrogen, USA).

Analysis of microsatellite sequence and primer design

Repetitive sequences were searched in the gained sequence using the software of Tandem Repeats Finder (Benson, 1999). Then pair primers were designed in the microsatellite sequences with the software Primer 5.0 and then used for the subsequent experiments..

Screening and analysis of polymorphisms microsatellite primers

PCR conditions were optimized for each pair of primers, and polymorphisms were assessed using 30 *H. cumingii* individuals that were collected from Dongting Lake of Hunan Province, China. Approximately 50 ng of template DNA were used in each 20-μl PCR reaction containing 0.5 μM of forward and reverse primers, 2.5 mM $MgCl_2$, 200 μM dNTPs, 1× Taq buffer (75 mM Tris-HCl, pH 8.8, 20 mM $(NH4)_2SO_4$, and 0.01% Tween 20), and 0.5 U of Taq polymerase (Ferments). The amplification procedure consisted of initial denaturing at 94°C for 5 min; 30 cycles at 94°C for 30 s, 52°C or 55°C for 45 s, 72°C for 1 min; and a final extension at 72°C for 10 min. PCR products were then separated on a 10% denaturing polyacrylamide gel and visualized by silver staining. The pBR322 marker (Tiangen) was used as a standard for scoring. Preliminary population genetics analysis was performed using Popgene version 1.32 (Yeh and Boyle, 1997) and Quantity-One software (Bio-Rad).

Deviations from Hardy-Weinberg equilibrium (HWE) and linkage equilibrium were determined using Popgene version 1.32. The frequency of the null allele at each locus was examined with Micro-Checker (Oosterhout et al., 2004). The polymorphic information content (PIC) was calculated using the following formula:

$$PIC = 1 - \left(\sum_{i=1}^{n} p_i^{\,2} \right) - \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2 p_i^{\,2} p_j^{\,2} \right),$$

where p_i and p_j are the frequency of the ith and j^{th} allele, respectively and n is the number of alleles (Botstein et al., 1980).

Results

Character of microsatellite enriched library
The recombinant clones were detected by PCR amplification with *Mse*I-N primers. An ABIPRISM 3730 automated sequencer (Invitrogen, USA) sequenced 100 clones with foreign inserts, of which 65 clones were screened for usable microsatellites. The average insert size was 320 bp and 95% of the inserts ranged from 149 to 452 bp, in which perfect microsatellite sequence was dominant with the percentage of 61.1%. In these microsatellite sequences, there were CT, GA, TCAC and TACA repetitive units except for the CA/GT. Length distribution of microsatellites in the genome of *H.cumingii* showed repetitive frequency concentrated on 11-30, which its percentage reached 80.95%, and the percentage was 11.09% for above 30 repetitive frequency. The biggest repetitive frequency reached 95 (Fig. 1).
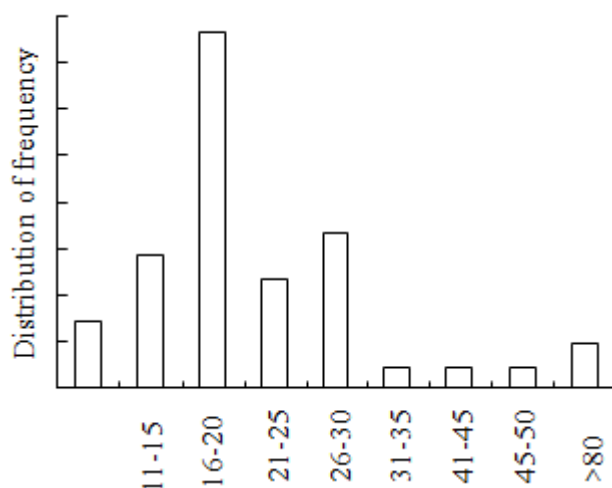
Figure 1: Length distribution of microsatellite in the genome of *H.cumingii*

Microsatellite characteristics.
Of 65 usable microsatellite sequences, 36 were deemed unique, possessing sufficient length (more than 8 repeats) and adequate flanking regions for primer design. These relevant sequences have been deposited in GenBank (GQ302634–GQ302668) (Table 1).

Table 1: Characterization of microsatellites and genetic diversity in 30 triangle mussels

Locus GenBank no.	Primer sequence (5′–3′)	Repeat motif	Expected size (bp)	Ta (°C)	N	K	H_O	H_E	PIC	FNA	Deviation from HWE
HC-clone1 GQ302634	F:CGTAATGTTGTCCCGATGAT R:ACCTCTACTGCTGGAGATTGTT	$(TG)_{19}$	305	52	30	6	0.6848	0.7926	0.5922	-0.1836	
HC-clone3 GQ302636	F:ATAGCGATCATTCTCACTGCAC R:TGCCTACTTTCAGTTCCATCAC	$(TG)_{21}$	199	55	30	5	0.3879	0.5109	0.3426	0.1287	
HC-clone6 GQ302639	F:ATGTCACTTGTTTGGCACCTC R:CAGTCCTCTTCAGGACTACAC	$(CA)_{19}$	162	52	29	7	0.6287	0.6829	0.5031	-0.0872	
HC-clone8 GQ302641	F:GCATTGTATAGCCAGGATTGT R:TGTCAGCAGGAGATGGATGG	$(CA)_{20}$	245	52	29	9	0.8913	0.7292	0.6974	-0.2091	
HC-clone12 GQ302645	F:GTGTGAATAGTAGATTATGGGTT R:TCAGGATAATGTAACTGGGTG	$(GT)_7GA(GT)_{10}$	165	55	28	5	0.4815	0.5608	0.4027	-0.1789	
HC-clone15 GQ302648	F:TTTTCTTTTCCGCTTTCATAC R:CAGAGGTGAGATGGCAGTGAT	$(CA)_{27}(TACA)_9$	257	55	29	9	0.9827	0.6791	0.7325	-0.5672	
HC-clone18 GQ302651	F:ATCCTCAACATCGCCCATAC R:TACTATACTCACATGTGCAA	$(CA)_{20}$	183	52	27	5	0.2543	0.7783*	0.4518	0.3872	*
HC-clone20 GQ302653	F:ATTTGTTGCCCGCCACTTT R:ATATCCCCAATTCCACCTTC	$(CA)_{21}$	106	52	30	4	0.5239	0.3629	0.3578	0.1768	
HC-clone21 GQ302654	F:AGAAACAAACCGCTGAAAGA R:GAATCAATGCGAGTAAAGTGG	$(TG)_{28}$	254	52	29	7	0.5417	0.6652	0.5982	-0.2419	
HC-clone24 GQ302657	F:ACTGTAAAATGGACCCGTGT R:CATAGAACTAAAGCCCGAAGA	$(CT)_{26}(CA)_{15}$	310	55	29	9	0.6238	0.7378	0.5931	-0.1487	
HC-clone25 GQ302658	F:ACGCAAGCAAAGGAGAATG R:GAATAGGCTCCGTGTCTTT	$(TC)_{14}GC(TC)_2TT(TC)_3$	124	52	28	7	0.4212	0.8217*	0.5129	0.4129	*

Table 1:Continued

					K	H_O	H_E	PIC	FNA
HC-clone26	F:AATAAGCCAAGAGGACGACA	(TC)$_{21}$	309	55	8	0.6269	0.6791	0.5946	-0.0872
GQ302659	R:TGCGGTTACAGATGAAAGAC								
HC-clone32	F:AGGAAAACAATGCTTCAGAT	(CA)$_{20}$	101	55	4	0.5432	0.3768	0.3781	-0.1389
GQ302665	R:AGTCAGTGCCCTATGTGTAAT								

Ta, annealing temperature; N, number of individuals successfully amplified; K = number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC= polymorphic information content; * Indicates significant deviation from Hardy–Weinberg equilibrium (HWE) after Bonferroni correction (adjusted P value = 0.005); FNA, null allele frequency.

PCR amplification with these 36 pair primers were conducted in the 30 individuals in the same population, in which no special bands were amplified using some primers (GenBank accession no: GQ302635, GQ302642 , GQ302650, GQ302662, GQ302667) in the 30 individuals and special bands were amplified with the primers (GenBank accession no: GQ302637, GQ302647, GQ302649, GQ302652, GQ302655, GQ302664) in partial individuals. So 25 pair primers were screened in the 36 pair primers.

Thirteen primer pairs of successfully amplified unambiguous PCR products were interpreted as possessing polymorphic allelic variation in the 25 pair usable primers (Table 1). The number of alleles per locus was 4–9 with an average of 6.54, the expected heterozygosity varied from 0.3629 to 0.8217 and the observed heterozygosity ranged from 0.2543 to 0.8913 (Table 1). The range of PIC was 0.3426 to 0.7325 with an average of 0.5198. Of the 13 polymorphic loci, 2 (HC clones 18 and 25) deviated from HWE in the tested population (adjusted P value = 0.005). An analysis using Micro-Checker v2.2.3 software (Oosterhout et al., 2007) suggested that these deviations were due to the presence of null alleles. The null allele frequency ranged from -0.5762 to 0.4129. No significant linkage disequilibrium was detected between any pairs of loci. These data suggest that high genetic diversity exists in the triangle mussel *H. cumingii* population.

Discussion

Microsatellite enriched library-combined biotin and streptavidin with beads are efficient markers in gaining microsatellite sequence due to their high efficiency, low cost and many usable microsatellite sequences. In the present study, 65 usable microsatellites were screened from 100 clones with the method. Among them, perfect was 61.1%, imperfect was 29.63% and compound was 9.26%, which was in accordance with the results in most of fish (Hao et al., 2006).

Microsatellites exist in every species genome, but repetitive motifs differ a lot. Dinucleotide repeat motif AC/TG was the most abundant in most animals. For example, (AC)n often existed in human genome and teleosts such as *Takifugu fasciatus*, *Ictalurus punctatus*, *Cyprinus carpio*, *Oncorhynchus mykiss*, *Ginglymostoma cirratum* and *Paralichthys olivaceus* (Brenner et al., 1993; Chen et al., 2005). For mollusks living in the same water as teleosts, there were many repetitive units. In the genome of *Mytilus galloprovincialis*, (AC/TG)n was the abundant repetitive motif except for the mononucleotide repeat and every 429 kb contains one $(AC)_{n>5}$ repeats. This result was also found in *Alasmidonta heterodon* (Shaw et al., 2006), *Margaritifera margaritifera* (Geist et al., 2003), *Mizuhopecten yessoensis* (Sato et al., 2005, *H.cumingii* (Li et al., 2007) and *C.plicata* (Jia et al., 2009). In other words, AC/TG motif was the most abundant dinucleotide repeat in eukaryote, while there was less frequency of dinucleotide and tetranucleotide in the genome of yeast and epiphyte (Tóth et al., 2000). Furthermore, the most common plant repeat motif was

AA/TT followed by AT/TA and CT/GA (Lagercrantz et al., 1993).

Ellegren et al. (2002) reported that repeat frequency was often less than 30 in eucaryote microsatellite sequences. In the study, repetitive frequency concentrated on 16-20, and its percentage reached 38.09%. This group comprised about 88.91 % of all microsatellites with a length of less than 30 repeats and the highest repeats frequency was 95, which was consistent with the result studied by Ellegren. All those showed that oligonucleotide probes with 13 repeats could screen enough high repeat of microsatellite sequence in *H.cumingii*.

H.cumingii is endemic to China and is the most productive freshwater pearl mussel. While mussel culture and pearl production are developing and more pearls are being produced, the quality of freshwater pearls is declining. Overexploitation and change of water quality also lead to germ degeneration in *H.cumingii* (Hua et al., 2003; Wang et al., 2007; Li et al., 2005). The study reported 13 microsatellite loci in *H.cumingii*. The high number of alleles, heterozygosity and polymorphic information content suggested their applicability for diverse purposes, e.g. in the determination of genetic diversity, the evaluation of genetic resource, and the study of population genetics of this mussel species.

Acknowledgments

References

Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research,* 27, 573-580.

Botstein, D., White, R. L., Skolnick, M., and Davis R. W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32, 314–331.

Brenner, S., Elgar, G., Sanford, R., Macrae, A., Venkatesh, B. and Aparicio, S., 1993. Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. *Nature*, 366, 256-259.

Chen, W., Zhang, Q. C., Yu, H. Y., Hu, J. J., Qi, J. and Bao, Z. M., 2005. Isolation and characterization of polymorphic microsatell ite loci in ol ive flounder (*Paralichthys olivaceus*). *Journal of Fishery Sciences of China*, 12, 682-687.

Dong, Z. G. and Li, L. J., 2004. Biodiversity and conservation of freshwater mollusks. *Acta Hydrobiologica Sinica,* 4, 440–444.

Ellegren, H., 2002. Microsatellite evolution: a battle between replication slippage and point mutation. *Trends Genet*, 2002, 18, 70.

Geist, J., Rottmann, O. and Schröer, W., 2003. Development of microsatellite markers for the endangered freshwater pearl mussel *Margaritifera margaritifera* L.(Bivalvia: Unionoidea). Molecular Ecology Notes, 2003, 3, 444-446.

Guo, B. Y., Xie, C. X., Li, H. J., Shan, L., Wang, H. L., Huo, B., Zhang, Y. and Zhou, Q., 2009. Isolation via

enrichment and characterization of 14 dinucleotide microsatellite loci in one catfish, *Glyptosternum maculatum* and cross-amplification in other related taxa. *Conservation Genetics*, 10, 547–550.

Hao, J., Sun, X. W., Liang, L. Q. and Lu, C. Y., 2006. Enrichment of large yellow croaker genome microsatell ite markers using magnet beads. *Journal of Fishery Sciences of China*, 13, 762-766.

Hua, D., Guo, R. B., Bai, Y.H. and Weng, H. B., 2003. Analysis of genetic diversity in cultured and wild freshwater pearl mussel using RAPD. *Journal of Fisheries of China*, 27, 540–544.

Jia, M. J., Li, J. L. and Wang, S. L., 2009. Isolation and characterization of 18 polymorphic microsatellite loci from freshwater pearl mussel (*Cristaria plicata*). Conservation genetics, 25, 120-127.

Lagercrantz, U., Ellegren, H. and Andersson, L., 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research*, 21, 1111-1115.

Li, J. L., Bai, Z. Y. and Qian, R.H., 2006. Comparative study on growth performance among the stocks of *Hyriopsis cumingii* from the five large lakes of China and the cultured stock from Zhuji city. *Fisheries Science and Technology Information*, 33, 243–246.

Li, J. L., Qian, R. H., Bao, B. L., Wang, G. L. and Wang, W. D., 2005. RAPD analysis on genetic diversity among the stocks of *Hyriopsis cumingii* from the

five large lakes of China. *Journal of Shanghai Fisheries University*, 14, 1–5.

Li, J. L., Wang, G. L., Bai, Z. Y. and Yue, G. H., 2007. Ten polymorphic microsatellites from freshwater pearl mussel, *Hyriopsis cumingii*. *Molecular Ecology Notes*, 7, 1357–1359.

O'Connell, M. and Wright, M., 1997. Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries*, 7, 331–363.

Oosterhout, C. V., Hutchinson, W. F., Wills, D. P. and Shipley, P., 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535–538.

Sato, M., Kawamata, K., Zaslavskaya, N., Nakamura, A. Ohta, T., Nishikiori, T., Brykov, V., Nagashima, K., 2005. Development of microsatellite markers for Japanese scallop (*Mizuhopecten yessoensis*) and their application to a population genetic study. Marine biotechnology, 7, 713-728.

Shaw, K. M., King, T. L., Lellis, W. A. and Eackles, M. S., 2006. Isolation and characterization of microsatellite loci in *Alasmidonta heterodon* (Bivalvia: Unionidae). Molecular Ecology Notes, 6: 365-367.

Tóth, G., Gáspári, Z. and Jurka, J., 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, 10, 967-981.

Wang, G. L., Yuan, Y., Li and J. L., 2007. SSR analysis of genetic diversity and phylogenetic relationships among different populations of *Hyriopsis*

cumingii from the five lakes of China. *Joural of Fisheries of China*, 31, 152–158.

Yeh, F. C. and Boyle, T. J. B., 1997. Population genetic analysis of codominant and dominant markers and qquantitative traits. *Belgian Journal of Botany*, 129, 157.