Identification and characterization of thirty novel microsatellite DNA markers from the Chinese mitten crab *Eriocheir sinensis* expressed sequence tags

Jingjing Li¹,², Xuyun Geng², Limei Chen³ and Jinsheng Sun¹,²*

**Abstract**

**Background:** The Chinese mitten crab *Eriocheir sinensis* is an economically important decapod crustacean in China. Despite a widespread distribution and production in China, the resources of *E. sinensis* have experienced a dramatic decline in the past decades. Here we describe a new set of novel polymorphic microsatellite loci to facilitate the investigation of genetic structure and artificial breeding.

**Results:** In this study, a set of 30 novel polymorphic microsatellite markers for *E. sinensis* was developed from EST databases. The number of alleles per locus ranged from three to twenty. The observed and expected heterozygosities ranged from 0.047 to 0.932 and from 0.047 to 0.935, respectively.

**Conclusions:** These informative microsatellite markers will be useful in studies of genetics, genomics and marker-assisted selection breeding in *E. sinensis*.

**Keywords:** Microsatellites, EST, Chinese mitten crab, *Eriocheir sinensis*

**Findings**

**Background**

The Chinese mitten crab *Eriocheir sinensis* is one of the most economically important aquaculture species in China [1] due to its taste and nutritious value, with a native range extending from the coastal estuaries of Korea in the north to the Fujian province of China in the south. However, the wild populations of *E. sinensis* have experienced a dramatic decline in the past decades due to overfishing and water pollution [2]. In China, the basic production technology of mitten crab populations has had a long history, with the conventional selective breeding programs based on phenotypic assessment. At present, the yield of *E. sinensis* is almost completely from artificial breeding. Unfortunately, like many other cultured species, the aquaculture performance of *E. sinensis* has declined significantly. In order to protect genetic diversity and prevent population degradation, understanding population genetic structure and genetic connectivity among populations and making a genetic linkage map are necessary.

Microsatellite markers provide a powerful tool in genome researches due to their wide distribution, codominant inheritance and high polymorphism. To date, approximately 83 microsatellite markers have been developed and applied for *E. sinensis* [3–8]. Although the number of described loci is relatively high, much more works is still needed because of the large diploid chromosome number of *E. sinensis* (*2n* = 146) [9]. In this study, we describe a new set of 30 EST-derived microsatellite markers which would aid in characterizing population structure, genetic diversity and constructing linkage map in *E. sinensis*. 
Experimental section

A total of 17067 *E. sinensis* ESTs obtained from the GenBank database (2013) were screened using SSRIT program [10] that was designed to find regions containing microsatellites. The parameters were set for detection of di-, tri- and tetranucleotide motifs with a minimum of six repeats, respectively. Eighty-five microsatellite loci were selected for microsatellite marker optimization. Primers flanking microsatellite were designed using the PRIMER PREMIER 5.0 program.

Sixty cultured *E. sinensis* individuals were randomly captured from Xieyuan Fishing Company in Qilihai region in Tianjin City, China. Genomic DNA was extracted from the leg muscles using a modified phenol-chloroform protocol [11]. Polymerase chain reaction (PCR) amplifications were performed in 10-μL volumes containing 0.25 U Taq DNA polymerase (Takara), 1× PCR buffer, 0.2 mM dNTP mix, 1 μM of each primer set, 1.5 mM MgCl₂ and about 100 ng template DNA. The PCR profiles for all loci were an initial denaturing at 94 °C for 3 min, followed by 35 cycles of 1 min at 94 °C, 1 min at the annealing temperatures listed in Table 1, and 1 min at 72 °C, with a final extension step of 5 min at 72 °C on a MJ Research PTC-200 DNA Engine (Peltier Thermal Cycler). Amplification products were resolved via 6 % denaturing polyacrylamide gel, and visualized by silver-staining. A 10-bp DNA ladder (Invitrogen) was used as a reference marker for allele size determination. The calculations of observed and expected heterozygosities were estimated with the program MICRO SATellite ANALYSE R software [12]. Tests for linkage disequilibrium and deviations from Hardy–Weinberg equilibrium (HWE) were performed using GENEPOP 4.2 [13, 14].

Results and discussion

Of the 85 potential microsatellite markers, forty-four loci were successfully amplified with the expected products. Thirty of them revealed polymorphism among the tested 60 individuals of *E. sinensis*. The number of alleles at each locus ranged from three to twenty with an average of 8.767 alleles per locus (Table 1). Observed heterozygosities ranged from 0.047 to 0.932 with an average of 0.527, while expected heterozygosities ranged from 0.047 to 0.935 with an average of 0.693. The mean number of alleles per locus, $H_O$ and $H_E$ demonstrated a relatively high genetic diversity within crab individuals. This was similar to reports from studies in other locations [3, 4, 6, 8]. Fourteen of the 30 loci significantly deviated from the Hardy–Weinberg equilibrium after Bonferroni correction. This might be due to the limited sample size, and/or the presence of null alleles at these loci. The high polymorphism of the loci suggests that they would be useful tool in studies of population structure, genetic diversity and the construction of genetic map for *E. sinensis*.

Conclusions

A set of 30 novel hypervariable microsatellite loci in *E. sinensis* was reported in this study. All the characterized microsatellite markers are suited for assessing the genetic diversity and the population structure, and also facilitate marker-assisted selection breeding of *E. sinensis*.

Ethics statement

Every effort was made to minimize animal pain, suffering and distress and to reduce the number of animal used. Sampling of the crabs was approved by Tianjin Diseases Prevention and Control Center of Aquatic Animals.

Availability of the supporting data

The microsatellite sequences are available through the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov); GenBank accession numbers see Table 1.
| Locus | GenBank accession no. | Repeat motif | Primer sequence (5′-3′) | $T_a$ (°C) | No. of individuals | No. of alleles | Size range (bp) | $H_O$ | $H_E$ | $P$ value |
|-------|----------------------|-------------|--------------------------|-------------|-------------------|----------------|----------------|--------|--------|-----------|
| ES27  | FL571941 (GT)$_{11}$ | F: TGTGATGAGAAGAAACCAAAGA  R: AATACCTGCTGGGGAAGTGA | 52           | 59                       | 8                | 107–123     | 0.712 0.797     | 0.0005*|
| ES39  | FG359711 (GT)$_{12}$ | F: AGGAGGAAAGTTGGAGGGG  R: AAATACAAATCTACGGGAGACAC | 55           | 56                       | 8                | 114–128     | 0.482 0.864     | 0.0000*|
| ES104 | FL569100 (TG)$_{12}$ | F: TCCAACACTACGAAAACCT  R: GAGTGTCAGTGATGAAAT | 48           | 53                       | 3                | 190–200     | 0.047 0.047     | 1.0000 |
| ES108 | FG984086 (GT)$_{10}$ | F: GTAACCCCTACGAAAACCTA  R: ACCTCCTAATACTACCTACCTACA | 54           | 59                       | 15               | 91–119      | 0.932 0.932     | 0.0000*|
| ES130 | FL570152 (GCG)$_{11}$ | F: CGTTGTTGAGGCGCTGTCG  R: GGCTGTCCTCCATCTACCTG | 57           | 60                       | 4                | 145–154     | 0.167 0.159     | 1.0000 |
| ES212 | FG984116 (CCCA)$_{6}$(GAA)$_{10}$ | F: GTGACACTGATGCTAGAGCA  R: TTACGCTTTCATGACCAAGAC | 55           | 56                       | 4                | 181–190     | 0.482 0.583     | 0.2114 |
| ES271 | FG359821 (TG)$_{12}$ | F: GCTCTCACCCTCCGTGATG  R: CTCTCTCTTTGCTTCTTTTA | 54           | 49                       | 6                | 176–186     | 0.615 0.754     | 0.0065 |
| ES352 | FL572494 (GT)$_{10}$ | F: CACTGGTACAAACATCAC  R: AATGGGATGTTAGTTAGTGT | 53           | 56                       | 17               | 91–127      | 0.768 0.929     | 0.0000*|
| ES582 | FL570362 (GA)$_{16}$ | F: ACCCTCAACGCCCTTACC  R: GAACAAACAGGGGAGCACAC | 53           | 58                       | 5                | 241–261     | 0.293 0.296     | 0.6958 |
| ES84  | FL574606 (CA)$_{22}$ | F: AGGGAAGTTGAAGTGAAGGA  R: ATGGGAATGAGATGAGGATAGA | 49           | 56                       | 9                | 222–240     | 0.429 0.863     | 0.0000*|
| ES645 | FL571837 (AC)$_{13}$ | F: GACCGCACAGAACAACCTC  R: CCACCTCTCAGTCAACGGGAAGA | 60           | 56                       | 11               | 122–158     | 0.522 0.912     | 0.0000*|
| ES709 | FL574505 (GCA)$_{7}$ | F: GCAGCCACCAACACCAGCGAAG  R: CTGGCCAATGAGCATCC | 60           | 60                       | 6                | 197–212     | 0.233 0.219     | 1.0000 |
| ES776 | FG357327 (CA)$_{34}$ | F: GTTGGTGTGGAAGGGAGCCA  R: CCTATACGGTGGCTGTCAC | 53           | 59                       | 4                | 204–220     | 0.410 0.557     | 0.1180 |
| ES789 | GE340666 (GT)$_{32}$ | F: TCGGCTGTTGGAGTGGTGG  R: AGCACTCACTGCTGCTGAGC | 52           | 53                       | 18               | 178–218     | 0.170 0.929     | 0.0000*|
| ES851 | GE340258 (CAC)$_{7}$ | F: TCCAACGCCGCGCAAAAG  R: AGCAAGTCCACCCAAACCAT | 54           | 54                       | 7                | 216–240     | 0.778 0.814     | 0.0580 |
| ES911 | FL569216 (TTCA)$_{12}$ | F: CGCGCGACACACTGAGAAT  R: CGAGGGTGAAGGAGCATT | 56           | 53                       | 5                | 242–258     | 0.453 0.634     | 0.0012*|
| Locus  | GenBank accession no. | Repeat motif | Primer sequence (5′–3′) | $T_a$ (°C) | No. of individuals | No. of alleles | Size range (bp) | $H_o$ | $H_e$ | $P$ value |
|--------|-----------------------|--------------|--------------------------|-----------|--------------------|---------------|----------------|--------|--------|----------|
| ES998  | FL572952              | (GT)$_1$(TG)$_3$(TG)$_{12}$ | F: CGACGGGTGTCAAGATTAGTGA<br>R: ACCACGGGCTCAAGAAGGCA | 56        | 51                 | 8             | 217–231        | 0.392  | 0.860  | 0.0000* |
| ES1045 | FL575077              | (GT)$_{38}$ | F: GGAAGACACCCAGTAAAGAAT<br>R: TCAACACGAAACCCGACCA | 55        | 55                 | 17            | 112–148        | 0.582  | 0.935  | 0.2415  |
| ES1053 | FL572054              | (CA)$_{10}$ | F: CTACACCAAGACCCCTGCCTG<br>R: GCCCTGCTGGTTGGTTAGCA | 57        | 58                 | 6             | 145–155        | 0.741  | 0.703  | 0.3476  |
| ES1126 | FG359126              | (AAGG)$_{12}$ | F: TGTCAGCTCTCCATGCA<br>R: TGGTATGGTTCGAATCTC | 54        | 60                 | 14            | 180–240        | 0.883  | 0.907  | 0.7140  |
| ES1139 | FL569992              | (GA)$_{11}$ | F: ACAACCCGACCTCCAGCA<br>R: TTGAACAAACACGGGACGAC | 55        | 59                 | 20            | 110–150        | 0.584  | 0.925  | 0.0000* |
| ES1171 | FG357452              | (AC)$_{3}$(TG)$_{6}$ | F: CAATCTGCCCTAATCTGCTT<br>R: GGAGAACCTAGGGAAGATA | 57        | 55                 | 8             | 156–172        | 0.900  | 0.871  | 0.0000* |
| ES1178 | GE341515              | (AC)$_{7}$ | F: TTCCTGCGGTTAGAACA<br>R: AGCCAGACTGCGCACGAC | 55        | 60                 | 3             | 138–146        | 0.150  | 0.144  | 1.0000  |
| ES1240 | FG982578              | (TACA)$_{11}$ | F: ATGTGACGACCTACGAGCAT<br>R: ACAACACCTCCAACGACG | 52        | 60                 | 13            | 152–200        | 0.883  | 0.890  | 0.3087  |
| ES1289 | FL574500              | (TTA)$_{13}$ | F: ACCTTGGGATACCCAGCAT<br>R: TCCCTTCAACACCATGAA | 50        | 55                 | 12            | 124–169        | 0.782  | 0.872  | 0.0000* |
| ES1293 | FL569028              | (GTA)$_{5}$(GTA)$_{3}$ | F: GCCTCATTACGGCGCTCTGAT<br>R: CCTCCCTGCACTTTCTCT | 55        | 59                 | 7             | 129–176        | 0.763  | 0.769  | 0.2826  |
| ES1300 | FL575122              | (TG)$_{10}$ | F: CCCCTGTTGATGCTGTA<br>R: GTGACAGGAAAGCCTCCTC | 55        | 52                 | 5             | 186–194        | 0.519  | 0.608  | 0.0846  |
| ES1482 | FL576108              | (TG)$_{5}$ | F: ACTATCCCTGGCTCACCTCTG<br>R: GACCAAAATTACCGTACGCG | 57        | 60                 | 7             | 115–129        | 0.250  | 0.522  | 0.0000* |
| ES1507 | FL572842              | (AGT)$_{3}$(TAG)$_{9}$ | F: TGAGTGTGCTGGTCTGCTG<br>R: TAGCAGATCCCTTCACCTTC | 57        | 58                 | 6             | 197–215        | 0.603  | 0.743  | 0.0299  |
| ES1513 | FL574610              | (AG)$_{8}$ | F: AACAGTTTGGCAGGAAACAGAAAAG<br>R: AGGGAAGGGATGAGTGTAGAGCA | 57        | 56                 | 7             | 100–114        | 0.217  | 0.739  | 0.0000* |

$T_a$ annealing temperature, $H_o$ observed heterozygosity, $H_e$ expected heterozygosity

* Significant departure ($P < 0.05$) from expected Hardy–Weinberg equilibrium conditions after correction for multiple tests ($k = 30$)
Authors’ contributions
JS was responsible for the design of this study, supervision of the work and contributed to the interpretation of results. JL performed field sampling, data analysis and marker validation, and drafted the manuscript. XG coordinated field sampling and was responsible for the implementation of the study. LC contributed to analysis of sequences. All authors read and approved the final manuscript.

Author details
1 Tianjin Key Laboratory for Animal and Plant Resistance, College of Life Sciences, Tianjin Normal University, 393 Binshuxi Rd, Xiqing District, Tianjin 300387, China. 2 Tianjin Key Laboratory of Aqua-Ecology and Aquaculture, Fisheries Science Department, Tianjin Agricultural University, 22 Jinjing Rd, Xiqing District, Tianjin 300384, China.

Acknowledgements
This work was supported by Grants of the National High-Tech Research and Development Program of China (863 programs, 2012AA10A401), National Key Technology R&D Program of Tianjin (12ZCDZNC05500) and Research and Extension Projects of Tianjin Fishery Bureau (J2013-21) (J2013-7).

Competing interests
The authors declare that they have no competing interests.

Received: 17 September 2015 Accepted: 3 February 2016
Published online: 17 February 2016

References
1. Zhang SY, Fu HT, Qiao H, Sun SM. Research progression genetics and breeding of Eriocheir sinensis. Chin Agric Sci Bull. 2013;29:39–45.

2. Li SF. The problem and measures of Enrocheir sinensis industry. Sci Fish Farm. 2006;6:1–2.

3. Hanfling B, Weetman D. Characterization of microsatellite loci for the Chinese mitten crab, Eriocheir sinensis. Mol Ecol Notes. 2003;3:15–7.

4. Chang YM, Liang LQ, Li SW, Ma HT, He JG, Sun XW. A set of new microsatellite loci isolated from Chinese mitten crab, Eriocheir sinensis. Mol Ecol Notes. 2006;6:1237–9.

5. Zhu ZY, Shi YH, Le GW. Isolation and characterization of polymorphic microsatellites from Chinese mitten crab, Eriocheir sinensis. Mol Ecol Notes. 2006;6:838–9.

6. Mao RX, Zhao Y; Liu FJ, Jia ZY, Hou N, Chang YM, Lu CY, Liang LQ, Sun XW. Development and characterization of new microsatellite loci from Chinese mitten crab (Eriocheir sinensis). Conserv Genet. 2009;10:1117–9.

7. Gao XG, Li HJ, Li YF, Sui LJ, Zhu B, Liang Y, Liu WD, He CB. Sixteen polymorphic simple sequence repeat markers from expressed sequence tags of the Chinese mitten crab Eriocheir sinensis. Int J Mol Sci. 2010;11:3035–8.

8. Xiong LW, Wang Q, Qiu GF. Large-scale isolation of microsatellites from Chinese mitten crab Eriocheir sinensis via a solexa genomic survey. Int J Mol Sci. 2012;13:16333–45.

9. Du NS, Lai W, Xue LZ. The chromosomes of the Chinese mitten-handed crab, Eriocheir sinensis (Crustacea, Decapoda). Zool Res. 1986;7:293–6.

10. Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S. Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res. 2001;11:1441–52.

11. Li Q, Park C, Kijima A. Isolation and characterization of microsatellite loci in the Pacific abalone, Haliotis discus hannah. J Shellfish Res. 2002;21:811–5.

12. Dieringer D, Schlötterer C. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes. 2003;3:167–9.

13. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumicism. J Heredity. 1995;86:248–9.

14. Rousset F. Genepop’007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Resour. 2008;8:103–6.