Genetics and Molecular Research

Genetic variations in the UTR of genes related to molting in *Eriocheir sinensis*

X.G. Gao[1], H.S. Pu[2], C.B. He[1], X.B. Bao[1] and W.D. Liu[1]

[1]Key Lab of Marine Fishery Molecular Biology of Liaoning Province, Liaoning Ocean and Fisheries Science Research Institute, Dalian, China
[2]School of Food Science Engineering, Dalian Polytechnic University, Dalian, China

Corresponding author: X.G. Gao
E-mail: xiangganggao@163.com

Genet. Mol. Res. 14 (4): 12472-12478 (2015)
Received May 3, 2015
Accepted July 24, 2015
Published October 16, 2015
DOI http://dx.doi.org/10.4238/2015.October.16.14

ABSTRACT. Crab grows by periodic molting, which is controlled by molt-inhibiting hormone (MIH) and ecdysteroids. Untranslated regions (UTRs) play crucial roles in the posttranscriptional regulation of gene expression. In this study, using crab collected from Changjiang (Yangtze), Huanghe (Yellow), Liaohe, and Yalujiang rivers, 33 haplotypes of the 3'-UTR of ecdysteroid-regulated protein (*ERP*) gene were identified, of which 14 haplotypes were observed in more than one individual. One hundred and forty clones of haplotype h2 (41.5%) were observed in samples from all the 4 rivers. Three haplotypes were observed to be insertions. For the whole crab sample, we found a positive Tajima's D value and a negative *Fu's* Fs test (Tajima's D value = 0.98726; *Fu's* Fs test = -27.382), although the P values were not significant (P > 0.10). The network profile of these 33 haplotypes presented a single core pattern with h2 as the core. In this study, we found that the UTR of *ERP* gene had a considerably high genetic polymorphism among crab from regions south to north of China. Furthermore, we observed a relatively high genetic divergence among different haplotypes, which would suggest a high diversity of the crab gene pool.

Key words: Chinese mitten crab; *Eriocheir sinensis*; UTRs; *MIH* gene; Ecdysteroid-regulated protein gene; Genetic variation

INTRODUCTION

Chinese mitten crab (*Eriocheir sinensis*) is a catadromous crustacean, native of East Asia, which lives predominantly in freshwater but migrates to sea for breeding. In China, *E. sinensis* is an economically important decapod crustacean cultured for its good taste. During its approximately 2-year life span, the crab grows by periodic molting, which is controlled by the molt-inhibiting hormone (MIH) and ecdysteroids (Keller, 1992). The primary role of MIH is to inhibit the synthesis and release of ecdysteroids. Ecdysteroids play an important role in blood glucose regulation and regulate several important physiological processes, including growth and reproduction (Burtis et al., 1990). Recently, the ecdysteroid-regulated protein (ERP), which controls the ecdysteroids, has been discovered (Buszczak et al., 1999; Sun et al., 2002). Untranslated regions (UTRs) play crucial roles in the posttranscriptional regulation of gene expression, including modulating the transport of mRNAs out of the nucleus, the efficiency of their translation, subcellular localization, and stability (Bashirullah et al., 2001; Tsai et al., 2002).

In view of the importance of UTRs, it is necessary to analyze their genetic variations in the *MIH* and *ERP* genes and to test whether these polymorphisms can be used as markers for selective breeding. In the present study, we detected a high level of genetic variation in the UTR of *ERP* gene in different populations of *E. sinensis*.

MATERIAL AND METHODS

DNA samples and data information

Adult Chinese mitten crabs were randomly collected from Changjiang (Yangtze; CJ), Huanghe (Yellow; HH), Liaohe (LH), and *Yalujiang* (YL) *rivers (*Table 1*).* Each population had 20 samples, with 10 females and males each. Leg muscle was taken from each individual and preserved in 80% ethanol.

Table 1. Samples, number of haplotypes, nucleotide diversity, and Tajima's D neutrality test in *E. sinensis* populations.

Population	Geo-coordinates	Abbrev-iation	Number of haplotypes	Nucleotide Diversity	Tajima's D	Statistical significance
Changjiang (Yangtze) River	31.14N, 121.29E	CJ	13	0.00743	0.76498	P > 0.10
Huanghe (Yellow) River	37.46N, 118.49E	HH	4	0.00229	-0.06501	P > 0.10
Liaohe River	41.07N, 122.03E	LH	15	0.00645	0.49333	P > 0.10
Yalujiang River	39.44N, 122.50E	YL	15	0.00460	0.68862	P > 0.10
Total			33	0.00655	0.98726	P > 0.10

DNA extraction, polymerase chain reaction (PCR) and sequencing

Genomic DNA was extracted from the leg muscle using the traditional phenol/chloroform extraction method (Sambrook and Russell, 2001). Primers were designed according to the published *E. sinensis* ERP and MIH sequences (GenBank accession No. GU443954 and AY310313; Table 2). Each reaction was conducted in a 25 µL volume containing 50 ng of genomic DNA, 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 200 nM of each primer, and 1U of La Taq polymerase (Takara, Dalian, China). The PCR program was as follows: initial denaturation at 94°C for 5 min followed by 35 cycles of 30 s at 94°C, 30 s at annealing temperature (see Table 2), and 40 s at 72°C, with a final extension step at 72°C for 10 min.

Table 2. Primers used in the PCR.

Primer	Sequence of primers (5'-3')	Product length (bp)	Amplification region	Annealing temperature
ERP	GGTGGTGACCTTCCAGTTG TTGTGGTCATTTTGTAGCCG	946	3'-UTR	58
MIH	GGAGTTCTGTCCACCGC CGTGCGTCCTCAAACTAA	465	5'-UTR	55

The purified PCR product from the leg muscle of each crab was cloned into pMD-18T simple vector (Takara, Dalian, China). Five–six clones were randomly selected from each ligation and sequenced in the forward direction using M13 forward primer (5'-CATGTGCTACTTCACCAAC GG-3'). DNA sequencing was performed with the ABI PRISM 310 Automated Genetic Analyzer (Applied Biosystems, Foster City, CA) using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, CA).

Data analyses

To determine the variable nucleotide sites and unique alleles, the sequences obtained from the DNA sequencing were aligned with the computer software ClustalX 1.8 (Thompson et al., 1997) and MEGA5 (Tamura et al., 2011), and the alignment corrected manually. All the polymorphisms were considered and DnaSP V5 (Librado and Rozas, 2009) was used to estimate the nucleotide diversity and to perform the neutrality tests, *Tajima*'*s* D (Tajima, 1989) and *Fu*'*s* Fs test (Fu, 1997). A haplotype network was constructed based on the polymorphic sites of the UTR of the *MIH* and *ERP* gene sequences using the Median-Joining method in the NETWORK version 4.6.1.3 computer software (Fluxus Technology Ltd, Clare, Suffolk, England).

RESULTS

Genetic variation in the 5'-UTR of *MIH* gene

This study characterized two fragments of the genes related to molting from a large number of crabs taken from different regions of China. Initially, characterization of the 5'-UTR of the *MIH* gene was performed for 40 crabs (10 individuals from each population). A total of 195 clones were sequenced and 465 bp sequences were obtained. All these sequences were identical to other known MIH sequences of *E. sinensis* (GenBank accession No. AY310313) indicating that the UTR of *MIH* gene is monomorphic in *E. sinensis.* Therefore, no further analysis was performed using this marker.

Genetic variation in the 3'-UTR of *ERP* gene

A total of 407 clones from 80 individual crabs were sequenced for the 3′-UTR region of the *ERP* gene. In all, 33 distinct sequences (haplotypes; h1-h33) were identified (GenBank accession No. KP780222-KP780254; Figure 1). Of the 33 haplotypes, 14 haplotypes were observed in more than one individual, whereas the remaining 19 alleles were only seen in one individual, but were observed in two or more clones.

Haplotype h2 was observed in all the four populations in 140 clones (41.5%), h6 in CJ, HH, and LH populations in 43 clones (12.8%), and h17 in HH, LH, and YL populations in 40

clones (11.7%). Three insertion haplotypes (76 Ins ATTTA，605 Ins CCTA, and 819 Ins ATCG) were observed in haplotypes h1, h3, and h4, respectively (Figure 1).

```
          1111 2334666667 7888888888 8889
     7777891235 3276000086 9012222355 6692
     6789009131 9664567869 3590129557 0193
h23  -----ATAC- CCGA----TA CG----AGGT --TT     9
h24  -----.A..T ....----.. ..----T... --..     2
h25  -----....T ....----AT A.----T... --..     2
h26  -----....- ....----.T A.----T... --A.     2
h27  -----....T ....----A. A.----T... --A.     2
h29  -----T..AA .T.G----A. A.----T..C --GA    12
h30  -----T..AA TTT.----A. A.----T..C --GA     2
h28  -----T..AA .T..----A. A.----T... --A.     2
h32  -----....- ....----.. AT----TT.C --GA     2
h33  -----....- ....----A. AT----TT.C --GA     2
h31  -----....T T.T.----A. AT----TT.C --GA     2
h21  -----....T TTT.----.. A.----T... --A.     4
h22  -----....T TT..----.. A.----T... --A.     2
h19  -----....T TTT.----A. A.----T... --A.     2
h4   -----.AT.T TTTG----AT A.ATCGT... --A.     3
h2   -----.AT.T TTTG----AT A.----T... --A.   140
h8   -----.AT.T TTT.----A. A.----T.A. --A.     3
h9   -----.AT.T TTT.----AT A.----T.A. --A.     2
h7   -----.AT.T TTTG----AT A.----T.A. --A.     3
h6   -----.AT.T TTT.----AT A.----T... --A.    43
h1   ATTTA.AT.T TTTG----AT A.----T... --A.     9
h5   -----.AT.T T.TG----AT A.----T... --A.     2
h13  -----.AT.T TTTG----.. ..----.... --..     2
h14  -----.AT.T TTT.----A. ..----.... --..     2
h12  -----.AT.T TTTG----A. ..----T... --..     6
h11  -----.AT.T TTTG----A. A.----T... --..     2
h10  -----.AT.T TTTG----A. A.----T... --A.    12
h3   -----.AT.T TTTGCCTAAT A.----T... --A.     3
h16  -----....T TTTG----AT A.----T... --A.     2
h17  -----....T TTTG----A. A.----T... --A.    40
h20  -----....T TTTG----.. A.----T... --A.    12
h18  -----....T TTTG----A. ..----T... --A.     2
h15  -----....T TTT.----AT A.----T... --A.     2
h34  -----....T TTT.----AT A.----T... ATA.  reported
```

Figure 1. Variations in the 3'-UTR sequences of the *ERP* gene from *E. sinensis* and one crab *ERP* sequence present in the database. The numbers of individuals sharing the same haplotype are listed in the right column. Dots (·) denote identical sites. The *ERP* genomic sequence of crab (GenBank accession No. GU443954, marked as “reported”) was also included for comparison.

Based on the 3'-UTR of the *ERP* gene, the nucleotide diversity of the four populations, CJ, HH, LH, and YL was 0.00743, 0.00229, 0.00645, and 0.00460, respectively (Table 1). The total nucleotide diversity was 0.00655. The *Tajima*'*s* D values for the four populations were 0.76498 (CJ), -0.06501 (HH), 0.49333 (LH), and 0.68862 (YL). For the whole crab sample, we found a positive *Tajima*'*s* D value and a negative *Fu*'*s* Fs test (*Tajima*'*s* D value = 0.98726; *Fu*'*s* Fs test = -27.382), although the P values were not significant (P > 0.10).

The network profile of these 33 haplotypes presented a single core pattern (with h2 as the core). Two main haplotypes (h6 and h17) were close to the haplotype h2. The remaining haplotypes had a low frequency and radiated from the central location. From h23 to h33, these haplotypes deviated from the core position (Figure 2).

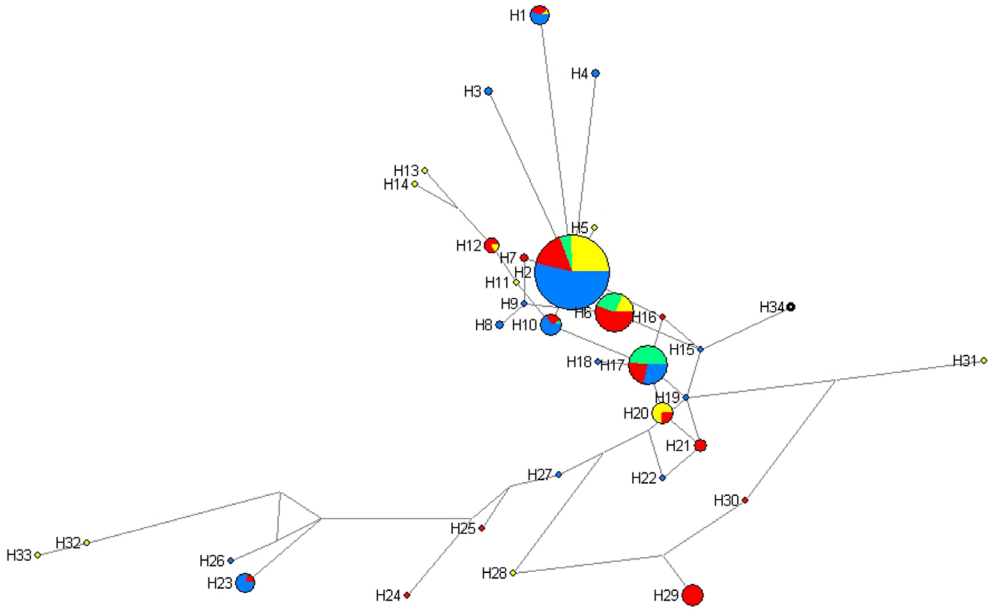

Figure 2. Sequence network profile of the 3'-UTR of the *ERP* gene of *E. sinensis* and one crab *ERP* sequence present in the database. The links are labeled by nucleotide positions to designate variations. The order of the variants on a branch is arbitrary. Each circle represents a haplotype, with the area of a circle being proportional to the frequency of the haplotype. Each color shows the geographical origin of the isolates. Yellow indicates isolates from CJ and green indicates isolates from HH. Red indicates isolates from LH and blue indicates isolates from YL.

DISCUSSION

Previous *in silico* analyses have shown that 3'-UTRs might contain AU-rich elements, RNA binding proteins, microRNAs, and other potent motifs, which may have biological functions (Barreau et al., 2005; Yu et al., 2014). Gene duplication is thought to be one of the major sources of evolutionary innovation (Conant and Wolfe, 2008). In the present study, 33 haplotypes were identified from the 3'-UTR of *ERP* gene from 80 individual crabs. The fact that all individuals had more than two haplotypes suggested that these haplotypes did not arise from the same locus and, thus, could be regarded as evidence for the existence of gene duplication at the UTR of the *ERP*

gene in *E. sinensis*. Gene duplications have been revealed in D-loop of mtDNA (Cjetva et al., 1992; Ludwig et al., 2000) and in many MHC loci, for example, the spotted seal (*Phoca largha*) DRB locus (Gao et al., 2015) and the cattle (*Bos taurus*) DQA and the DQB loci (Glass et al., 2000).

In this study, we found that the UTR of *ERP* gene had a considerably high genetic polymorphism among the Chinese mitten crabs from the south to north of China. Furthermore, we observed a relatively high genetic divergence among different haplotypes (Figures 1 and 2) which could suggest a high diversity of the crab gene pool.

In conclusion, in the present study, we characterized the genetic variation in the UTR of the *ERP* gene. In the future, we intend to investigate the correlation between polymorphism and growth traits to identify suitable candidate markers that may be related to crustacean growth and development. The results presented herein could potentially lead to the use of genetic strategies for improved crab breeding (Jia et al., 2014).

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (Grant #31302178).

REFERENCES

Barreau C, Paillard L and Osborne HB (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* 33: 7138-7150.

Bashirullah A, Cooperstock RL and Lipshitz HD (2001). Spatial and temporal control of RNA stability. *Proc. Natl. Acad. Sci. U. S. A.* 98: 7025-7028.

Burtis KC, Thummel CS, Jones CW, Karim FD, et al. (1990). The *Drosophila* 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two ets-rehted proteins. *Cell* 61: 85-99.

Buszczak M, Freeman MR, Carlson JR, Bender M, et al. (1999). Ecdysone response genes govern egg chamber development during mid-oogenesis in *Drosophila*. *Development* 126: 4581-4589.

Cjetva JB, Cook D and Zouros E (1992). Repeated sequences and large-sacle size variation of mitochondrial DNA: a common feature among scallops (Bivalvia: Pectinidae). *Mol. Biol. Evol*. 9: 106-124.

Conant GC and Wolfe KH (2008). Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet*. 9: 938-950.

Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915-925.

Gao XG, Han JB, Lu ZC, Zhang PJ, et al. (2015). Sequence variation and gene duplication at the MHC DRB loci of the spotted seal *Phoca largha*. *Genet. Mol. Res*. 14: 2055-2062.

Glass EJ, Oliver RA and Russell GC (2000). Duplicated DQ haplotypes increase the complexity of restriction element usage in cattle. *J. Immunol*. 165: 134-138.

Jia JL, Zhang LP, Wu JP, Ha ZJ, et al. (2014). Study of the correlation between GH gene polymorphism and growth traits in sheep. *Genet. Mol. Res*. 13: 7190-7200.

Keller R (1992). Crustacean neuropeptides structures, functions and comparative aspects. *Experientia* 48: 439-448.

Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

Ludwig A, May B, Debus L and Jenneckens I (2000). Heteroplasmy in the mtDNA Control Region of Sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). *Genetics* 156: 1933-1947.

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

Sun G, Zhu J, Li C, Tu Z, et al. (2002). Two isoforms of the early E74 gene, an Ets transcription factor homologue, are implicated in the ecdysteroid hierarchy governing vitellogenesis of the mosquito, *Aedes aegypti*. *Mol. Cell Endocrinol*. 190: 147-157.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol*. 28: 2731-2739.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 25: 4876-4882.

Tsai FJ, Lin CC, Lu HF, Chen HY, et al. (2002). Urokinase gene 3'-UTR T/C polymorphism is associated with urolithiasis. *Urology* 59: 458-461.

Yu D, Xu L, Peng L, Chen SY, et al. (2014). Genetic variations of mitochondrial antiviral signaling gene (*MAVS*) in domestic chickens. *Gene* 545: 226-232.