Acta Oceanol. Sin., 2019, Vol. 38, No. 8, P. 41–47
https://doi.org/10.1007/s13131-019-1428-0
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

Identification of SNP markers correlated with the tolerance of low-salinity challenge in swimming crab (*Portunus trituberculatus*)

Yanyan Feng[1], Dening Zhang[1], Jianjian Lv[1], Baoquan Gao[1], Jian Li[1]*, Ping Liu[1]

[1] Key Laboratory of Sustainable Development of Marine Fisheries of Ministry of Agriculture, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China

Received 19 July 2018; accepted 28 February 2019

© Chinese Society for Oceanography and Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Water salinity condition is an important factor for artificial propagation of the swimming crab (*Portunus trituberculatus*). Low salinity (LS)-resistant strains are preferred by crab industries. Single nucleotide polymorphism (SNP), the third generation of molecular markers, can be utilized in the breeding of LS-resistant species of *P. trituberculatus*. Our earlier study identified 615 genes differentially expressed in low-salinity stress compared to the controls. Although thousands of SNP loci have been found, it is hard to identify a SNP marker in correlation with a desired trait. In this study, time-of-flight mass spectrometry (TOF-MS), as an efficient method to select SNPs for the tolerance of LS challenge, was utilized for SNP typing. Fifty gene segments were amplified based on comparative transcriptomics in our earlier study, a total of 18 511 bp DNA fragments were amplified, and eighty-five SNP markers were found. The frequency of the SNPs was estimated to be 0.46 per 100 base pairs of DNA sequences. The rate of the conversion mutation was 81%, while the transversion mutation was 19%. The mutation rate of the G/T (C/A), A/T and G/C was 26%, 12% and 7%, respectively. Eight SNP markers were found to significantly correlate with the adaption of low salinity. Of the eight SNP markers, three linked-SNPs were found in the cuticle proportion gene, and another three SNPs were found in three new genes, and the rest two were found in aquaporin gene and chloride channel gene. The development of these SNP markers found in our study could be primarily used for breeding LS-resistant strains of *P. trituberculatus*.

Key words: *Portunus trituberculatus*, low salinity, time-of-flight mass spectrometry, single nucleotide polymorphism, SNP

Citation: Feng Yanyan, Zhang Dening, Lv Jianjian , Gao Baoquan, Li Jian, Liu Ping. 2019. Identification of SNP markers correlated with the tolerance of low-salinity challenge in swimming crab (*Portunus trituberculatus*). Acta Oceanologica Sinica, 38(8): 41–47, doi: 10.1007/s13131-019-1428-0

1 Introduction

Portunus trituberculatus belongs to Arthropoda, a commercially important fishery species widely distributed in the coastal waters of Japan, Korea and China (Dai et al., 1986). As a euryhaline species, it can survive in a wide range of salinity conditions though different salinity conditions can influence its distribution and migration route (Dai et al., 1977, 1986; Xue et al., 1997).

Salinity is one of the most important abiotic factors in aquaculture. Optimal salinity levels for growth, survival and production efficiency are often species-specific (Rouse and Kartamulia, 1992; Bray et al., 1994; Kumlu and Jones, 1995; Kumlu et al., 2001; Soyel and Kumlu, 2003; Ruscoe et al., 2004), while many crustaceans exhibit some degree of euryhalinity (Péqueux, 1995). Summer stormy weather increases mortality and leads to the degeneration of production traits, impacting the crab seriously in the past decades, which indicates the need of a strain that is resistant to low salinity. Some studies focusing on the salinity response of *P. trituberculatus* found that variable salinity could affect larval development and that high salinity condition would inhibit metamorphosis from zoea to megalops (Ji, 2005). Most of these studies concentrated on physiological characterization of *P. trituberculatus* only a little information on genomic response of swimming crab exposed to environmental salinity stress.

The SNP has many advantages that other markers do not possess, such as wide distribution, high density and correlation with traits. With the development of DNA-based marker analysis and high-throughput genotypic technologies, SNPs have become the best choice of markers for large scale mapping and genotyping (Rafalski, 2002; Black IV et al., 2001). There have been many reports of finding SNPs in crustaceans, such as *Litopenaeus vannamei* (Yu et al., 2014), *Scylla paramamosain* (Ma et al., 2011) and *P. trituberculatus* (Li et al., 2013). But it remains difficult to identify the SNPs that are associated with important traits. Genes that are associated with important traits are becoming the resources for the identification of SNP markers in some species lacking whole genome sequencing. Maybe it is more important to identify SNPs based on transcription data because those genes correlate with traits significantly (Yu et al., 2014; Jin et al., 2014). The high throughput computational methods for detection of

Foundation item: The Modern Agro-industry Technology Research System under contract No. CARS-48; the program of Shandong Leading Talent under contract No. LJNY2015002; the National Natural Science Foundation of China under contract No. 31472275; the Qingdao Industrial Development Program (Science and Technology Benefit Special Project) under contract No. 17-3-3–62–nsh.
*Corresponding author, E-mail: lijian@ysfri.ac.cn

SNPs and small INDELs (insertion/deletion) have been used widely in the field of molecular markers.

A number of techniques for SNP genotyping have been reported, including those based on allele-specific PCR (Sommer et al., 1992), single-base extension (Syvänen, 1999), probe hybridization (Livak et al., 1995; Piatek et al., 1998; Sauvage et al., 2007; Schütz et al., 2000), Tm-shift (Germer and Higuchi, 1999; Wang et al., 2005) and microarray technology (Hirschhorn et al., 2000), and others (Kwok, 2001). Each method has its advantages and disadvantages compared with the rest in regard to cost, specificity and throughput. There is no clearly superior method that has been established. Time-of-flight mass spectrometry (TOF-MS) is built on multiplex PCR. Its working principle is based on the difference of base weight (G>A>T>C). The signals strength was amplified in experiment, and then tested by mass spectrometer.

In our previous study, transcript sequencing of the *P. trituberculatus* that was exposed to two different salinity challenges (11 and 33) were performed (Lv et al., 2013). Our cDNA microarray data suggest that there are gene expression patterns of *P. trituberculatus* for low salinity acclimation, and that a series of genes are potential key elements for low salinity environment. Comparison of gene expression revealed a total of 615 unigenes that were differentially expressed between normal salinity condition and low salinity challenge. GO functional enrichment analysis revealed that some differentially expressed genes were associated with crucial processes in osmoregulation, such as ion transport processes, amino acid metabolism and synthesis processes, proteolysis process and chitin metabolic process. In the present study, we screened out and amplified the candidate gene fragments related to low salinity regulation based on comparative transcriptomics of *P. trituberculatus* in our earlier study (Lv et al., 2013), discovered SNP markers, typed SNPs by TOF-MS, and analyzed the association between SNP and LS-resistance. These results will increase the efficiency of the selective breeding of *P. trituberculatus*.

2 Materials and methods

2.1 *Portunus trituberculatus*

Swimming crabs at age of 80 d with average body weight of 50 g were collected from Changyi Haifeng Aquaculture Co., Ltd. (Shandong, China). A total of 600 crabs from ten families, each with sixty crabs, of fast-growing strain of Huang Xuan No.1 in an outdoor pond were used only for the low salt stress experiment. The salinity of seawater in outdoor culture pond is 23. All crabs were acclimatized at 26.0°C in aerated seawater for 5 d prior to the salinity challenge. The crab stocking density was 3 per cubic meter. They were fed with *Potamocorbula laevis* twice a day. Seawater was refreshed every other day. The low salinity challenge experiment was carried out at three stages and acclimated to low salinity challenges in turn: Salinity 11 for 7 d, 6 for 7 d, and 4 for 10 d. During acclimation, the crabs were checked for death every 2 h. Myoideum was collected from the dead individuals and all those survived at the end, and stored in –20°C freezer for later analysis.

2.2 *Filtering gene segments by differential expression and GO enrichment analysis*

Differential expression analysis was carried out with the DEGseq (2010) R package. The *P* value was adjusted with *q* value (Storey and Tibshirani, 2003). The *q* value less than 0.005 and $|\log_2(\text{fold change})| > 1$ was set as the threshold for significantly differential expression. The smaller the *q* value is, the more relevant the gene fragments are to osmoregulation. Based on comparative transcriptomics of *P. trituberculatus* in our earlier study (Lv et al., 2013), 615 differentially expressed genes (DEGs) were found.

Generally, RPKM (gene expression level) value is 0.1 or 0.01 as the threshold value to judge if the gene is expressed. Different thresholds are adopted in separate documents. For instance, it would be regarded as a gene with low abundance expression level if its RPKM is between 0.1 and 3.57, and middle abundance expression level for RPKM from 3.57 to 15; and high abundance expression level for RPKM more than 15. GO enrichment analysis of the DEGs was performed using GOseq-based Wallenius non-central hyper-geometric distribution, which was adapted for gene length bias in DEGs. In combination with GO enrichment analysis, the RPKM of unigenes were selected as high as possible. And then the lowest 50 unigenes were chose according to the *q* value.

2.3 *DNA isolation*

Genomic DNA was extracted from belly muscle by a standard protocol using proteinase K digestion and phenol–chloroform extraction. The quality and concentration of DNA were assessed with agarose gel electrophoresis and measured with Nano Photometer® spectrophotometer (IMPLEN, CA, USA). DNA was adjusted to 100 μg/mL and stored at –20°C.

2.4 *Primer design, polymerase chain reaction (PCR) and DNA sequencing*

We compared our unigenes with *P. trituberculatus* sequences from NCBI and found conserved regions for primer design with PrimerPremier 5.0. The primers were synthesized by Shanghai Biological Technology Company (Shanghai, China).

Total DNA was used as the template for PCR amplification of the target sequence with a PCR supermix kit (TaKaRa, Dalian, China).

PCR products were purified and sequenced with ABI3730 DNA Analyzer at the Sangni Company (China). We ran two PCR reactions using one pair of primers on two different DNA templates: one was from susceptible individuals and the other was from resistant ones. Both PCR products were purified and mixed for the sequencing in the same tube.

2.5 *Analysis of sequencing results and search for SNPs*

By identifying double peaks in the chromatogram of sequencing, we found variations in the same SNPs of the two PCR products mentioned above. For the identification of the putative SNPs, we found the distinct double peaks in the Atlas of chromatograms.

2.6 *Genotyping of SNPs*

The DNA of 90 crabs was used for typing. TOF-MS, a method based on different time of flight among four bases (A, G, T, C) in the magnetic field, was used for SNP typing. SNPs identified were genotyped with the Mass ARRAY platform (Sequenom, San Diego, CA, USA) following the protocols and recommendations provided by the manufacturer. Briefly, the technique consists of an initial locus-specific PCR, followed by single-base extension using mass-modified dideoxynucleotide terminators of an oligonucleotide primer that anneals immediately upstream of the polymorphic site (SNP) of interest. The distinct mass of the extended primer identifies the SNP allele. MALDI-TOF mass spectrometry analysis in an Autoflex spectrometer was used for allele scoring.

2.7 *SNP association with LS-resistance*

A chi-square test was applied to the analysis of SNP and LS-resistance association. For each SNP site, one-way ANOVA was conducted to determine whether there was a significant difference in osmoregulation among crabs with different genotypes. Statistical analysis and data log transformation were carried out using the software SPSS 17.0. The distributions of allele frequencies of each SNP site were calculated by SPSS 17.0 and compared by the Pearson's chi-square test or Fisher exact test to confirm their associations with the susceptibility/tolerance to salinity challenge. $P<0.05$ was considered as statistically significant.

3 Results

3.1 *Mortality rate of different low salinity challenges*

Mortality of *P. trituberculatus* with different low salinity challenges: Salinity 11 for 7 d, 6 for 7 d, and 4 for 10 d, is shown in Table 1. It shows that the lowest salinity caused the most death among the three conditions tested. Myoideum was collected from the deceased individuals and all those survived until the study was completed.

Table 1. Mortality rate of the *P. trituberculatus*

	Salinity			Remained alive individuals	Total
	11 (7 d)	6 (7 d)	4 (10 d)		
Count	70	62	329	103	564
Mortality rate/%	12.41	11.00	58.33	0	81.74

3.2 *Gel analysis of genomic DNA isolated from the swimming crab*

The quality of DNA was assessed with agarose gel electrophoresis. Main bands of genomic DNA with a slight degradation were observed (Fig. 1). It shows that the quality of DNA meets the needs of follow-up study.

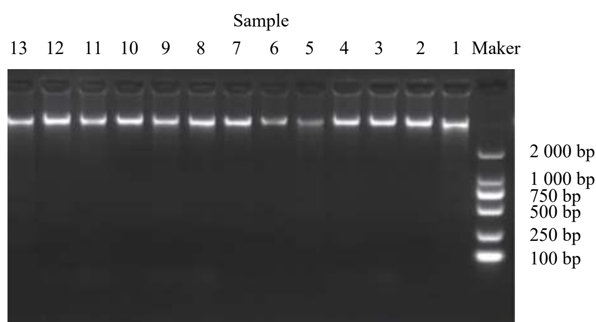

Fig. 1. Gel electrophoresis of *P. trituberculatus* genome DNA.

3.3 *Analysis of sequencing results and search for SNPs*

The results of primer design and PCR are shown in Table 2. For identification of putative SNPs, we conducted DNA sequencing and then found distinct double peaks in the atlas of chromatograms.

In this research, we identified 85 SNP sites in a total of 18 511 bp fragments of DNA, which indicates the SNP frequency is 0.46 per 100 bp DNA. The proportion of C/T or G/A conversion was 55.29%. The rate of the G/T (C/A), A/T and G/C was 26%, 12% and 7%, respectively.

3.4 *SNP tying by TOF-MS and SNP polymorphism*

Among 30 SNPs selected for tying, 23 SNPs were successful with TOF-MS. We used POPgene to analyze gene polymorphism. As shown in Table 3, the observed and expected heterozygosities range from 0 to 0.780 5 and from 0 to 0.506, respectively. And the SNP marker Comp68973_c0(382), comp55325_c0(384), comp55325_c0(485), comp65582_c0(421), comp43413_c0(75), comp65155_c0(158), comp73731_c0(150), comp73731_c0(339), comp3(563), comp6(1 199), and comp206(2 463) deviate significantly ($P<0.05$) or very significantly ($P<0.01$) from HWE. Except for comp65582_c0(163) and comp6(1 199), other locus' MAF was larger than 0.05, which illustrated these SNPs had high polymorphism and are suitable for association analysis.

3.5 *SNP association with LC-resistance*

Previous studies have shown that synonymous coding SNPs are likely to become involved in disease mechanisms. So synonymous coding SNPs discovered from GWAS (Genome Wide Association Study) should also be discussed with functional studies (Chen et al., 2010). To identify the SNP frequency, the concatenation of the 795 sequences was done to generate a consensus sequence of 142 711 bp. The precondition of one-way ANOVA is that concomitant probability needs to be larger than 0.05. Some sites do not meet the criteria of one-way ANOVA.

The haplotype frequencies of isolate comp55325_c0(384), comp68972_c2(289), comp77848_c0(75), comp73731_c0(150), comp73731_c0(339), comp8(1 269), and comp206(2 463) significantly differed ($P<0.05$), while others showed minimum association between the SNPs and osmoregulation. Typing of 23 SNPs were successful and 8 SNPs were related to osmoregulation, shown in Tables 4 and 5.

4 Discussion

In this research, we identified 85 SNP sites in a total of 18 511 bp fragments of DNA with a frequency of 0.46 SNPs per 100 bp. This result was similar to that reported for *Scylla paramamosain* (0.3/100 bp) (Ma et al., 2011), but higher than those reported for sperm whale (0.19/100 bp) (Morin et al., 2007) and fineness porpoise (*Neophocaena phocaenoides*) (0.18/100 bp) (Li et al., 2009), and lower than scallops (*Aequipecten opercularis* and *Mimachlamys varia*) (1/100 bp) (Arias et al., 2009). The proportion of base transition is 81.18% and the proportion of transversion is 18.82%. The ratio of transition/transversion (1:0.23) in this study was lower than those in scallops (Arias et al., 2009), zebrash (*Danio rerio*) (Stickney et al., 2002), Pacific oyster (*Crassostrea gigas*) (Sauvage et al., 2007), and Pacific salmon (*Oncorhynchus tshawytscha*) (Smith et al., 2005). There is a large difference in SNP density among different species, which may be related to the complexity of the growing environment. C/T or G/A conversion is the greatest mutation types, because the 5-methyl C to T transition mutation occurs about 10 times as frequently as other transitions. The higher level of C/T (G/A) SNPs is probably related to 5-methylcytosine deamination reactions that are known to occur frequently, particularly at CpG dinucleotides (Holliday and Grigg, 1993).

In this study, we found that the proportion of A+T is greater than G+C before the sites mutated, which suggests the A/T bases mutated much more often than G/C bases. This can be explained through thermodynamics. In DNA, a C base has hydrogen-bond with a G base and an A base with a T base. These interactions keep the two DNA strands together. There are one more hydrogen bonds between G and C than A and T. So a G paired with a C can be more stable than an A with a T. We found the proportion of A+T decreased by calculating the G+C to A+T ratio in the mutational sites after mutation. There had been a deviation with (G+C) / (A+T). The study of drosophila and mammals non-

Table 2. List of the primers used in the study

Location	Primer sequence (5′ to 3′)	Design length/bp	Objective strap size/bp	Intron or not
comp65155_c0	CTTCGGCTGGGGCGGTTACTCC	310	310	No
	CTCGTTGGAGCGGAGGTTGGTC			
comp68972_c2	CGGCGGTCAATGGCTAAACGGA	490	490	No
	GTGGAGCGAACACGGAAACAGG			
comp64836_c0	GACGATGCTTGCTCAGGGCTAATG	561	410	No
	TTCTCAGTACCACGCCCAAGACG			
comp37438_c0	AGAAAACCGCCCCTAATGGTAAGATC	775	1 100	Yes
	CAGGTTGCTGCCAGGGGTGATG			
comp65155_c0	CTCGTTGGAGCGGAGGTTGGTC	783	1 133	Yes
	GGTGGCGTTCTACAGTTCTGGTGC			
comp64816_c0	ACGGCATAGGCTACAGCAAGTGG	776	874	Yes
	CGCCACATCCTACTGCGACAAAA			
comp71227_c2	ACGGTTGGTTGGTGACAGCCTAG	230	230	No
	TCAGCTTCCAAAGAGGTCGAGTGA			
comp68973_c0	GCTGTTGGTGGGCTTCGTGGCG	520	520	No
	CGTGGGTAGTGTCGTGATTGAAGG			
comp77722_c0	CAAGGCTTTCTGGCTCTGTGGG	840	845	No
	CCTCCATAGACGGGATGAGAACGAA			
comp55325_c0	ATGAGTTTACGTTCACCCCGCTGTTC	709	1 094	Yes
	CGCTTGCGGTTCTGAGGGTAAGTC			
comp77179_c1	GCTGCGGCTGCGTGGCTGGAAT	464	921	Yes
	CACTTGGAGGCTCGGAACTTGTCA			
comp77561_c0	AGCGTGTATTGGGTTTGTCTGGC	503	506	No
	GCCCTGCTGCCCCACTCCCTTA			
comp65582_c0	AGCATCAAAAGCAACGCCAACG	705	963	Yes
	GTGTCAAAGGCACTGATGGTGGAAAC			
comp68972_c2	TGTAGGCGACGCACCAGACAGC	237	527	Yes
	CGCAGGTTCACGATCAACTACGG			
comp43413_c0	AATGGAGCCGTTGTACCTGTAGGC	56	315	Yes
	TGTGCGGGCTACGGAAGCACTA			
comp41025_c1	ACTTCACAAAGACTGGCAGACACCC	314	317	No
	TGGACTCCCAGAGTTAAGGATCAGC			
comp77448_c1	TCGTCAGGCAGAGGCAGGAAGG	549	612	Yes
	TTTGGGTTTGGTTGGTCAGGGA			
comp76555_c2	GGTTTGTGGTCTTTATTATTTCGTG	780	548	No
	TCTTCCATCGAGGTCTGAACAGG			
comp69000_c0	ACACGCATCATGGTGAGATTCAG	667	708	Yes
	TTGGCAGTAATCCAATCCACAAAGT			
comp71809_c0	AACAAATCATCATTGCTTCCCTG	531	825	Yes
	GAGGAGCCTTGTCGGTTTATTGC			
comp73192_c1	GCCTGAGAAACGGCTACCACATC	939	939	No
	ACCAGACAAATCGCTCCACCAAC			
comp73731_c0	CCAGTAGCCGCATCTTGGAGTGC	875	917	Yes
	AAGCCATAATTCCTTGTATCGTTGTAC			
comp77682_c0	TGTTCTCGATGTTATCCGCAAAG	892	927	Yes
	TTCGGCTTCAGTGAACTCCATTT			
comp71227_c1	CATCGCTGGACAGGAGCCTGGGC	420	891	Yes
	TCGAACTGGGACTCCAAGCCTGC			
comp54927_c0	CAGATTCCGACTTGTCCACATTG	619	909	Yes
	TTTGGTGCAGCATAACCATTCCT			
comp70134_c0	AGTTTGGTTACTACCCGCACGAC	680	680	No
	TTGTTGCTTGTACCGTAGCTTAGAG			

coding region also found the proportion of A+T increased after base substitutions (Petrov and Hartl, 1999).

In our study, 23 SNPs were successful with TOF-MS. The observed and expected heterozygosities range from 0 to 0.780 5 and from 0 to 0.505 9, respectively. However, the observed and expected heterozygosities of SSR range from 0.250 to 1.000 and from 0.365 to 0.945, respectively (Cui et al., 2012). The differences may be mainly due to different markers used, as SNPs show less in-

Table 3. The analysis of SNP polymorphism

Location	H_o		H_e		MAF		HWE(P)	
	Susceptible	Resistant	Susceptible	Resistant	Susceptible	Resistant	Susceptible	Resistant
comp65155_c0(135)	0.256	0.310	0.226	0.354	0.128	0.226	0.362	0.402
comp65155_c0(182)	0.267	0.381	0.234	0.312	0.133	0.191	0.325	0.142
comp68973_c0(223)	0.222	0.286	0.200	0.248	0.111	0.143	0.429	0.304
comp68973_c0(279)	0.415	0.548	0.360	0.423	0.232	0.298	0.325	0.053
comp68973_c0(382)	0.600	0.439	0.442	0.438	0.322	0.317	0.015	0.993
comp77722_c0(478)	0.500	0.564	0.471	0.490	0.370	0.410	0.689	0.340
comp55325_c0(384)	0.091	0.325	0.421	0.435	0.296	0.313	0.000	0.104
comp55325_c0(485)	0.195	0.286	0.398	0.391	0.268	0.262	0.001	0.075
comp65582_c0(421)	0.091	0.262	0.496	0.503	0.432	0.464	0.000	0.002
comp65582_c0(163)	0.050	0.075	0.141	0.073	0.075	0.038	0.000	0.842
comp68972_c2(289)	0.133	0.342	0.126	0.287	0.067	0.171	0.664	0.206
comp43413_c0(75)	0.048	0.000	0.248	0.444	0.143	0.324	0.000	0.000
comp77179_c1(138)	0.273	0.364	0.271	0.379	0.159	0.250	0.958	0.780
comp65155_c0(158)	0.477	0.432	0.368	0.342 6	0.239	0.216	0.043	0.077
comp73731_c0(150)	0.523	0.289	0.391	0.250	0.261	0.144	0.022	0.279
comp73731_c0(339)	0.615	0.333	0.432	0.281	0.308	0.167	0.007	0.233
comp65155_c0(208)	0.273	0.400	0.238	0.324	0.136	0.200	0.318	0.105
comp64816_c0(246)	0.605	0.578	0.473	0.505	0.372	0.489	0.064	0.331
comp3(563)	0.781	0.659	0.506	0.505	0.488	0.476	0.000	0.049
comp6(1199)	0.068	0.000	0.067	0.000	0.034	0.000	0.849	0.000
comp204(614)	0.244	0.439	0.217	0.347	0.122	0.212	0.376	0.082
comp206(2463)	0.578	0.310	0.499	0.458	0.444	0.345	0.286 7	0.033
comp331(518)	0.405	0.333	0.403	0.406	0.274	0.278	0.970	0.223
comp334(1217)	0.442	0.364	0.348	0.355	0.221	0.227	0.072	0.873
comp335(1394)	0.429	0.364	0.341	0.329	0.214	0.205	0.087	0.477

Note: H_o represents observed heterozygosity, H_e expected heterozygosity, MAF minor allele frequency, and P-value probability of chi-square test for Hardy-Weinberg equilibrium (HWE).

Table 4. The correlation analysis of DNA SNPs

Location	Mutation type	X^2 (p)	One-Way ANOVA (p)	Significant correlation (Yes/No)
comp65155_c0(135)	G-A	3.776(0.151)	0.148	No
comp65155_c0(158)	T-G	0.183(0.669)	0.348	No
comp65155_c0(182)	A-C	1.300(0.254)	0.390	No
comp65155_c0(208)	G-C	1.613(0.204)	0.339	No
comp68973_c0(223)	C-T	0.464(0.496)	0.098	No
comp68973_c0(279)	C-T	2.510(0.285)	0.447	No
comp68973_c0(382)	T-G	3.756(0.153)	0.020*	Yes
comp77722_c0(478)	C-T	0.751(0.687)	0.766	No
comp55325_c0(384)	T-C	7.341(0.025)*	0.004**	Yes
comp55325_c0(485)	G-T	1.213(0.545)	0.283	No
comp65582_c0(421)	C-T	4.289(0.117)	0.013*	Yes
comp65582_c0(163)	G-A	2.214(0.331)	NO	No
comp68972_c2(289)	C-T	5.207(0.022)*	0.004**	Yes
Comp77848_c0(75)	T-A	6.258(0.044)*	0.002**	Yes
comp77179_c1(138)	C-T	2.214(0.331)	0.109	No
comp73731_c0(150)	A-C	5.050(0.025)*	NO	Yes
comp73731_c0(339)	T-G	6.221(0.013)*	NO	Yes
comp64816_c0(246)	T-C	3.623(0.163)	NO	No

Note: * Significant correlation (P<0.05); ** very significant correlation (P<0.01). p represents probability, and NO means it does not satisfy the analysis condition.

formation content than microsatellites (Aitken et al., 2004). Eleven SNP markers deviate significantly (P<0.05) or very significantly (P<0.01) from HWE. The main reason we surmised is that Huang Xuan No.1 was not a random mating population, but bred by mating system of artificial design. The second factor is that these two groups were separated by low salinity stress. Deviation from HWE indicates that these SNPs are correlated with osmoregulation. Although comp55325_c0(485), comp65155_c0(158),

Table 5. The correlation analysis of cDNA SNPs

Location	Mutation type	X^2 (p)	One-Way ANOVA (p)	Significant correlation (Yes/No)
comp3(563)	T-A	2.696(0.260)	NO	No
comp6(1199)	C-T	2.759(0.097)	NO	No
comp204(614)	G-A	3.634(0.057)	0.057	No
comp206(2463)	T-C	6.616(0.037)*	0.010**	Yes
comp331(518)	C-T	0.714(0.700)	0.556	No
comp334(1217)	G-A	2.326(0.313)	0.127	No
comp335(1394)	T-C	2.404(0.301)	0.130	No

Note: * Significant correlation (P<0.05); ** very significant correlation (P<0.01). p represents probability, and NO means it does not satisfy the analysis condition.

comp3(563), and comp6(1199) deviate significantly (P<0.05) or very significantly (P<0.01) from HWE, association analysis does not show much correlation between the SNPs and osmoregulation. One possible cause of this phenomenon is that these SNPs were selected in the whole breeding group.

Previous studies have shown that within coding exon, the nucleotide diversity is some four-fold lower, with about half resulting in non-synonymous codon changes (Li and Sadler, 1991; Wang et al., 1998; Lai et al., 1998; Nickerson et al., 1998; Harding et al., 1997; Taillon-Miller et al., 1998). About 20% missense mutations would change the significance of a protein, as well as its property. With differences between individuals in the nucleic acid, it is impossible that a synonymous mutation has no significance to the change of biology character at all. Synonymous cSNP (coding-region single nucleotide polymorphism) can modify the structure and function of a protein by the change of its secondary structure (Shen et al., 1999) and translation speed (Komar, 2007) of mRNA. It may generate six known effects. We can use one or several effects to explain its mechanism of action: Due to base mutation, the conventional codons are turned into rare codons, which affect the final protein conformation; mutation generates a hidden RNA splicing site or tail sequence; stability and location of mRNA can also be changed by base mutation; mutation changes the activity of cis-acting elements, which results in changed intensity of gene transcription; if mutation happens on unknown overlap genes, it may influence the activity of another gene; mutation changes the DNA or RNA sequence, which can lead to change in interaction with proteins. SNPs in the introns area can change the activity of the splice sites and the gene function (Tran et al., 2005). The functional mechanism of SNPs in this study is unclear. Further studies are needed to understand their mechanisms.

The selection of the SNPs is the first stage in the study of their functions. Further work is planned to locate these markers in the genome, find their correlation with predicted genes, and strengthen the association analysis with functional genes. This study lays a foundation for the long-term development of high density genetic linkage map, gene mapping, trait transmission tracking, gene function model and the research of gene expression pathway. The data of this study will facilitate the research of crab LS-tolerance mechanism and the development of novel LS-related genes.

Acknowledgements

We were grateful to all the laboratory members for experimental material preparation and technical assistance.

References

Aitken N, Smith S, Schwarz C, et al. 2004. Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. Molecular Ecology, 13(6): 1423–1431, doi: 10.1111/j.1365-294x.2004.02159.x

Arias A, Freire R, Boudry P, et al. 2009. Single nucleotide polymorphism for population studies in the scallops *Aequipecten opercularis* and *Mimachlamys varia*. Conservation Genetics, 10(5): 1491–1495, doi: 10.1007/s10592-008-9766-z

Black IV W C, Baer C F, Antolin M F, et al. 2001. Population genomics: genome-wide sampling of insect populations. Annual Review of Entomology, 46: 441–469, doi: 10.1146/annurev.ento.46.1.441

Bray W, Lawrence A L, Leung-Trujillo J. 1994. The effect of salinity on growth and survival of *Penaeus vannamei*, with observations on the interaction of IHHN virus and salinity. Aquaculture, 122(2–3): 133–146, doi: 10.1016/0044-8486(94)90505-3

Chen R, Davydov E V, Sirota M, et al. 2010. Non-synonymous and synonymous coding SNPs show similar likelihood and effect size of human disease association. PLoS One, 5(10): e13574, doi: 10.1371/journal.pone.0013574

Cui Zhaoxia, Liu Yuan, Wang Hongxia, et al. 2012. Isolation and characterization of microsatellites in *Portunus trituberculatus*. Conservation Genetics Resources, 4(2): 251–255, doi: 10.1007/s12686-011-9518-0

Dai Aiyun, Feng Zhongqi, Song Yuzhi, et al. 1977. Primary investigation on the fishery biology of the *Portunus trituberculatus*. Chinese Journal of Zoology (in Chinese), (2): 30–33, doi: 10.13859/j.cjz.1977.02.015

Dai Aiyun, Yang Siqiong, Song Yuzhi, et al. 1986. Marine Crabs in China Sea (in Chinese). Beijing: China Ocean Press, 194–195

Germer S, Higuchi R. 1999. Single-tube genotyping without oligonucleotide probes. Genome Research, 9(1): 72–78

Harding R M, Fullerton S M, Griffiths R C, et al. 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. American Journal of Human Genetics, 60(4): 772–789

Hirschhorn J N, Sklar P, Lindblad-Toh K, et al. 2000. SBE-TAGS: an array-based method for efficient single-nucleotide polymorphism genotyping. Proceedings of the National Academy of Sciences of the United States of America, 97(22): 12164–12169, doi: 10.1073/pnas.210394597

Holliday R, Grigg G W. 1993. DNA methylation and mutation. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 285(1): 61–67, doi: 10.1016/0027-5107(93)90052-H

Ji Dongsheng. 2005. Techniques of pond-farming of swimming crab, *Portunus trituberculatus*. Special Economic Animal and Plant (in Chinese), 8(3): 12–13, doi: 10.3969/j.issn.1001-4713.2005.03.012

Jin Yulin, Kong Lingfeng, Yu Hong, et al. 2014. Development, inheritance and evaluation of 55 novel single nucleotide polymorphism markers for parentage assignment in the Pacific oyster (*Crassostrea gigas*). Genes & Genomics, 36(2): 129–141, doi: 10.1007/s13258-013-0150-0

Komar A A. 2007. SNPs, silent but not invisible. Science, 315(5811): 466–467, doi: 10.1126/science.1138239

Kumlu M, Eroldogan O T, Saglamtimur B. 2001. The effects of salinity and added substrates on growth and survival of *Metapenaeus monoceros* (Decapoda: Penaeidae) post-larvae. Aquaculture,

196(1–2): 177–188, doi: 10.1016/S0044-8486(00)00580-9

Kumlu M, Jones D A. 1995. Salinity tolerance of hatchery-reared postlarvae of *Penaeus indicus* H. Milne Edwards originating from India. Aquaculture, 130(2–3): 287–296, doi: 10.1016/0044-8486(94)00319-J

Kwok P Y. 2001. Methods for genotyping single nucleotide polymorphisms. Annual Review of Genomics and Human Genetics, 2(2): 235–258, doi: 10.1146/annurev.genom.2.1.235

Lai E, Riley J, Purvis I, et al. 1998. A 4-Mb high-density single nucleotide polymorphism-based map around human APOE. Genomics, 54(1): 31–38, doi: 10.1006/geno.1998.5581

Li Xihong, Cui Zhaoxia, Liu Yuan, et al. 2013. Polymorphisms of anti-lipopolysaccharide factors in the swimming crab *Portunus trituberculatus* and their association with resistance/susceptibility to *Vibrio alginolyticus*. Fish & Shellfish Immunology, 34(6): 1560–1568, doi: 10.1016/j.fsi.2013.03.373

Li Shuzhen, Wan Huirong, Ji Heyi, et al. 2009. SNP discovery based on CATS and genotyping in the finless porpoise (*Neophocaena phocaenoides*). Conservation Genetics, 10(6): 2013–2019, doi: 10.1007/s10592-009-9882-4

Li W H, Sadler L A. 1991. Low nucleotide diversity in man. Genetics, 129(2): 513–523

Livak K J, Marmaro J, Todd J A. 1995. Towards fully automated genome-wide polymorphism screening. Nature Genetics, 9(4): 341–342, doi: 10.1038/ng0495-341

Lv Jianjian, Liu Ping, Wang Yu, et al. 2013. Transcriptome analysis of *Portunus trituberculatus* in response to salinity stress provides insights into the molecular basis of osmoregulation. PLoS One, 8(12): e82155, doi: 10.1371/journal.pone.0082155

Ma Hongyu, Ma Qunqun, Ma Chunyan, et al. 2011. Isolation and characterization of gene-derived single nucleotide polymorphism (SNP) markers in *Scylla paramamosain*. Biochemical Systematics and Ecology, 39(4-6): 419–424, doi: 10.1016/j.bse.2011.05.024

Morin P A, Aitken N C, Rubio-Cisneros N, et al. 2007. Characterization of 18 SNP markers for sperm whale (*Physeter macrocephalus*). Molecular Ecology Notes, 7(4): 626–630, doi: 10.1111/j.1471-8286.2006.01654.x

Nickerson D A, Taylor S L, Weiss K M, et al. 1998. DNA sequence diversity in a 9. 7-kb region of the human lipoprotein lipase gene. Nature Genetics, 19(3): 233–240, doi: 10.1038/907

Péqueux A. 1995. Osmotic regulation in crustaceans. Journal of Crustacean Biology, 15(1): 1–60, doi: 10.1163/193724095X00578

Petrov D A, Hartl D L. 1999. Patterns of nucleotide substitution in *Drosophila* and mammalian genomes. Proceedings of the National Academy of Sciences of the United States of America, 96(4): 1475–1479, doi: 10.1073/pnas.96.4.1475

Piatek A S, Tyagi S, Pol A C, et al. 1998. Molecular beacon sequence analysis for detecting drug resistance in *Mycobacterium tuberculosis*. Nature Biotechnology, 16(4): 359–363, doi: 10.1038/nbt0498-359

Rafalski A. 2002. Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology, 5(2): 94–100, doi: 10.1016/S1369-5266(02)00240-6

Rouse D B, Kartamulia I. 1992. Influence of salinity and temperature on molting and survival of the Australian freshwater crayfish (*Cherax tenuimanus*). Aquaculture, 105(1): 47–52, doi: 10.1016/0044-8486(92)90160-m

Ruscoe I M, Shelley C C, Williams G R. 2004. The combined effects of temperature and salinity on growth and survival of juvenile mud crabs (*Scylla serrata* Forskål). Aquaculture, 238(1–4): 239–247, doi: 10.1016/j.aquaculture.2004.05.030

Sauvage C, Bierne N, Lapègue S, et al. 2007. Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. Gene, 406(1–2): 13–22, doi: 10.1016/j.gene.2007.05.011

Schütz E, Von Ahsen N, Oellerich M. 2000. Genotyping of eight thiopurine methyltransferase mutations: three-color multiplexing, "two-color/shared" anchor, and fluorescence-quenching hybridization probe assays based on thermodynamic nearest-neighbor probe design. Clinical Chemistry, 46(11): 1728–1737

Shen L X, Basilion J P, Stanton V P Jr. 1999. Single-nucleotide polymorphisms can cause different structural folds of mRNA. Proceedings of the National Academy of Sciences of the United States of America, 96(14): 7871–7876, doi: 10.1073/pnas.96.14.7871

Smith C T, Elfstrom C M, Seeb L W, et al. 2005. Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in *Pacific salmon*. Molecular Ecology, 14(13): 4193–4203, doi: 10.1111/j.1365-294X.2005.02731.x

Sommer S S, Groszbach A, Bottema C. 1992. PCR amplification of specific alleles (PASA) is a general method for rapidly detecting known single-base changes. Biotechniques, 12(1): 82–87

Soyel H I, Kumlu M. 2003. The effects of salinity on postlarval growth and survival of *Penaeus semisulcatus* (Decapoda: Penaeidae). Turkish Journal of Zoology, 27(3): 221–225

Stickney H L, Schmutz J, Woods I G, et al. 2002. Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays. Genome Research, 12(12): 1929–1934, doi: 10.1101/gr.777302

Storey J D, Tibshirani R. 2003. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America, 100(16): 9440–9445, doi: 10.1073/pnas.1530509100

Syvänen A C. 1999. From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. Human Mutation, 13(1): 1–10, doi: 10.1002/(SICI)1098-1004(1999)13:1<1:AID-HUMU1>3.0.CO;2-I

Taillon-Miller P, Gu Zhijie, Li Qun, et al. 1998. Overlapping genomic sequences: a treasure trove of single-nucleotide polymorphisms. Genome Research, 8(7): 748–754, doi: 10.1101/gr.8.7.748

Tran H T T, Takeshima Y, Surono A, et al. 2005. A G-to-A transition at the fifth position of intron-32 of the dystrophin gene inactivates a splice-donor site both in vivo and in vitro. Molecular Genetics and Metabolism, 85(3): 213–219, doi: 10.1016/j.ymgme.2005.03.006

Wang Jun, Chuang Karen, Ahluwalia M, et al. 2005. High-throughput SNP genotyping by single-tube PCR with Tm-shift primers. Biotechniques, 39(6): 885–893, doi: 10.2144/000112028

Wang D G, Fan Jianbing, Siao C J, et al. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science, 280(5366): 1077–1082, doi: 10.1126/science.280.5366.1077

Xue Junzeng, Du Nanshan, Lai Wei, et al. 1997. A review of studies on *Portunus trituberculatus* in China. Donghai Marine Science (in Chinese), 15(4): 60–65

Yu Yang, Wei Jiankai, Zhang Xiaojun, et al. 2014. SNP discovery in the transcriptome of White Pacific Shrimp *Litopenaeus vannamei* by next generation sequencing. PLoS One, 9(1): e87218, doi: 10.1371/journal.pone.0087218