Scientific Notes

Polymorphism in *Litopenaeus vannamei* genes and cross-species amplification in other shrimp species

Camilla Alves Santos[1], Flavio Farias[2], Ana Karina Teixeira[2], João Luís Rocha[3], Ana Carolina Guerrelhas[3] and Patrícia Domingues de Freitas[1]

[1]Universidade Federal de São Carlos, Departamento de Genética e Evolução, Via Washington Luiz, Km 235, Caixa Postal 676, CEP 13565-905 São Carlos, SP, Brazil. E-mail: camilla.alves@yahoo.com.br, patdf@ufscar.br [2]Genearch Aquacultura, Pititinga, CEP 59578-000 Rio do Fogo, RN, Brazil. E-mail: flaviofarias@genearch.com.br, karinateixeira@genearch.com.br [3]Aquatec Larvicultura de Camarão Marinho, Barra do Cunhaú, CEP 59190-000 Canguaretama, RN, Brazil. E-mail: johnrocha@genearch.com.br, anaguerrelhas@genearch.com.br

Abstract – The objective of this work was to assess polymorphisms within *Litopenaeus vannamei* candidate genes related to growth performance, and a possible association between weight genotypes and phenotypes in two shrimp closed lines. Three EST-SSR (serine-arginine, 60S ribosomal, and troponin) and one EST-SNP (crustacyanin) were selected and evaluated in *L. vannamei* and seven other marine shrimp species, besides two freshwater species. Genotypes had no effect on total body weight. Although no association was observed in the sampled population, these markers showed amplification for most of the evaluated species and, therefore, may be useful in further genetic studies on shrimp species.

Index terms: candidate genes, crustacyanin, microsatellite, SNP, weight-association, 60S ribosomal.

Polimorfismo em genes de *Litopenaeus vannamei* e amplificação heteróloga em outras espécies de camarão

Resumo – O objetivo deste trabalho foi avaliar polimorfismos em genes candidatos de *Litopenaeus vannamei* relacionados ao crescimento, e uma possível associação entre genótipos e fenótipos de peso em duas linhagens fechadas de camarão. Três SSR-EST (serina-arginina, 60S ribossomal e troponina) e um SNP-EST (crustacianina) foram selecionados e avaliados em *L. vannamei* e sete outras espécies de camarão marinho, além de duas espécies de água doce. Os genótipos não tiveram efeito significativo sobre o peso corporal total. Embora não tenham sido observadas associações na população amostrada, esses marcadores puderam ser amplificados para a maioria das espécies avaliadas e podem ser úteis em estudos genéticos adicionais em espécies de camarão.

Termos para indexação: genes candidatos, crustacianina, microssatélite, SNP, associação à massa, 60S ribossomal.

Molecular markers such as simple sequence repeats (SSR) and single nucleotide polymorphisms (SNP) within genes have showed their wide application in aquaculture in the last decade (Yu & Li, 2008). The association between genotypes and productive traits can be highlighted for several aquaculture species, including the description of a SSR region linked to resistance to White Spot Syndrome Virus (WSSV) disease, in *Penaeus monodon* (Mukherjee & Mandal, 2009), and the description of SNPs associated with survival and disease resistance in *Litopenaeus vannamei* (Marti et al., 2010). Besides, molecular markers in coding regions have high-transferability rates among closely related species. Several studies have described the potential use of SSR and SNP linked to expressed sequence tags (EST) in shrimp species with still poorly known genomes (Freitas et al., 2007; Santos et al., 2012).

The objective of this work was to evaluate polymorphisms in *L. vannamei* shrimp in three microsatellite loci and one SNP, within candidate genes related to the growth performance. All molecular markers were into the protein main chain. A possible association between the genotypes and high growth performance rates in two *L. vannamei* families was also tested for high-growth performance rates. In

addition, we performed cross-species amplification tests for the EST-microsatellite loci in seven other marine shrimp: *Metapenaeus monoceros, Penaeus monodon, Fenneropenaeus indicus, Xiphopenaeus kroyeri, Farfantepenaeus brasiliensis, Litopenaeus schmitti* and *Farfantepenaeus paulensis*; and two freshwater species: *Macrobrachium amazonicum* and *Macrobrachium jelskii.*

Three genes containing SSR – Serine-Arginine (Ser-Arg), 60S ribosomal (60Srib), and Troponin (Trop) –, and one containing one SNP marker – Crustacyanin (Crust) – were obtained from the ShEST Project, available at TSA database, BioProject PRJNA77889, under JR494104, JR494145, JR494085, and JR494393 accession numbers, respectively. Sets of specific primer pairs were designed for each loci using the Primer 3 software (Rozen & Skaletsky, 1999), considering the flanking regions and the respective annealing temperatures for the SSR and SNP, as follows: serine-arginine (F: GATATGGCAGGCTTCGTGAT, and R: TACCTGTCAAGCCAAGCT, at 58°C); 60S ribosomal (F: CAGCTACTGACATCAACTTC, and R: CTCCTTCTTGGCTTCCTCCT, at 58°C); troponin (F: AGCCTGTGTACGAGGAGGAA, and R: AAAATAACGCGACGACATCC, at 64°C), and crustacyanin (F: ATCGAAATGAGCGGCTACAA, and R: TGTGTGTCAAGCTAG, at 55°C).

Polymerase chain reactions (PCRs) were performed for DNA samples of 200 specific pathogen-free shrimps, (SPF) (Genearch Ltda., Rio do Fogo, RN, Brazil). The specimens belonged to families challenged for high-survival and growth rates. Pleopods of 100 animals (50 males and 50 females) from two families (F1 and F3) were sampled and stored in alcohol 100% at -20°C. Whole body weight of each animal was measured with analytical balance, and sex was identified by visualization of the external morphology. The samples of each family were firstly separated into two groups: the heavier group of 50 animals (25 females and 25 males) from F1 and F3 families, with the greatest weight gain rates; and the lighter group of 50 animals (25 females and 25 males), from those same families, with the lowest weight gain rates. The heaviest and lightest animals from each group were then separated into four subgroups.

Biological samples for the nine shrimp species were obtained from the Tissue Bank, at the Laboratory of Molecular Biodiversity and Conservation (University of São Carlos, São Carlos, SP, Brazil), and used in cross-amplification tests. DNA extraction was performed using the phenol protocol proposed by Sambrook et al. (1989). PCR was carried out in 10 µL reaction mixtures containing 50 ηg μL^{-1} DNA, 0.2 mmol L^{-1} dNTPs, 1 x PCR buffer [200 mmol L^{-1} Tris-HCl, pH 8.4, and 500 mmol L^{-1} KCl (Invitrogen Life Technologies)], 1.5 mmol L^{-1} $MgCl_2$, 1 U of Taq DNA polymerase, 8 pmol μL^{-1} of reverse primer, 2 pmol μL^{-1} of the M13-tailed forward primer, and 8 pmol μL^{-1} of the fluorescently-labeled M13 primer for SSR loci (Schuelke, 2000). The M13 primer was not used for the locus with SNP marker. The thermal cycler was programmed for an initial denaturation of 5 min at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at the annealing temperature, 30 s at 72°C, and a final extension of 20 min at 72°C.

The annealing temperatures were established using gradient tests that were performed in a Mastercycler gradient thermal cycler (Life Technologies). Alleles were genotyped in a Illumina automated sequencer MiSeq and scored using the Geneious software (Kearse et al., 2012). Gene product was identified using blastN and blastX against nrNCBI (National Center for Biotechnology Information), SwissProt (Protein Sequence Database), InterPro (Protein Sequence Analysis and classification), and GO (Gene Ontology Consortium) databases. The PCR products obtained from the cross-species amplifications were also sequenced, and blastN (nrNCBI) searches were performed in order to confirm the described protein and its respective species. Allele frequencies were estimated using Genepop version 3.4 (Raymond & Rousset, 1995). The genotype frequencies and analysis of variance were calculated using GLM (general linear model) procedure with SAS (SAS Institute Inc., Cary, NC, USA) version 9.3. The effects from family, sex, and genotypes, for each locus, were also considered. Sex was treated as a fixed effect, and family and genotypes as random effects. The animal whole body weight was considered a dependent variable, and data significance was evaluated at 5% probability.

For the 60Srib protein, the frequencies obtained for the 333 bp and 239 bp alleles were 73 and 27%, respectively. The most frequent genotypes were: 333/333 (66%) in F3 family, in the animals with weight means (WM) equal to 15.3 g; 239/333 (62%) in F1 (WM = 14.6 g); and 239/239 (71%) also in F1 (WM =

13.6 g). Considering the SNP C153T observed in the Crust gene, the frequency for the C allele was 94% and, for the T allele was 6%. The CC (WM = 15.0 g) and CT (WM = 13.9 g) genotypes were equally distributed between F1 and F3 families.

Although the weight means differed among the animals, significant differences were not observed. The SSR into Trop and Ser-Arg genes were both removed from the association analysis, as the first marker exhibited monomorphism, and the second one showed frequency of 98% for the 594/594 genotype.

The effects and significance of family and group were tested in a nested model comparison (test A in the present study), considering sex and genotype effects individually for each locus. Only families and groups significantly affected the total variance for weight ($p<0.001$). The determination coefficient (R^2) was close to 83% for all loci, and family and group showed high-mean square (MS) values, which highlights the great effect of family and group on weight variation for the assumed model (Table 1).

As described earlier, no significant values for genotype means were observed in the previous analysis. Therefore, the two groups were subdivided into four subgroups according to family, in order to verify whether the weight variance could be partially assigned to the genotypes (test B). The four subgroups were considered as follows: subgroups 1 and 2, higher and lower weights, respectively, for animals from the F1 family; and subgroups 3 and 4, higher and lower weights, respectively, for animals from the F3 family. The results for test B were the same as those for test A, with no association observed between genotypes and phenotype weights (Table 1). Some possible interactions between two loci, or among the loci, and the other effects were also tested, but, once more, no association could be observed. Considering that weight is a quantitative trait usually influenced by several genes of small effects each, it is expected that a greater number of loci would explain the variance observed on weight (Jung et al., 2013). Additionally, this variance may be assigned to several factors, such as environmental variance in phenotype expression or bias in the sampling (Visscher et al., 2008). Hence, further studies are recommended to perform association analyses for these and other genes, in different populations with larger sampling.

Regarding the cross-species amplification reactions, the best rates were found for the SSR into the Ser-Arg and 60Srib genes (89% each). This can be justified by the fact that their encoded proteins participate in major metabolic steps, such as splicing control and repairs in translation (Ge et al., 1991; Zhai et al., 2008), tending to be highly conserved among species. Patterns for all SSR loci were observed for *P. monodon, F. indicus, F. brasiliensis* and *F. paulensis* amplification (Table 2). Thus, it is possible that these

Table 1. Analysis of variance of the effects[(1)] of genes, families, and lineages on shrimp weight (*Litopenaeus vannamei*).

Parameter	Family and lineage	Sex	Locus	R^2
		60S ribossomal		
p-value	<0.001	0.42	0.81	0.82
Mean square	557.10	1.26	0.41	
		Crustacyanin		
p-value	<0.001	0.32	0.15	0.83
Mean square	583.76	1.83	3.76	

[(1)]Test A had the same results as test B.

Table 2. Success rates for cross-species amplification using three EST-SSR loci of *Litopenaeus vannamei* in seven marine and in two freshwater shrimp species.

Shrimp species	EST microsatellite			Success (%)
	Serine-arginine	Troponin	60S ribosomal	
Penaeus monodon	X	X	X	100
Fenneropenaeus indicus	X	X	X	100
Farfantepenaeus brasiliensis	X	X	X	100
Farfantepenaeus paulensis	X	X	X	100
Metapenaeus monocerus	X	-	X	66
Litopenaeus schmitti	-	X	X	66
Xiphopenaeus kroyeri	X	-	-	33
Macrobrachium jelskii	X	-	X	66
Macrobrachium amazonicum	-	-	-	0
Success (%)	89	55	89	-

findings may be extrapolated to shrimp species of closely related genera, such as *Farfantepenaeus* and *Litopenaeus* (Ma et al., 2011). Overall, the markers that did not show polymorphisms in the association analyses herein may be potentially useful for species with a small amount or even with no molecular markers described yet, such as most of the shrimp mentioned in this study. Therefore, the markers identified in this study, including those with no allele variants, such as Ser-Arg and Trop, may be applied in further population genetic approaches for *L. vannamei* and other closely related species.

Acknowledgments

To Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp 2012/17322-8 and 2012/13069-6), to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), for grants and financial support.

References

FREITAS, P.D.; CALGARO, M.R.; GALETTI JR, P.M. Genetic diversity within and between broodstocks of the white shrimp *Litopenaeus vannamei* (Boone, 1931) (Decapoda, Penaeidae) and its implication for the gene pool conservation. **Brazilian Journal of Biology**, v.67, p.939-943, 2007. Supplement 4. DOI: 10.1590/S1519-69842007000500019.

GE, H.; ZUO, P.; MANLEY, J.L. Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. **Cell**, v.66, p.373-382, 1991. DOI: 10.1016/0092-8674(91)90626-A.

JUNG, H.; LYONS, R.E.; HURWOOD, D.A.; MATHER, P.B. Genes and growth performance in crustacean species: a review of relevant genomic studies in crustaceans and other taxa. **Reviews in Aquaculture**, v.5, p.77-110, 2013. DOI: 10.1111/raq.12005.

KEARSE, M.; MOIR, R.; WILSON, A.; STONES-HAVAS, S.; CHEUNG, M.; STURROCK, S; BUXTON, S.; COOPER, A.; MARKOWITZ, S.; DURAN, C.; THIERER, T.; ASHTON, B.; MEINTJES, P.; DRUMMOND, A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, v.28, p.1647-1649, 2012. DOI: 10.1093/bioinformatics/bts199.

MA, K.Y.; CHAN, T.-Y.; CHU, K.H. Refuting the six-genus classification of *Penaeus* (Dendrobranchiata, Penaeidae): a combined analysis of mitochondrial and nuclear genes. **Zoologica Scripta**, v.40, p.498-508, 2011. DOI: 10.1111/j.1463-6409.2011.00483.x.

MARTI, S.M.; ONTERU, S.K.; DU, Z.Q.; ROTHSCHILD, M.F. Short communication. SNP analyses of the 5HT1R and STAT genes in Pacific white shrimp, *Litopenaeus vannamei*. **Spanish Journal of Agricultural Research**, v.8, p.53-55, 2010. DOI: 10.5424/sjar/2010081-1162.

MUKHERJEE, K.; MANDAL, N. A microsatellite DNA marker developed for identifying disease-resistant population of giant black tiger shrimp, *Penaeus monodon*. **Journal of the World Aquaculture Society**, v.40, p.274-280, 2009. DOI: 10.1111/j.1749-7345.2009.00250.x.

RAYMOND, M.; ROUSSET, F. Genepop (version 1.2): population genetics software for exact tests and ecumenicism. **Journal of Heredity**, v.86, p.248-249, 1995. DOI: 10.1093/oxfordjournals.jhered.a111573.

ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. **Bioinformatics Methods and Protocols**, v.132, p.365-386, 1999. DOI: 10.1385/1-59259-192-2:365.

SAMBROOK, J.; FRITSHI, E.F.; MANIATIS, T. **Molecular cloning**: a laboratory manual. 2nd ed. New York: Cold Spring Laboratory, 1989.

SANTOS, C.A.; ROSSINI, B.C.; MARQUES, C.G.; GALETTI JR, P.M.; FREITAS, P.D. Characterization and genomic annotation of polymorphic EST-SSR loci in *Litopenaeus vannamei* shrimp. **Aquaculture Research**, v.43, p.1567-1570, 2012. DOI: 10.1111/j.1365-2109.2011.02956.x.

SCHUELKE, M. An economic method for the fluorescent labeling of PCR fragments: a poor man's approach to genotyping for research and high-throughput diagnostics. **Nature Biotechnology**, v.18, p.1-2, 2000.

VISSCHER, P.M.; HILL, W.G.; WRAY, N.R. Heritability in the genomics era - concepts and misconceptions. **Nature Reviews Genetics**, v.9, p.255-266, 2008. DOI: 10.1038/nrg2322.

YU H.; LI, Q. Exploiting EST databases for the development and characterization of EST-SSRs in the Pacific oyster (*Crassostrea gigas*). **Journal of Herdability**, v.99, p.208-214, 2008. DOI: 10.1093/jhered/esm124.

ZHAI, B.; VILLÉN, J.; BEAUSOLEIL, S.A.; MINTSERIS, J.; GYGI, S.P. Phosphoproteome analysis of *Drosophila melanogaster* embryos. **Journal of Proteome Research**, v.7, p.1675-1682, 2008. DOI: 10.1021/pr700696a.

Received on December 29, 2016 and accepted on June 27, 2017