

Available online at: <http://ejournal-balitbang.kkp.go.id/index.php/iaj>

## IDENTIFICATION OF SNP MARKER FOR CRUSTACEAN HYPERGLYCEMIC HORMONE GENE: A SOMATIC GROWTH-RELATED IN GIANT FRESHWATER PRAWN (*Macrobrachium rosenbergii*)

Asep Sopian<sup>1) #</sup>, Alimuddin<sup>2)</sup>, Imron<sup>3)</sup>, Harry Krettiawan<sup>3)</sup>, Fajar Anggraeni<sup>3)</sup>, and Desy Nurul Astuti<sup>3)</sup>

<sup>1)</sup> Post graduate student, Department of Aquaculture, Faculty of Fisheries and Marine Science, Bogor Agricultural University

<sup>2)</sup> Department of Aquaculture, Faculty of Fisheries and Marine Science, Bogor Agricultural University

<sup>3)</sup> Research Institute for Fish Breeding

(Received 30 May 2017; Final revised 2 August 2017; Accepted 2 August 2017)

### ABSTRACT

High size variation of giant freshwater prawn was found in harvest and resulting in low productivity. Marker assisted selection may be useful to generate broodstock that produces progeny with high growth and homogeneity. This study was conducted to obtain growth related molecular marker in giant freshwater prawn. Genomic DNA was extracted from swimming leg (pleiopods) of 10 giant freshwater prawns fifth Generation for existence of SNP identification, consisted of 5 fast growth (FG) and 5 slow growth (SG). While for SNP confirmation and resolving power of specific primer studies. The pleiopods sample was taken from six generation of 201 giant freshwater prawns, consisted of 129 fast-growth (FG) with  $16.06 \pm 2.48$  g body weight and 72 slow-growth (SG) with  $6.05 \pm 0.90$  g body weight. Oligonucleotide primers were designed according to Gene Bank database of crustacean hyperglycemic hormone (CHH) gene sequence. The amplified DNA fragment was then sequenced. The results of sequencing showed there was one base different in nucleotides of FG and SG prawns. Six set of primers were designed based on those CHH gene sequence. PCR analysis resulted one set of primers which showed a specific amplification product of 280 bp for growth. The result of sequence analysis using the basic local alignment search tools showed that the nucleotide sequence of those PCR products had similarity of 99%-100% with CHH gen of *M. rosenbergii*. Thus, a candidate of growth related molecular marker have been identified for giant freshwater prawn.

**KEYWORDS:** giant freshwater prawn; SNP molecular marker; fast growth; slow growth

### INTRODUCTION

The giant freshwater prawn (*Macrobrachium rosenbergii*) is the largest known palaemonid among other freshwater prawns. It is native to the Vietnam, Malaysia, Myanmar, Indonesia, India, Bangladesh, Thailand, Pakistan, Philippines, and Australia. It has been introduced for culture in other areas beyond its natural distribution such as in United State of America, China, and other American countries. In 2013, total farmed *M. rosenbergii* volume reached more than 203.028 tons, and 98% of them come from Asia (FAO, 2016).

Prospect of market opportunity is still open for giant freshwater prawn (GFP), locally or internationally, especially to EU countries, USA, and Japan. However, the productivity of GFP aquaculture is still left behind with the *Litopenaeus vannamei* and other penaeid prawns. One of the causes of low productivity of giant freshwater prawn is the very high variation in size at the end of the growth-out phase.

High variability of individual size during harvesting time become one reason causing low productivity in cultivated GFP. Growth, survival rate, and disease resistant are known to depend not only on genetics factor but also on environmental factors (Hartnoll, 2001). One genetic factor was improper broodstock selection, while some of environmental factors are variation on water temperature, survival rate, food and nutrition, and stocking density.

# Correspondence: Research Institute for Fish Breeding.  
Jl. Raya Sukamandi No. 2, Subang 41256, Indonesian.  
Phone: + (0260) 520500  
E-mail: [publikasi.bppi@gmail.com](mailto:publikasi.bppi@gmail.com)

In breeding, one of the most important aspect to explore is the broodstock selection because it is related to the high seed quality. Broodstock genetic improvement can be pursued to gain high quality seed. Conventional breeding program through selection is still being used and proven to be successful to increase production. However, there are still limitations on conventional selection, especially in time period to bring up the desired genes.

Currently, lots of information about selection activities are based on the molecular technology. Theoretically, molecular marker technologies can provide diverse information of genes to explore. Various techniques and molecular markers have been developed for screening allele polymorphism in the gene. One of the DNA markers that can be applied for breeding program is the single nucleotide polymorphism (SNP). SNP is a single nucleotide difference between DNA sequences from compared individuals that can cover more than 90% of the genetic variation, so it can be a marked the difference between the varieties of a species. SNP markers have been used recently to screen candidate gene for growth in a number of species, including *Litopenaeus vannamei* and *Penaeus monodon* (Glenn *et al.*, 2005), and *Macrobrachium rosenbergii* (Thanh *et al.*, 2010; Jung *et al.*, 2014). A previous study reported that four intronic SNPs showed positive associations with three growth trait (body weight, carapace length, and standard length) (Thanh *et al.*, 2010).

Crustacean hyperglycemic hormone (CHH) gene has high potential to impact body weight variation in crustacean and should, therefore, be considered as a primary gene of interest in growth studies (Jung *et al.*, 2014). Identification of CHH genes has never been studied on Indonesian GFP. Based on Thanh *et al.* (2010), a study for marker associated growth identification by PCR amplification using specific primer was designed for GI Macro-2 population. The specific primer was based on the CHH growth gene sequence from GenBank, with GI Macro-2 DNA template which have the characteristics of fast and slow growth.

This study was conducted to obtain a specific marker that can differentiate the growth speed of GFP and evaluate whether SNPs exist in CHH gene of GFP populations showing slow and fast growth performances. Further, it was also aimed to find specific primer design that has ability to differentiate between those two contrasting phenotypes.

## MATERIAL AND METHOD

The experiments conducted for the research consisted of pond and laboratory activities, from April

2015 to October 2016, at Research Institute for Fish Breeding (RIFB), Sukamandi, West Java, Indonesia.

### Experimental Animals

The fifth generation (G5) and six generation (G6) of Genetically improved *Macrobrachium* II (GI Macro-2), a selectively bred population produced by Research Institute for Fish Breeding, Sukamandi (Released Document Unpublished) were used. They were harvested following three months of rearing in a grow out pond. At the time of harvest, the population was sexually separated, and only females were used to avoid the social rank effect on male GFP (Thanh *et al.*, 2009). G5 and G6 following sexual based separation, the population was separated into small, medium, and large groups based on its size. The GFP grouping was carried out based on standard length trait using criteria described by Imron *et al.* (2015) as presented on Table 1. For the purposes of this study, only small and large groups were used. Large group was designated as fast growth (FG) GFP, while small group was designated as slow growth (SG) GFP.

Pleopod samples of GI Macro-2 from the fifth generation were collected for existence SNP identification. The numbers of samples available for FG and SG were 5, respectively. For SNP confirmation and resolving power of specific primer studies, the pleopod samples of GI Macro-2 from six generation were collected at harvest. The numbers of samples available for FG were 129 with  $16.06 \pm 2.48$  g and 72 SG with  $6.05 \pm 0.90$  g body weight, seven was chosen randomly for PCR amplification analysis.

### CHH Gene Amplification and DNA Sequencing

CHH gene detection from FG and SG of GI Macro-2 population referred to Thanh *et al.* (2010) primer, called CH3. The forward primer for CH3 was 5'-GGTCATTGCGTGGAAGATTT-3' and the reverse primer was 5'-GGCAGA TGAGAGGGACTGAG-3' with 608 bp PCR products. DNA PCR products from G5 consisting of 5 FG and 5 SG respectively, were sent to the laboratory of 1<sup>st</sup> Base Sequencing INT Singapore using the ABI PRISM 3730 XL (Genetic Analyzer Developed by Applied Biosystems, USA) using the Sanger method, for sequencing. The Basic Local Alignment Search Tool (BLAST) [www.blast.ncbi.nlm.nih.gov/] (Altschul *et al.*, 1990) was used to determine the similarity of the gene sequenced result with the GenBank data.

### Designing Specific Primers Having Differential Amplification Profile on FG and SG

The CH3 primer used in the previous step successfully amplified PCR products of both FG and SG

Table 1. The GFP grouping into small, medium, and large based on Imron *et al.* (2015). The Y and X indicate observation values belong to a particular group and sample average, respectively. SD indicates standard deviation value

Group	Criteria
Small	$Y < X - 1SD$
Medium	$(X + 1SD > Y > (X - 1SD))$
Large	$Y > X + 1SD$

at exactly the same size. Therefore, it was not possible to differentiate directly both groups using PCR product of genomic DNA amplified using this primer. For practical purposes, it would be very convenience to have primer sequences that result in different amplification profiles between the FG and SG of GFP. The different amplification profiles can be in the forms of size differences of PCR products or in the present or absent profile of PCR products. SNPs on the gene were analyzed and identified using Bioedit (Hall, 1999) and Clustal W (Thompson *et al.*, 1994). Specific primers design was conducted using Primer 3 Plus program, where the first SNP identified became forward primer and the last SNP identified became the reverse primer. To ensure the amplified DNA is the target gene, the PCR products, consist of 6 FG amplicons and 4 SG amplicons, were sent and sequenced at the laboratory of 1<sup>st</sup> Base Sequencing INT Singapore and analyzed using BLAST. Polymorphics DNA bands resulted from specific primer was then statistically analyzed using t-student (SPSS 2.2).

#### Specific Primer Confirmation on Large Population

The specific primer was then used to amplify larger population (n=201), consisting of 129 FG and 72 SG. Genomic DNAs were collected from pleiopods. Pleiopods were taken aseptically and put in a 1.5 mL sterile tube. Genomic DNA from FG and SG were extracted using Gene Jet DNA Purification kit (thermo Scientific). Each genomic DNA was dissolved in 50  $\mu$ L of elution buffer, and then stored at -20°C to be used in the PCR process. The amplification processes were using the PCR thermocycler (ESCO). The total of PCR mix was 25  $\mu$ s, consist of: 12.5  $\mu$ L FastStart PCR mastermix kit (Rosche, Germany), 2  $\mu$ L of primer (10  $\text{nmol}/\mu\text{L}$ ), 2  $\mu$ L of genomic DNA, and 8.5  $\mu$ L water free RNase. Gel electrophoresis with Tris-Acetate-EDTA solution (TAE 1x) was run using 2% agarose (Vivantis) for 35 minutes at 100 volts. Peggreen (Peclab) dye for about 2  $\mu$ L was then added in agarose. The DNA visualization in agarose utilized UV transilluminator.  $\beta$ -actin primer used as the internal control of DNA loading.

#### Data Analysis

##### SNPs Identification

The presence or absence of SNPs within the CHH gene using CH3 primer was also meant to identify the presence of SNPs between the two groups. SNPs identification was conducted by aligning the DNA sequences of the two groups and identifying the SNPs by visual judgment. Specific information of interest regarding the SNPs included the number and position of SNPs, type of mutations whether they were transition or transversion, and deduced amino acid substitution, whether they were synonymous or non-synonymous.

##### Resolving Power of Specific Primer

The ability of specific primer to distinguish between FG and SG groups was determined using differences in amplification product, either as present or absent state, or as size differences when both groups were successfully amplified. The distinguishing ability of the primer was counted as percentage of correctly classifying phenotypes according to the expected under genotype (profile of PCR products).

#### RESULTS AND DISCUSSION

##### Crustacean Hyperglycemic Hormone (CHH) Gene Amplification and SNPs Identification

CHH gene came from 5 FG and 5 SG samples of G5 GI Macro-2 GFP strain in this study were successfully amplified using CH3 primer, with 608 bp in length (Figure 1). The result was similar to Thanh *et al.* (2010), and indicated the existence of CHH gene in GI Macro-2 population.

The BLAST analysis result indicated that the sequenced FG and SG populations from GI Macro-2 strain had 99% to 100% similarity to that known CHH gene in the GenBank (Table 2). This high similarity showed that the indicated CHH gene from GI Macro-2 strain is the same as the CHH gene in the GenBank.

SNPs identification can be detected using sequences of CHH gene fragment. Based on the align-

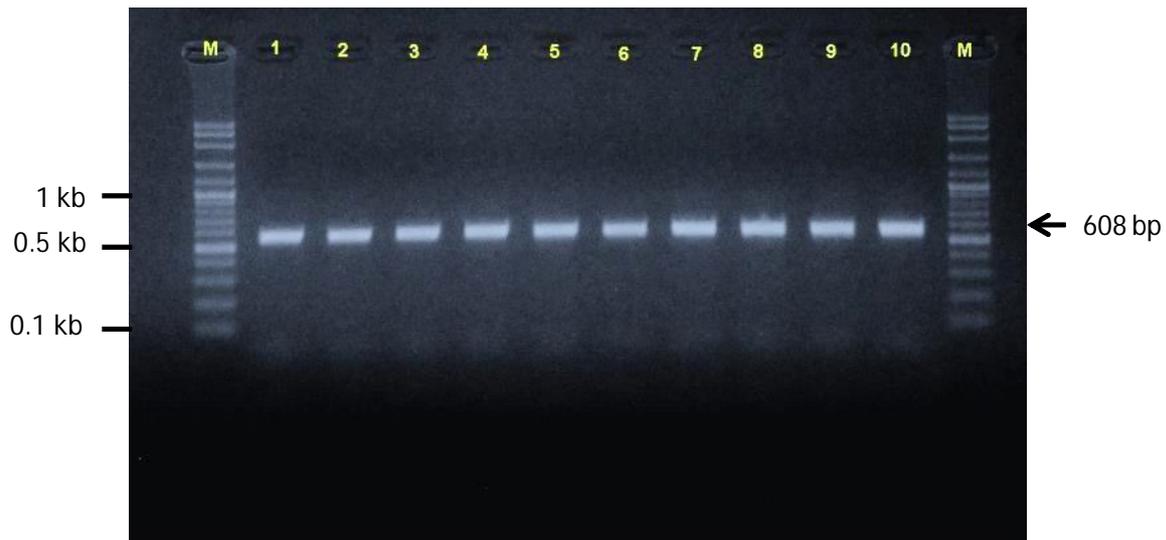


Figure 1. The PCR amplification product of CHH gene fragments of Gimacro-2 strain. M= DNA fragment size marker (Vivantis); 1-5= fast growth; 6-10= slow growth. The arrow indicated the DNA amplified fragment.

Table 2. BLAST similarity of crustacean hyperglycemic hormone (CHH) gene from GI Macro-2 strain and known as GenBank sequences

Accession no.	Description	Homology (%)
AF372657.1	<i>M. rosenbergii</i> CHH gene, complete cds	99
AF219382.1	<i>M. rosenbergii</i> CHH mRNA, complete cds	99
AJ130968.1	<i>M. rosenbergii</i> CHH-H mRNA, partial cds	100

ment result, some allele variations (SNP) were found between FG and SG GFP, those were: changes from A to G (A>G), C to T (C>T), and deletion (Table 3). Changes from G to A and T to C occur more frequent in FG than SG.

De Santis & Jerry (2007) declare that identified SNPs in the DNA sequence can potentially change the structure and activity of the protein coding, or cause changes in the regulation of mRNA transcription level. Thus, detected SNPs in this study potentially can distinct the FG from SG in GI Macro-2 GFP strain.

According to analysis result the nucleotide sequence of CHH gene fragment came from FG and SG GFP showed five SNP which were at 147, 158, 257, 324, and 402 (Table 4). Allel variation that happened in CHH gene fragment were transition substitutions, while transversion substitution was not found. The deletion was only found for one base at the start of the fragment, for T base at 48 base positions.

Shu *et al.* (2010) confirmed that most of the identified SNPs directly correlated with gene expression

were found in the coding region or exon. Genissel *et al.* (2004) reported that 1% of SNPs in exon region are able to perform the changes in the amino acid product. While Zhang & Zhang (2005) reported that SNPs in the intron region can affect the gene expression level, RNA stability, and perform the quantitative variances.

Identified SNPs on CHH gene fragment in this study were all found in non-coding region or intron. This was occasionally happened, because natural selection generally conserves coding regions (Liu, 2007). Those SNPs were able to distinct the FG and SG from GI Macro-2 population. The same result was reported by Thanh *et al.* (2010) that 4 SNPs found in intron on CHH gene fragment had positive correlation with growth performance in GFP.

Based on Table 3, SNP at 147 bp, 158 bp, and 402 bp were more significant in changes then that of the other three. Thus, the specific primer design was followed the three SNPs. The forward primer included SNP at 147 bp (F: 5'-CGTTGCATTTTATGTTAG-3'), and

Table 3. Allelic variation (SNP) in CHH gene alignment of GI Macro-2 GFP strain

SNPs (bp)	Fast growth (FG)					Consensus	Slow growth (SG)					Consensus
	1	2	3	4	5		1	2	3	4	5	
48	-	-	-	-	<b>T</b>	-	-	-	-	-	-	-
147	<b>G</b>	A	A	A	A	A	G	G	<b>A</b>	G	G	G
158	A	<b>G</b>	A	A	<b>G</b>	A	A	A	A	A	A	A
257	T	T	<b>C</b>	T	T	T	T	T	<b>C</b>	T	T	T
324	A	<b>G</b>	A	A	A	A	A	A	A	A	A	A
402	C	C	<b>T</b>	<b>T</b>	C	C	C	C	C	C	C	C

Remarks: - = deletion. Bold letters indicates different bases and SNPs are shown in bold letters

the reverse primer included the 402 bp SNP (R: 5'-CGTATTCGTCGAGTTG-3'). This specific primer was applied to the previous DNA samples and able to alter FG from SG in GI Macro-2 strain. Most of the FG population showed negative bands while the SG population showed positive bands. The amplified fragment was at 280 bp, while the β-actin marker as an internal control produced bands on both populations at 480 bp (Figure 2).

### Specific Primer Confirmation on Large Population

Result of BLAST analysis from amplified product of specific primer presented on Table 4. The similarity of DNA sequence from GI Macro-2 strain and known GFP GenBank accessions were high, mean that the specific primer still able to amplify the CHH gene. The specific primer then applied on large population of GI Macro-2 GFP strain.

Table 4. BLAST result from amplified specific primers for CHH gene on the FG and SG populations of GI Macro-2 GFP strain

Population	GenBank GFP Accession	Query cover (%)	Homology (%)
FG80	<i>Macrobrachium rosenbergii</i> CHH gene (AF372657.1)	100	99
SG7	<i>Macrobrachium rosenbergii</i> CHH gene (AF372657.1)	100	99

Remarks: FG = fast growth, SG = slow growth

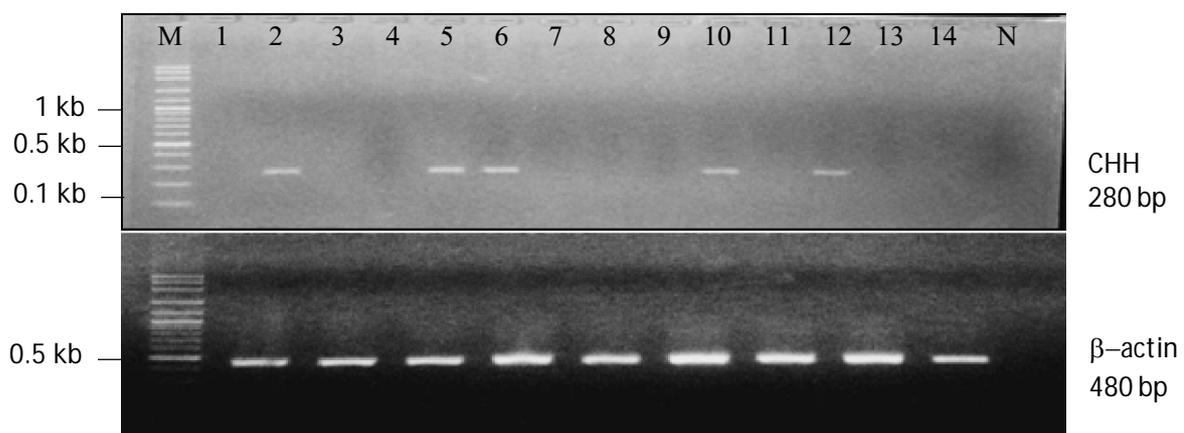


Figure 2. PCR amplification products of specific primer in CHH gene in GI Macro-2 strain. β-actin marker was used as control loading. M = Marker of DNA fragment size (Vivantis); 1-7 = fast growth GFP; 8-14 = slow growth GFP; and N = none DNA template (negative control).

From 129 FG and 72 SG samples, as many as 58 FG (45%) and 22 SG (31%) showed positive bands. Either positive or negative bands were appeared on both populations. The inconsistencies of bands appearance on both populations were then analyzed by t-student test (Figure 3). The statistical test proved that there was significant difference between FG and SG populations. This means that the specific primer was able to distinguish the FG population from SG population on GI Macro-2 GFP strain, statistically.

To clarify the inconsistency of bands appearance on both populations, the CHH gene alignment was then analyzed (Table 5). Three strong SNPs were used to analyze the bases consistency alignment on 5 FG and 5 SG samples. Most of the FG samples had nega-

tive bands, and most the SG samples had positive bands during the research.

Based on the alignment, the SG samples showed consistency on (G-A-C) bases pattern for positive bands. The first SNP (G/A) will indicate positive band if the base was G and negative band if the base was A. This confirmed why the SG3 and most of the FG samples had the negative band. When the base was C, it confirmed for the SG population; while if the base was T, it confirmed for the FG population. Based on the analysis, the SG population will have (G-A-C) or (A-A-C) bases pattern; while for the FG population will have (A-G-C) or (A-A-T) bases pattern.

There is anomaly for FG1, where the pattern was (G-A-C), that should be the pattern for SG popula-

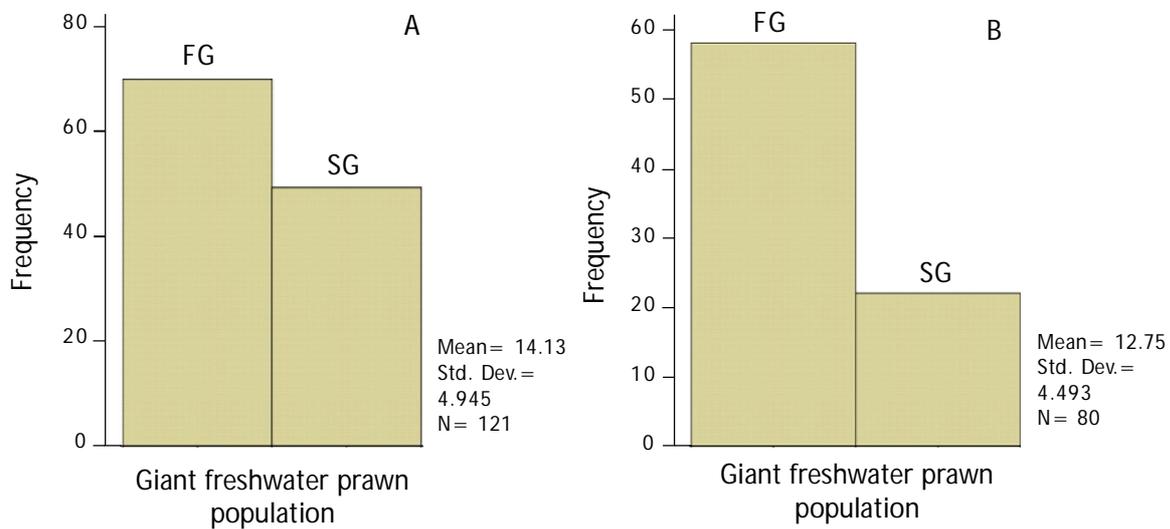


Figure 3. Evaluation on specific primer of generation 6 GI Macro-2. (A) for negative bands and (B) for positive bands.

Table 5. The pattern of difference SNP of gene CHH fragment at FG and SG of GFP

Population samples	Weight	SNP bases			Amplified bands
		1	2	5	
FG1	107	G	A	C	Positive
FG2	99	A	G	C	Negative
FG3	89	A	A	T	Negative
FG4	101	A	A	T	Negative
FG5	96	A	G	C	Negative
SG1	74	G	A	C	Positive
SG2	71	G	A	C	Positive
SG3	78	A	A	C	Negative
SG4	76	G	A	C	Positive
SG5	73	G	A	C	Positive

tion, having 107 g in weight. This anomaly cannot be answered from this study alone. Some hypothesis for this anomaly are offered: 1) CHH gene is not the only gene that control growth on GFP; and 2) more SNP specific markers needed to apply in CHH gene, to determine the differences growth of GFP.

## CONCLUSIONS

SNP markers associated genes identified growth to GFP the size of the PCR product of about 280 bp. The use of specific primers in this study was statistically proven to differentiate between FG and SG of GFP.

## ACKNOWLEDGEMENT

The authors appreciate the funding support from APBN with DIPA No. 032.11.2.660052/2016 in Research Institute for Fish Breeding, Sukamandi. We thank to Mr. Hasan, Mr. Khairul Syahputra, Mrs. Huria Marnis, Mr. Ali Akbar, Mr. Dede, Mr. Uus, Mr. Galih, Miss Dessy for the support of this research.

## REFERENCE

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.*, 215, 403-410.
- De-Santis, C. & Jerry, D.R. (2007). Candidate growth genes in finfish – where should we be looking? *Aquaculture*, 272, 22-38.
- Food and Agriculture Organization [FAO]. (2016). FAO yearbook Fishery and Aquaculture Statistics 2014. Rome, Italy.
- Genissel, A., Pastinen, T., Dowell, A., Mackay, T.F.C., & Long, A.D. (2004). No evidence for an association between common nonsynonymous polymorphism in delta and bristle number variation in natural and laboratory populations of *Drosophila melanogaster*. *Genetics Society of America*, 166, 291-306.
- Glenn, K.L., Grapes, L., Suwanasopee, T., Harris, D.L., Li, Y., Wilson, K., & Rothschild, M.F. (2005). SNP analysis of *AMY2* and *CTSL* genes in *Litopenaeus vannamei* and *Penaeus monodon* shrimp. *Animal Genetics*, 36, 235-236.
- Hall, T.A. (1999). Bio Edit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.*, 41, 95-98.
- Hartnoll, R.G. (2001). Growth in crustacea-twenty years on. *Hydrobiologia*, 449, 111-112.
- Imron, Iswanto, B., Ridzwan, N.S., Suprpto, R., & Marnis, H. (2015). Association of microsatellite genetic diversity with growth related traits in the base population of African catfish, (*Clarias gariepinus*), breeding program. *Indonesian Aquaculture Journal*, 10(1), 1-11.
- Jung, H., Lyons, R.E., Li, Y., Thanh, N.M., Dinh, H., Hurwood, D.A., Salin, K.R., & Mather, P.B. (2014). A candidate gene association study for growth performance in an improved giant freshwater prawn (*Macrobrachium rosenbergii*) culture line. *Mar. Biotechnol.*, 16, 161-180.
- Liu, Z. (2007). Single nucleotide polymorphism (SNP). In Liu, Z. (Ed.), *Aquaculture genome technologies* (p. 59-72). Iowa, USA: Blackwell.
- Shu, Y.J., Li, Y., Wu, N.L.H., Bai, X., Cai, H., Ji, W., & Zhu, Y.M. (2010). Mining and identification of SNPs from EST sequences in soybean and converting SNP markers into CAPS. *Acta Agron. Sin.*, 36(4), 574-579.
- Thanh, N.M., Ponzoni, R.W., Nguyen, N.H., Vu, N.T., Barnes, A., & Mathe, P.B. (2009). Evaluation of growth performance in a diallel cross of three strains of giant freshwater prawn (*Macrobrachium rosenbergii*) in Vietnam. *Aquaculture*, 287, 75-83.
- Thanh, N.M., Barnes, A.C., Mather, P.B., Li, Y., & Lyons, R.E. (2010). Single nucleotide polymorphisms in the actin and crustacean hyperglycemic hormone genes and their correlation with individual growth performance in giant freshwater prawn (*Macrobrachium rosenbergii*). *Aquaculture*, 301, 7-5.
- Thompson, J.D., Higgins, D.G., & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22(22), 4673-4680.
- Zhang, D. & Zhang, Z. (2005). Single nucleotide polymorphism (SNPs) discovery and linkage disequilibrium (LD) in forest trees. *Forestry Studies in China*, 7(3), 1-14.