WALAILAK JOURNAL
http://wjst.wu.ac.th Applied Sciences

RNA Sequence Analysis of Growth-Related Genes in *Penaeus monodon*

Nifareesa CHEALOH[1,2], Sataporn DIREKBUSARAKOM[1,2], Piyapong CHOTIPUNTU[1,2], Pitchanee JARIYAPONG[2,3], Hidehiro KONDO[4], Ikuo HIRONO[4] and Suwit WUTHISUTHIMETHAVEE[1,2,*]

[1]*School of Agricultural Technology, Walailak University, Nakhon Si Thammarat 80161, Thailand*
[2]*Research Center of Excellence on Shrimp, Walailak University, Nakhon Si Thammarat 80161, Thailand*
[3]*School of Medicine, Walailak University, Nakhon Si Thammarat 80161, Thailand*
[4]*Laboratory of Genome Science, Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, Minato, Tokyo 1088477, Japan*

(*Corresponding author's e-mail: suwitwuth@gmail.com)

Received: 25 April 2018, Revised: 14 June 2018, Accepted: 20 July 2018

Abstract

Penaeus monodon is one of the most economically important shrimp species in Thailand. However, little information is available about the functional genomics related to its growth performance. In this study, Illumina paired-end sequencing was used to analyze transcriptomes related to growth performance in *P. monodon* muscle. A total of 38.4 million reads were generated. The pooled reads, from 10 libraries, were *de novo* assembled into 113,991 genes, with an average length of 337 bp. Gene expression was analyzed with the edgeR program, which revealed 705 differentially expressed contigs ($p < 0.05$) in fast-growth shrimp compared to slow-growth shrimp. The results show the 234 up-regulated contigs in fast-growth shrimp are mostly underlined genes involving the metabolic pathway. Quantitative real-time polymerase chain reaction (qRT-PCR) revealed seven genes involved in the cell cycle that were expressed more in fast-growth shrimp ($p < 0.05$) than in slow-growth shrimp, and moderately to strongly correlated with shrimp body weight. These genes may be good candidates for growth performance improvement in *P. monodon*.

Keywords: *Penaeus monodon*, RNA sequencing, qRT-PCR, growth-related genes, growth improvement

Introduction

The shrimp farming industry began as inland aquaculture over 30 years ago in response to increasing consumption of and demand for shrimp worldwide. In 2007, the estimated global production of shrimp farms reached 3.36 million metric tons, with a value of over US$14 billion [1]. In Southeast Asia, Thailand has been a major producer and exporter of farmed shrimp since 2015 [2]. Black tiger shrimp (*P. monodon*) is one of the most economically important shrimp species in Southeast Asia, and is distributed throughout the Indo-Pacific region [3]. Current production of black tiger shrimp in Thailand has dramatically decreased due to a number of problems. One of the most significant of these is the slow growth performance of black tiger shrimp compared to other species, particularly Pacific white shrimp (*Litopenaeus vannamei*). However, black tiger shrimp has been long domesticated in Thailand, and shrimp growers are acquainted to their behavior; this makes renovation of black tiger shrimp culture in Thailand a challenge. Innovation and knowledge in improving growth performance are essential to re-establish the aquaculture industries of this species, not only in Thailand, but throughout the region.

Growth performance of shrimp is naturally dependent on internal, external [3,4], and genetic factors that contribute to phenotypic variation. The candidate genes potentially associated with individual growth performance have been reported in finfish [5] and *Macrobrachium rosenbergii* [6]. For *P. monodon,*

expression analysis revealed growth-related candidate genes in the eyestalk and muscle [7,8]. While there are no available gene expression data to compare fast and slow growth in this species, this could be a useful tool to create direct strategies to improve growth performance.

Through transcriptome analysis using next-generation sequencing technology, which considers extensive gene expression and does not require prior knowledge of gene information [9], the growth-related genes for fast and slow growth were identified in order to enrich and fill gaps in knowledge about the molecular information for *P. monodon*. The RNA sequence data were *de novo* assembled, and the gene abundances were estimated. The up-regulated contigs in fast-growth shrimp are mostly underlined genes involving the metabolic pathway. The selected gene participates in the cell cycle was found to be moderately to strongly correlated with shrimp body weight and may be a good candidate gene for growth performance improvement in *P. monodon*.

Materials and methods

Sample collection and RNA extraction

Specific-pathogen-free *P. monodon* (four months of age) were reared and collected from the Shrimp Quarantine Center, Walailak University. First, the body weight and body length of 50 individual shrimp at 4 months old were recorded, and the mean and standard deviation were determined and used as criteria to divide the shrimp into two groups (fast-growth group: shrimp > mean + sd, and slow-growth group: shrimp < mean + sd). Muscle tissue was collected from five shrimp specimens in each group and individually separated for next-generation sequencing (NGS) analysis. Total RNA was extracted from shrimp muscle tissue using TriPure Isolation Reagent® (Roche, Basel, Switzerland). The tissue samples were ground in liquid nitrogen and mixed with TriPure Isolation Reagent®. Phenol-chloroform was then added to the suspension to separate DNA and protein from the RNA sample. Total RNA was precipitated using isopropanol, washed with ethanol, and digested with DNase treatment to eliminate contaminated genomic DNA. After centrifugation, the RNA was dried and used for NGS and quantitative RT-PCR.

Gene expression profiling analysis

mRNA was isolated from individual DNase-treated total RNA using a Truseq RNA sample preparation kit (Illumina, San Diego, U.S.A.) according to the manufacturer's low-throughput protocol instruction. The cDNA library, with fragment sizes of about ~240 bp and different barcodes from 10 individual libraries, was pooled and sequenced on the Illumina MiSeq platform (Illumina, San Diego, California, U.S.A.). All sequence reads were filtered with trimmomatic software to remove adapter sequencing and very short sequences [10]. Sequence reads were combined from all 10 libraries and *de novo* assembled into contigs using the Trinity software [11]. The sequence reads were assembled using default parameters with a k-mer of 25 to generate the transcriptome assembled contigs [12]. Reads from individual shrimp samples were mapped onto the assembled contigs using the Bowtie program. Gene abundances were determined by the gene count number using RSEM software [13]. The gene count data of each library from fast-growth and slow-growth shrimp were used to detect the differential expressed genes. The Exact test was used to identify differentially expressed genes across samples in the two groups and normalize the number of counts per million reads [14-17], using the support tool from Blast2GO bioinformatics software [18]. Statistical tests at p-value < 0.05 [19,20] and calculated fold changes were included for candidates of differential gene expression. The differentially expressed genes were annotated with Blast2GO bioinformatics software version 4.0. The blast-hit results were mapped to gene ontology to identify their function and improve the protein domain information with InterPro annotation (http://www.ebi.ac.uk/ interpro/search) [21-23]. The pathways of differentially expressed genes were predicted by the Kyoto Encyclopedia of Genes and Genome (KEGG) (http://www.genome.jp/kegg) [15,19,24].

Validation of differential gene expression by quantitative RT-PCR

Specific primers for high differentially expressed genes between fast-growth and slow-growth shrimp were designed based on contig sequences using the Primer 3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3) (**Table 1**). Individual body weight from 20 shrimps in each group were recorded. Muscle tissue was collected for individual RNA extraction. The purified DNase I-treated total RNA was subjected to cDNA synthesis using the iScript Select cDNA Synthesis Kit according to the manufacturer's instructions (Bio-rad, California, USA). The quantitative PCR (qPCR) reaction contained 20 ng of total RNA equivalent, 1x HOT FIREPol Evagreen qPCR Mix Plus (2.5 nM $MgCl_2$), and 100 - 250 nM forward and reverse primer, respectively. The amplification profile included an initial denaturation step at 95 ºC for 15 min and 40 cycles of 95 ºC for 15 s, 50 - 60 ºC for 15 s, and 72 ºC for 30 s using the Applied Biosystems 7300 Real-Time PCR System (Applied Biosystems). Beta-actin (GenBank accession no. JQ241179.1) was used as the reference gene. The melting curves were directly plotted from ABI-7300 software (Applied Biosystems). The obtained cycle threshold (Ct) values were calculated by the relative expression ratio. Relative mRNA expression of fast-growth and slow-growth shrimp were calculated using the real-time PCR Ct method with qPCR efficiencies (E=$10^{[-1/\text{slope}]}$) [25,26]. PCR efficiency was determined by creating standard curves from five-fold serial dilutions from the pooled total RNA (n = 5 for each group). qPCR was performed using SPSS Statistical Package version 17 (Chicago, IL, USA), and statistical significance was set at $p < 0.05$. *t*-tests were performed to compare growth performance between the groups of shrimps. The mRNA expression values of the groups of shrimp were compared using non-parametric comparisons from the Mann-Whitney U-test [27,28]. Correlations between relative gene expression and shrimp body weight were analyzed using Pearson's correlation and the corresponding probability values at $p < 0.05$.

Table 1 List of the candidate genes and primer sequences for qPCR analysis.

Contig	BLASTX annotation	Primer sequences (5' to 3')	Melting temperature	qPCR Efficiency
c32943g1	Cyclin A (*Penaeus monodon*)	TGTCGACTGGCTTGTTGAAG TCTCCTCATATTTCGCAGCA	89.0	1.88
c34539g1	Cyclin B (*Litopenaeus vannamei*)	TGGAAGTGCTACATTTTGAACG ACACCTTGGGAAGATCATGG	84.7	1.82
c23687g2	Peritrophin (*Fenneropenaeus chinensis*)	CGAACTGGTAAATGCAAGGA ATGCACTGTTTCTGGGGTCT	85.1	2.02
c34027g2	Polehole-like protein (*Penaeus monodon*)	GGTTTCCATGGTCATCCATC CTGCTCCAACATCAGTGGAA	83.7	1.85
c34340g1	Uncharacterized protein (*Hyalella azteca*) ; Amphipod	GGAGACCGTGACGAGGAC AAACCGGTGTGGTGGTAAC	90.1	1.81
c23687g1	Unknown gene	TGAGACACAATCGAAGTTCAGG AGACACCCGATAAGTGGACC	78.6	1.86
c25370g1	Unknown gene	AGCCCAGTAGTGGTCTGTGG TGTAGTTCCCTCCTGCTGCT	90.2	1.90

Results and discussion

RNA sequencing and *de novo* assembly

Five shrimps per group were collected for NGS analysis; their growth data are presented in the **Table 6**. The results of Illumina paired-end sequencing and read processing from fast-growth and slow-growth shrimp were 37,272,958 and 37,270,669 raw reads, before and after trimming, respectively (**Table 2**). The clean reads were submitted to the SRA database (https://www.ncbi.nlm. nih.gov/sra/SRP132755, SRA accession: SRP132755). All clean reads from 10 cDNA libraries were merged and *de novo* assembled using the Trinity program into 117,265 contigs with a total length of 40,813,691 nt. The median and average contig length were 264 bp and 384 bp, respectively. From cluster analysis, 113,991 genes were obtained from 38,474,021 bp with an average length of 337.52 bp (**Table 3**). **Figure 1** shows the sequence length distribution of genes and transcripts in this study.

Table 2 Summary of NGS data.

	Slow growth	**Fast growth**
Before trimming		
Number of reads	19,289,140	17,983,818
Sequence length (bp)	35 - 151	35 - 151
Percent of GC	39.9	42.2
After trimming		
Number of reads	19,288,354	17,982,315
Sequence length (bp)	36 - 151	36 - 151
Percent of GC	39.9	42.2
Percent of remaining	99.9	99.9

Table 3 Summary statistics of *de novo* assembled data.

	Transcripts	**Genes**
Total number (contigs)	117,265.00	113,991.00
N50 (b)	320.00	312.00
Median length (b)	264.00	263.00
Average length (b)	348.05	337.52
Maximum length (b)	10980.00	10953.85
Minimum length (b)	201.00	201.00
Total assembled bases (b)	40,813,691.00	38,474,021.00

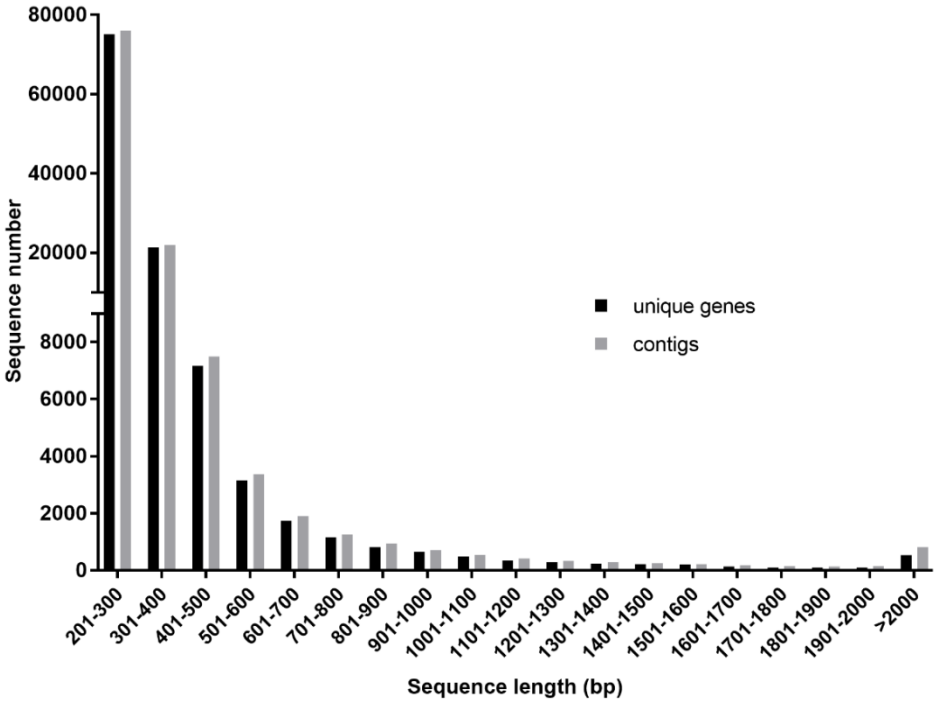

Figure 1 Length distribution of *de novo* assembled genes and transcripts of *P. monodon.*

Differentially expressed genes

Individual gene count data between the two groups were compared using the edgeR program. A total of 705 unique genes demonstrated differentially expressed value between fast-growth and slow-growth shrimp (p-value < 0.05) which contained 234 up-regulated genes and 471 down-regulated genes in fast-growth *P. monodon*. Alternatively, the genes that matched the threshold False discovery rate (FDR) < 0.001 were adjusted as high differentially expressed genes (17 up-regulated and two down-regulated genes) and used for qRT-PCR analysis (**Table 4**).

Table 4 Summary of differentially expressed genes annotation analysis.

	Genes	**Contigs**
Cut of at p-value < 0.05		
Up-regulated genes	234 (33.2 %)	324 (34.9 %)
Down-regulated genes	471 (66.8 %)	605 (65.1 %)
Total	**705**	**929**
Cut of at FDR < 0.001		
Up-regulated genes	17	
Down-regulated genes	2	
Total	**19**	

Gene ontology analysis

A BLASTX search was performed for each set of up-regulated and down-regulated genes in fast-growth *P. monodon.* After alignment against the NCBI non-redundant (nr) protein database at a cut-off *E*-value ≤ 1E-5, most sequences matched genes of crustaceans or insects, such as *P. monodon*, *Hyalella Azteca* (amphipod crustacean), *Fenneropenaeus chinensis* (Chinese white shrimp), *Pediculus humanus corporis* (louse), or *Zootermopsis nevadensis* (eusocial termite) (**Figure 2**).

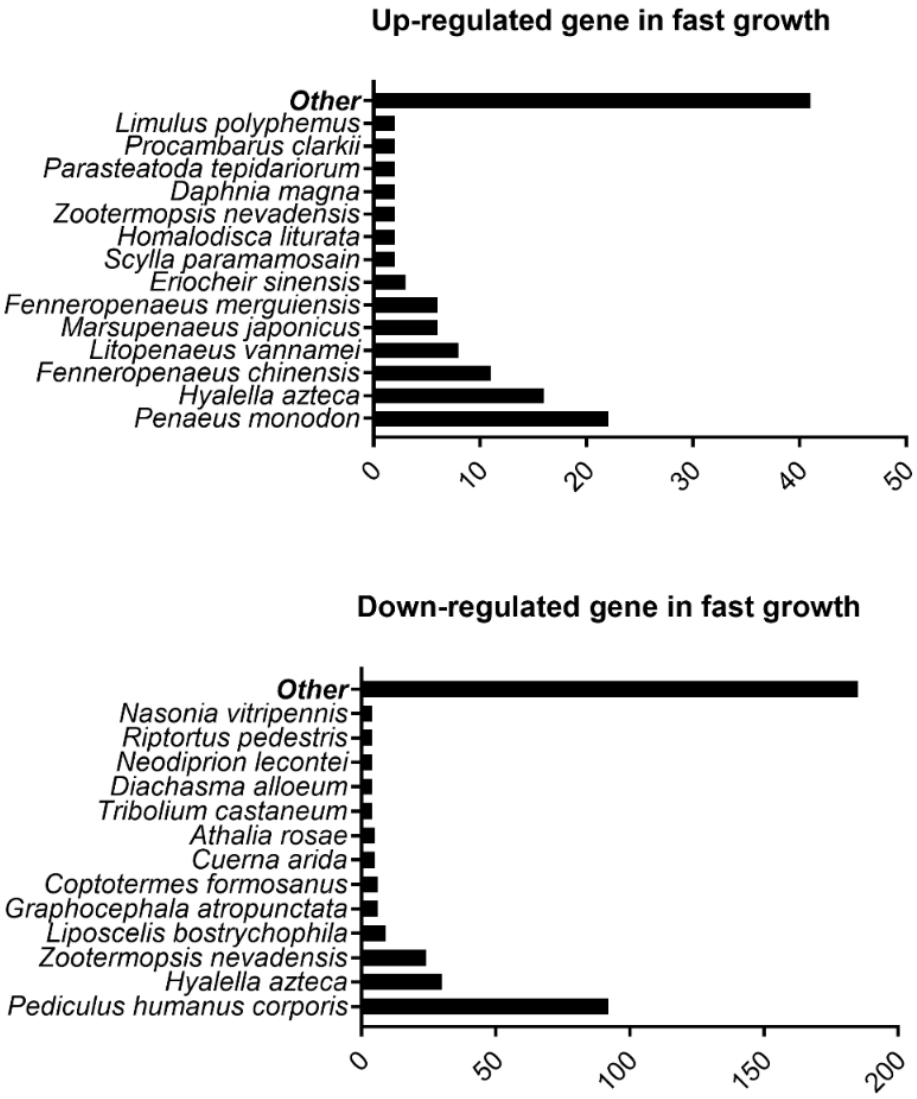

Figure 2 Species distribution searches against the database.

Table 5 List of protein domain prediction in up-regulated and down-regulated genes.

IPR accession number	Description	Protein type	Expression
IPR001304; IPR016186; IPR016187	C-type lectin-like; C-type lectin-like/link domain; C-type lectin fold	domain	up
IPR008197	WAP-type 'four-disulfide core' domain	domain	up
IPR024509; IPR024716	Anti-lipopolysaccharide factor/Scygonadin; Anti-lipopolysaccharide factor	family	up
IPR000867	Insulin-like growth factor-binding protein, IGFBP	domain	up
IPR009030	Growth factor receptor cysteine-rich domain	domain	up
IPR011390	Insulin-like growth factor binding protein-related and protein (IGFBP-rP)	family	up
IPR002181; IPR014715; IPR014716	Fibrinogen, alpha/beta/gamma chain, C-terminal globular domain; subdomain1;2	domain	up
IPR002557	Chitin binding domain	domain	up, down
IPR012336	Thioredoxin-like fold	domain	up, down
IPR001254; IPR009003	Serine proteases, trypsin domain; Peptidase S1, PA clan	domain	up, down
IPR018114	Serine proteases, trypsin family, histidine active site	active site	up, down
IPR001314	Peptidase S1A, chymotrypsin family	family	up, down
IPR027417	P-loop containing nucleoside triphosphate hydrolase	domain	down
IPR008991	Translation protein SH3-like domain	domain	down
IPR012340; IPR014722	Nucleic acid-binding, OB-fold; Ribosomal protein L2 domain 2	domain	down
IPR010987	Glutathione S-transferase, C-terminal-like	domain	down
IPR011332	Zinc-binding ribosomal protein	domain	down
IPR016040	NAD(P)-binding domain	domain	down
IPR004045; IPR004046	Glutathione S-transferase, N-terminal; C-terminal	domain	down
IPR009000	Translation protein, beta-barrel domain	domain	down
IPR000859	CUB domain	domain	down

A total of 175 of 234 up-regulated genes and 389 of 471 down-regulated genes were matched to known proteins in the InterPro database. **Table 5** presents the protein domains adjusted by a high number of matched genes. Insulin-like growth-factor-binding protein, growth-factor receptor cysteine-rich domain, and insulin-like growth-factor binding protein-related protein (IGFBP-rP) up-regulation were presented in fast-growth *P. monodon*. Interestingly, proteins including C-type lectin-domain and anti-lipopolysaccharide factor, which play a role in immunity, were found to be up-regulated in this study.

The GO term at level 2 was classified into three sub-ontologies (**Figure 3**). The up-regulated genes were mapped into five molecular function categories. Most of the proteins involved in binding, catalytic activity, and the top of the biological process are metabolic and cellular processes. Furthermore, direct GO analysis revealed that most genes encode protein involved in the binding function, while genes involved in biological processes encode protein involved in the chitin metabolic process, cell adhesion, and regulation of cell growth (**Figure 4**). Based on the KEGG pathway, up-regulated genes in fast-growth shrimp are well-represented in metabolism pathways.

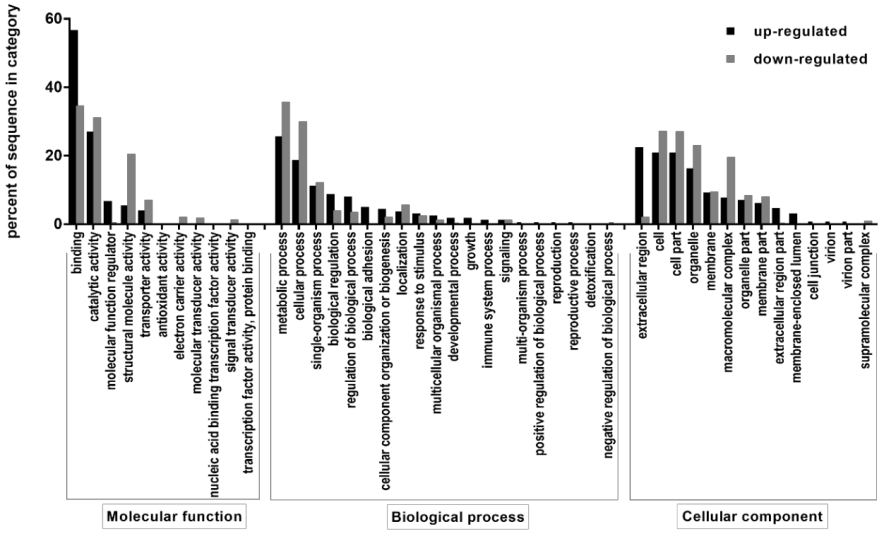

Figure 3 Gene ontology (level 2) of up-regulated and down-regulated genes in fast-growth *P. monodon*.

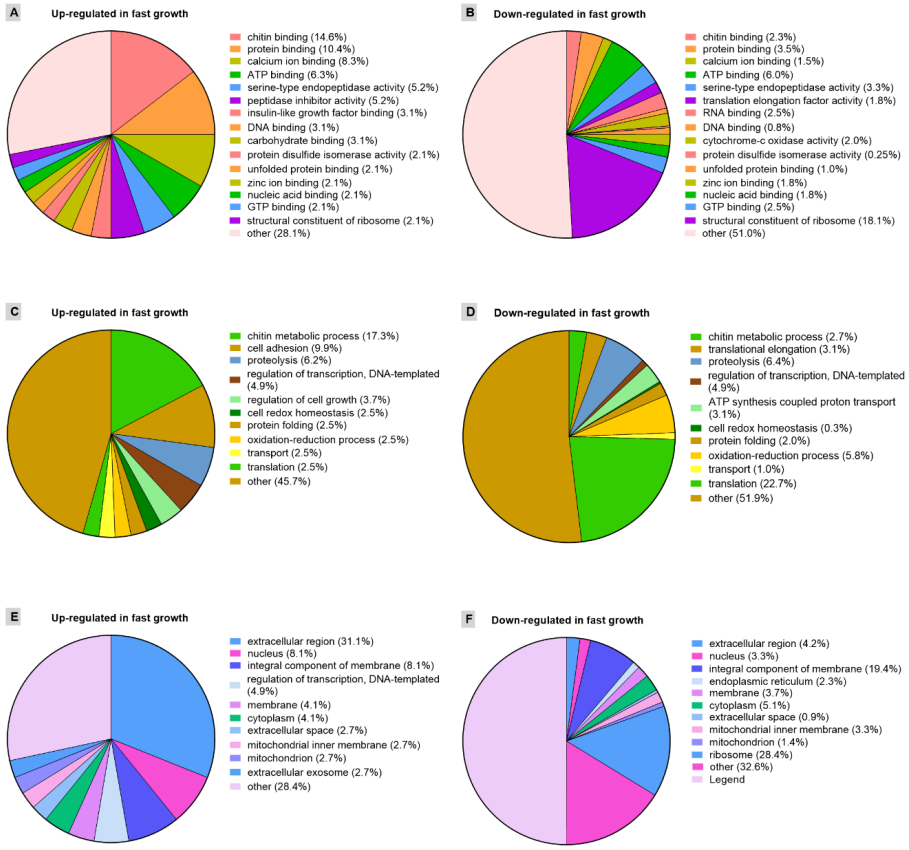

Figure 4 Gene ontology categorization of up-regulated and down-regulated genes in fast-growth *P. monodon*; molecular function (A, B), biological process (C, D), and cellular component category (E, F).

Table 6 Averages of growth data between fast growth and slow growth *P. monodon* for qPCR analysis.

	Body weight (g)	**Body length (cm)**
Slow growth	2.53 ± 0.56 [b]	5.51 ± 0.43 [b]
Fast growth	13.50 ± 3.86 [a]	9.42 ± 0.98 [a]
p-value	0.000	0.000

*Mean ± SD having the same superscript in the same column are not significantly different at 95 % confident interval of independent-sample *t*-test at *p*-value < 0.05.

Quantitative PCR validation

Quantitative PCR validation was performed in two groups of shrimp. The individual body weight of *P. monodon* of each group in average data between fast-growth and slow-growth was significantly different ($p < 0.05$) (**Table 6**). The qRT-PCR results show that seven candidate genes are highly expressed in fast-growth shrimp compared to slow-growth shrimp. These seven genes, namely cyclin-A, cyclin-B, peritrophin, polehole-like protein, uncharacterized gene I, unknown gene I, and unknown gene III, were found to be significantly highly up-regulated ($p < 0.05$) (**Figure 5**). Moreover, the mRNA levels of six genes, namely cyclin A, cyclin B, peritrophin, polehole-like protein, uncharacterized protein I, and unknown gene I ($p < 0.05$), were found to be significantly expressed and correlated with growth performance (**Table 7**).

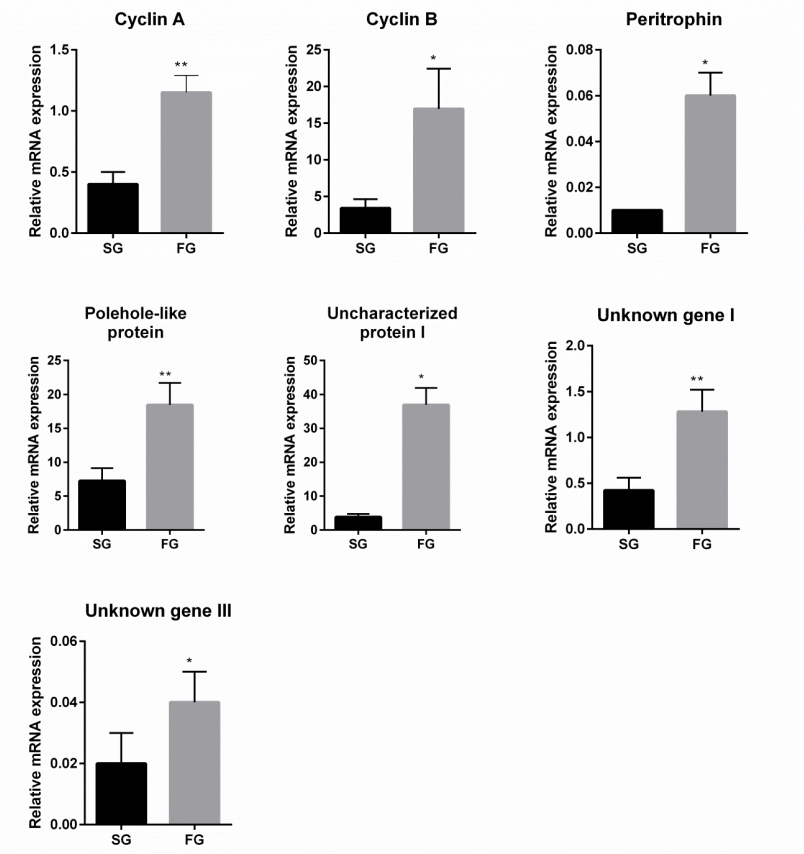

Figure 5 qRT-PCR validation results of candidate genes from RNA sequencing of fast-growth (FG) and slow-growth (SG) *P. monodon*. These data are expressed as the mean ± SEM. *; $p < 0.05$, **; $p < 0.01$.

Table 7 Correlation results of relative mRNA expression levels and shrimp body weights.

BLASTX annotation	Contig number	r[a]	*p*-value[b]
Cyclin A	c32943g1	0.728**	0.005
Cyclin B	c34539g1	0.568*	0.034
Peritrophin	c23687g2	0.677**	0.008
Polehole-like protein	c34027g2	0.520*	0.027
Uncharacterized protein I	c34340g1	0.640*	0.014
Unknown gene I	c23687g1	0.553*	0.032

[a] Correlation value: 0.00 - 0.19; very weak, 0.20 - 0.39; weak, 0.40 - 0.59; moderate, 0.60 - 0.79; strong, 0.80 - 1.00; very strong.
[b] *p*-value accepted as significantly different at $p < 0.05$.

Discussion

High-throughput RNA sequencing technologies have been used to create a reference to guarantee coverage of transcript detection [29,30]. RNA may be differentially expressed according to the tissue, stage of development, and the physiological condition in which they are accessed [30]. In this study, we focus on differentially expressed genes from *P. monodon* muscle, which is the main organ contributing to the weight and growth of shrimp [31]. RNA from the abdominal muscle of fast-growth and slow-growth shrimp was sequenced using Illumina MiSeq. The Trinity *de novo* assembly generated 113,991 genes with an average length of 337.52 bp. The difference in sequence quality is fundamentally dependent on the sample size and the sampled tissue. Equally mixed RNA from different tissues and specimens was utilized for transcriptome sequencing to obtain as many gene transcripts as possible and avoid the negative effects of individual variation [24,32]. However, the excessive number of samples presents disadvantages related to the quality of assembled sequences due to individual variation. This results in more unigenes but shorter average length, which may be partly caused by the different *de novo* assembler [19,33]. The Trinity assembly software used in this study performed consistently well [34] across the mentioned problems. Identification of protein domain for annotated transcriptome showed some sequences have roles in growth, including insulin like-growth and fibrinogen [8]. Regarding GO categories, our results correspond well with earlier studies on penaeid shrimp [19,33,35]. Focusing on fast-growth shrimp, the up-regulated genes were categorized by peptidase inhibitor activity and protein binding function. Most down-regulated gene expression is associated with structural molecule activity (ribosome). These differentially expressed candidate genes, which are highly expressed in fast-growth shrimp, may play important roles in promoting growth in *P.monodon*. Based on the KEGG database, the information is related to the metabolic pathways, which are involved in general metabolism, spliceosomes, RNA transport, phagosomes, and antioxidants [19,33,36].

Our prime interest, the genes that were observed to be highly expressed in fast-growth shrimp and related to shrimp body weight, were cyclin A, cyclin B, polehole-like protein, and peritrophin. In animal subjects, cyclins have been shown to have functional roles in cell cycle regulation [37]. Cyclins A and B have distinct, specialized roles in the cell cycle, but in all cells, both have a function in the regulation of mitotic control [38]. Cyclins A and B have been characterized in *P. monodon* (PMCyA and PMCyB) during ovarian development and spawning in normal shrimp broodstock [39]. Temporal expression of the cyclin B gene in various organs, including the liver, ovaries, muscle tissue, brain, stomach, and heart, indicate that it may play a key role in the cell division stage [40].

The polehole gene was first identified in drosophila [41]. The expression of this gene plays a fundamental role in dividing larval cells and may be linked to receptor tyrosine kinase (RTK) [42], which is an important mechanism linking the extracellular signal process in cell growth and differentiation to changes in gene expression and cell behavior [43,44]. It has also been identified as being involved in ovarian development of *Portunus trituberculatus* (swimming crab) [45]. The up-regulation of the gene in

fast-growth shrimp demonstrates the mechanism of this gene in cell proliferation, which is related to growth performance in *P.monodon*. The peritrophin gene, one of the components of the peritrophic matrix, was first isolated in the intestines of insects [46] and plays an important role in stimulating the digestion of food and providing protection from microorganisms [46,47]. In crustaceans, this protein is called shrimp ovarian peritrophin (SOP) and has been reported as a component of the egg layer in *P.semisulcatus* (marine shrimp) because it assists in the formation of a protective layer around eggs immediately after fertilization [48]. SOP has also been isolated from the oocytes of *M. japonicas*, *P. monodon*, and *F. merguiensis* [49-51]. Interestingly, peritrophin-like proteins were found to be highly expressed in fast-growth shrimp in this study, suggesting that this gene may not only be involved in ovarian development but also be related to growth development of shrimp.

The characterization of candidate genes involved in the growth of black tiger shrimp has been studied before. Eyestalk tissue from mature female shrimp (7 months old shrimp) was used in suppressive subtractive hybridization (SSH). Cyclophilin, fibrillarin, SPARC, and PC2 were reported as being involved in cell cycle and cell differentiation; they showed negative correlation with shrimp body weight [7]. In this study, we focused on positively correlated genes with growth of juvenile shrimp. The up-regulation genes involved in the cell cycle and cell division were also observed, such as cyclin A and cyclin B; from statistical analysis, they presented positive correlation to shrimp body weight. This is as per another study [8]; the genes from shrimp muscle tissue were observed, such as actin, profilin, and myosin genes. However they were not significantly different between groups of shrimp.

The up-regulation of unidentified genes in other animals or shrimp in this study may also be important for growth performance and should be further studied to develop markers related to the growth of *P. monodon*. From this study, concerning other highly expressed genes in fast-growth shrimp from the metabolic pathway, biological process, and domain in protein, from the NGS results, some of these were reported as being possibly involved with growth in other animals; these also will be further studied to obtain more information and understanding.

Conclusions

In the present study, transcriptome analysis of abdominal muscle in fast-growth and slow-growth *P. monodon* was performed using the RNA sequencing technique. *De novo* assembly resulted in 113,991 genes with an average length of 337.52 bp and transcriptome coverage of about 38.47 Mb. A comparison of gene expression between fast-growth and slow-growth shrimp revealed a total of 705 genes, including 234 up-regulated and 471 down-regulated genes ($p < 0.05$). Quantitative RT-PCR revealed highly expressed genes in fast-growth shrimp, namely the cyclin A and cyclin B genes, peritrophin gene, and polehole-like protein gene, all of which are related to growth performance. Moreover, the interesting gene ontology established in molecular functional protein binding is related to biological processes such as the cell regulation of cell growth. All of these are candidate genes for further study to develop a marker for growth performance detection in *P. monodon*.

Acknowledgement

This research was supported by the Program Strategic Scholarships Fellowships Frontier Research Networks (Specific for Southern Region) for the Ph.D. Program, from the Office of the Higher Education Commission, Thailand, and also by the Research Center of Excellence on Shrimp.

References

[1] M Saroglia and Z Liu. *Functional Genomics in Aquaculture*. Wiley, 2012, p. 1-403.
[2] FAO. *The State of World Fisheries and Aquaculture 2016*. Food and Agriculture Organization of the United Nations, 2016.
[3] AW Fast and LJ Lester. *Marine Shrimp Culture: Principles and Practices*. Elsevier, 1992, p. 879.
[4] H Jung, RE Lyons, DA Hurwood and PB Mather. Genes and growth performance in crustacean species: A review of relevant genomic studies in crustaceans and other taxa. *Rev. Aquaculture* 2013; **5**, 77-110.
[5] C De-Santis and DR Jerry. Candidate growth genes in finfish: Where should we be looking? *Aquaculture* 2007; **272**, 22-38.
[6] H Jung, RE Lyons, Y Li, NM Thanh, H Dinh, DA Hurwood, KR Salin and PB Mather. A candidate gene association study for growth performance in an improved giant freshwater prawn (*Macrobrachium rosenbergii*) culture line. *Mar. Biotechnol.* 2014; **16**, 161-80.
[7] A Tangprasittipap, M Tiensuwan and B Withyachumnarnkul. Characterization of candidate genes involved in growth of black tiger shrimp *Penaeus monodon*. *Aquaculture* 2010; **307**, 150-6.
[8] C Nguyen, TG Nguyen, LV Nguyen, HQ Pham, TH Nguyen, HT Pham, HT Nguyen, TT Ha, TH Dau, HT Vu, DD Nguyen, NTT Nguyen, NH Nguyen, DV Quyen, HH Chu and KD Dinh. *De novo* assembly and transcriptome characterization of major growth-related genes in various tissues of *Penaeus monodon*. *Aquaculture* 2016; **464**, 545-53.
[9] CA Santos, DV Blanck and PD de Freitas. RNA-seq as a powerful tool for penaeid shrimp genetic progress. *Front. Genet.* 2014; **5**, 1-6.
[10] AM Bolger, M Lohse and B Usadel. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**, 2114-20.
[11] MG Grabherr, BJ Haas, M Yassour, JZ Levin, DA Thompson, I Amit, X Adiconis, L Fan, R Raychowdhury, Q Zeng, Z Chen, E Mauceli, N Hacohen, A Gnirke, N Rhind, FD Palma, BW Birren, C Nusbaum, K Lindblad-Toh, N Friedman and A Regev. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol*. 2011; **29**, 1-11.
[12] J Feng, W Li and T Jiang. Inference of isoforms from short sequence reads. *J. Comput. Biol.* 2011; **18**, 1-21.
[13] B Li and CN Dewey. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; **12**, 1-16.
[14] M-A Dillies, A Rau, J Aubert, C Hennequet-Antier, M Jeanmougin, N Servant, C Keime, G Marot, D Castel, J Estelle, G Guernec, B Jagla, L Jouneau, D Laloë, C Le Gall, B Schaëffer, SL Crom, M Guedj and F Jaffrézic. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings Bioinformatics* 2013; **14**, 671-83.
[15] S Li, X Zhang, Z Sun, F Li and J Xiang. Transcriptome analysis on Chinese shrimp *Fenneropenaeus chinensis* during WSSV acute infection. *PLoS One* 2013; **8**, 1-15.
[16] C Trapnell, BA Williams, G Pertea, A Mortazavi, G Kwan, MJ van Baren, SL Salzberg, BJ Wold and L Pachter. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 2010; **28**, 511-18.
[17] BJ Haas, A Papanicolaou, M Yassour, M Grabherr, PD Blood, J Bowden, MB Couger, D Eccles, B Li, M Lieber, MD MacManes, M Ott, J Orvis, N Pochet, F Strozzi, N Weeks, R Westerman, T William, CN Dewey, R Henschel, RD LeDuc, N Friedman and A Regev. *De novo* transcript sequence reconstruction from RNA-Seq: Reference generation and analysis with Trinity. *Nat. Protocol.* 2013; **8**, 1-43.
[18] MD Robinson, DJ McCarthy and GK Smyth. edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; **26**, 1-2.
[19] S Sookruksawong, F Sun, Z Liu and A Tassanakajon. RNA-Seq analysis reveals genes associated with resistance to Taura syndrome virus (TSV) in the Pacific white shrimp *Litopenaeus vannamei*. *Dev. Comp. Immunol.* 2013; **41**, 523-33.

[20] J Gao, X Wang, Z Zou, X Jia, Y Wang and Z Zhang. Transcriptome analysis of the differences in gene expression between testis and ovary in green mud crab (*Scylla paramamosain*). *BMC Genomics* 2014; **15**, 1-15.
[21] H Jung, RE Lyons, H Dinh, DA Hurwood, S McWilliam and PB Mather. Transcriptomics of a giant freshwater prawn (*Macrobrachium rosenbergii*): *De novo* assembly, annotation and marker discovery. *PLoS One* 2011; **6**, 1-14.
[22] A Conesa, S Gotz, JM Garcia-Gomez, J Terol, M Talon and M Robles. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005; **21**, 3674-76.
[23] S Götz, JM García-Gómez, J Terol, TD Williams, SH Nagaraj, MJ Nueda, M Robles, M Talón, J Dopazo and A Conesa. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic. Acids. Res.* 2008; **36**, 3420-35.
[24] H Guo, CX Ye, AL Wang, JA Xian, SA Liao, YT Miao and SP Zhang. Trascriptome analysis of the Pacific white shrimp *Litopenaeus vannamei* exposed to nitrite by RNA-seq. *Fish Shellfish Immunol.* 2013; **35**, 2008-16.
[25] MW Pfaffl. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic. Acids Res.* 2001; **29**, 1-6.
[26] MW Pfaffl, GW Horgan and L Dempfle. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic. Acids Res.* 2002; **30**, 1-10.
[27] JP Baren, GD Stewart, A Stokes, K Gray, CJ Pennington, R O'Neill, DAC Deans, S Paterson-Brown, ACP Riddick, DR Edwards, KCH Fearon, JA Ross and RJE Skipworth. mRNA profiling of the cancer degradome in oesophago-gastric adenocarcinoma. *Br. J. Canc. Manag.* 2012; **107**, 143-49.
[28] JD Neve, O Thas, JP Ottoy and L Clement. An extension of the Wilcoxon-Mann-Whitney test for analyzing RT-qPCR data. *Stat. Appl. Genet. Mol. Biol.* 2013; **12**, 333-46.
[29] A Oshlack, MD Robinson and MD Young. From RNA-seq reads to differential expression results. *Genome. Biol.* 2010; **11**, 1-10.
[30] Z Wang, M Gerstein and M Snyder. RNA-Seq: A revolutionary tool for transcriptome. *Nat. Rev. Genet.* 2009; **10**, 1-7.
[31] JRO Cesar, B Zhao, S Malecha, H Ako and J Yang. Morphological and biochemical changes in the muscle of the marine shrimp *Litopenaeus vannamei* during the molt cycle. *Aquaculture* 2006; **261**, 688-94.
[32] X Fu, N Fu, S Guo, Z Yan, Y Xu, H Hu, C Menzel, W Chen, Y Li, R Zeng and P Khaitovich. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 2009; **10**, 1-12.
[33] S Xue, Y Liu, Y Zhang, Y Sun, X Geng and J Sun. Sequencing and de novo analysis of the hemocytes transcriptome in *Litopenaeus vannamei* response to white spot syndrome virus infection. *PLoS One* 2013; **8**, 1-12.
[34] QY Zhao, Y Wang, YM Kong, D Luo, X Li and P Hao. Optimizing de novo transcriptome assembly from short-read RNA-Seq data: A comparative study. *BMC Bioinformatics* 2011; **12**, 1-12.
[35] D Zeng, X Chen, D Xie, Y Zhao, C Yang, Y Li, N Ma, M Peng, Q Yang, Z Liao, H Wang and X Chen. Transcriptome analysis of Pacific white shrimp (*Litopenaeus vannamei*) Hepatopancreas in response to taura syndrome virus (TSV) experimental infection. *PLoS One* 2013; **8**, 1-8.
[36] C Li, S Weng, Y Chen, X Yu, L Lü, H Zhang, J He and X Xu. Analysis of *Litopenaeus vannamei* transcriptome using the next-generation DNA sequencing technique. *PLoS One* 2012; **7**, 1-12.
[37] J Pines and T Hunter. Cyclins A and B1 in the human cell cycle. *Ciba. Found. Symp.* 1992; **170**, 187-96.
[38] CF Lehner and PH O'Farrell. The roles of drosophila cyclins A and B in mitotic control. *Cell* 1990; **61**, 535-47.

[39] V Visudtiphole, S Klinbunga and K Kirtikara. Molecular characterization and expression profiles of cyclin A and cyclin B during ovarian development of the giant tiger shrimp *Penaeus monodon*. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 2009; **152**, 535-43.
[40] L Qiu, S Jiang, F Zhou, J Huang and Y Guo. Molecular cloning and characterization of a cyclin B gene on the ovarian maturation stage of black tiger shrimp (*Penaeus monodon*). *Mol. Biol. Rep.* 2007; **24**, 24.
[41] N Perrimon, L Engstrom and AP Mahowald. A pupal lethal mutation with a paternally influenced maternal effect on embryonic development in *Drosophila melanogaster*. *Dev. Biol.* 1985; **110**, 480-91.
[42] M Furriols and J Casanova. In and out of Torso RTK signalling. *EMBO J* 2003; **22**, 1947-52.
[43] G Jiménez, A Guichet, A Ephrussi and J Casanova. Relief of gene repression by torso RTK signaling: Role of capicua in Drosophila terminal and dorsoventral patterning. *Gene. Dev.* 2000; **14**, 224-31.
[44] P Li, K Wood, H Mamon, W Haser and T Roberts. Raf-1: A kinase currently without a cause but not lacking in effects. *Cell* 1991; **64**, 479-82.
[45] ZB Yu, CK Mu, WW Song, RH Li, YE Chen and CL Wang. Screening of genes related to ovarian development in the swimming crab, *Portunus trituberculatus*, by suppression subtractive hybridization. *Genet. Mol. Res*. 2015, **14**, 18675-86.
[46] XJ Du, JX Wang, N Liu, XF Zhao, FH Li and JH Xiang. Identification and molecular characterization of a peritrophin-like protein from fleshy prawn (*Fenneropenaeus chinensis*). *Mol. Immunol.* 2006; **43**, 1633-44.
[47] L Wang, F Li, B Wang and J Xiang. A new shrimp peritrophin-like gene from *Exopalaemon carinicauda* involved in white spot syndrome virus (WSSV) infection. *Fish Shellfish Immunol.* 2013; **35**, 840-46.
[48] M Khayat, PJ Babin, B Funkenstein, M Sammar, H Nagasawa, A Tietz and E Lubzens. Molecular characterization and high expression during oocyte development of a shrimp ovarian cortical rod protein homologous to insect intestinal peritrophins. *Biol. Reprod.* 2001; **64**, 1090-99.
[49] Y Chen, T Souaiaia and T Chen. PerM: Efficient mapping of short sequencing reads with periodic full sensitive spaced seeds. *Bioinformatics* 2009; **25**, 2514-21.
[50] YK Kim, N Tsutsui, I Kawazoe, T Okumura, T Kaneko and K Aida. Localization and developmental expression of mRNA for cortical rod protein in kuruma prawn *Marsupenaeus japonicus*. *Zool. Sci.* 2005; **22**, 675-80.
[51] W Loongyai, JC Avarre, M Cerutti, E Lubzens and W Chotigeat. Isolation and functional characterization of a new shrimp ovarian peritrophin with antimicrobial activity from *Fenneropenaeus merguiensis*. *Mar. Biotechnol.* 2007; **9**, 624-37.