



# Comparative transcriptome analysis reveals three potential antiviral signaling pathways in lymph organ tissue of the red swamp crayfish, *Procambarus clarkii*

Z.-Q. Du

School of Life Science and Technology,  
Inner Mongolia University of Science and Technology, Kundulun District,  
Baotou, Inner Mongolia Autonomous Region, China

Corresponding author: Z.-Q. Du  
E-mail: nmdzq1981@163.com

Genet. Mol. Res. 15 (4): gmr15048858  
Received June 7, 2016  
Accepted July 20, 2016  
Published November 21, 2016  
DOI <http://dx.doi.org/10.4238/gmr15048858>

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

**ABSTRACT.** The red swamp crayfish (*Procambarus clarkii*) is one of the most economically important farmed aquatic species in China. Compared with its relatively well-known antibacterial and antifungal mechanisms, the antiviral mechanism is still unclear. We used Illumina-based RNA sequencing and bioinformatic technology to obtain high-quality sequence reads from the crayfish lymph organ. A total of 5933 differentially expressed genes (DEGs) were identified between normal and white spot syndrome virus-challenged samples. Of these, 4638 genes were differentially upregulated and 1295 differentially downregulated by more than two-fold. The DEGs were then mapped to different signaling pathways; the Janus kinase/signal transducers and activators of transcription, insulin, and Wnt signaling pathways

were predicted to be involved in crayfish antiviral innate immunity. These results provide new insights into crayfish antiviral immunity mechanisms.

**Key words:** *Procambarus clarkii*; White spot syndrome virus; Lymph organ; Illumina sequencing; *De novo* assembly; Comparative transcriptome

## INTRODUCTION

Over the course of evolution, organisms have developed advanced immune systems in order to defend against pathogen invasion by continuous confrontations between hosts and pathogens. Both vertebrates and invertebrates have developed specific immune responses. Unlike vertebrates, invertebrates lack an acquired immune system, but develop an innate immune system that includes both cellular and humoral immune responses (Du et al., 2010). When hosts are challenged by pathogens, genes can be synergistically mobilized to play their respective roles in defense, particularly in the humoral immune response (Taffoni and Pujol, 2015). It is crucial to study immune-related gene functions in order to understand the coordination mechanisms of the innate immune system.

The red swamp crayfish (*Procambarus clarkii*) is a typical invertebrate, and is used as a model organism to study the invertebrate innate immune system. It is native to northeastern Mexico and the Southern USA, and was introduced to China from Japan in the 1930s (Shen et al., 2014). Because of its resilience, adaptability to changing environments, and high fecundity, the species has been widely cultured in China (Manfrin et al., 2015), and has become one of the most economically important farmed aquatic species in the country. The excellent resistance of the red swamp crayfish to bacteria, fungi, and viruses is well known. Recent research has elucidated the species' antibacterial and antifungal mechanisms, such as the Toll and Imd pathways (Ermolaeva and Schumacher, 2014), but the antiviral mechanism is still not known. Micro RNAs might play a role in antiviral innate immune responses (Zeng et al., 2015); however, it is necessary to identify antiviral genes and antiviral-related signaling pathways using transcriptome sequencing. The transcriptome sequencing of crayfish tissues, including the hepatopancreas, muscle, ovary, testis, eyestalk, cuticular epidermis, branchia, intestine, and stomach has recently been conducted (Shen et al., 2014; Manfrin et al., 2015); however, transcriptome data for the lymph organ and white spot syndrome virus (WSSV)-challenged tissues have not been reported. The crayfish lymph organ is vital for the removal of pathogens, and is an important research tool for investigating the innate immune response mechanism.

In recent years, next-generation sequencing (NGS) technology has been widely used to explore the genetic information of model organisms (Jiang et al., 2014). NGS technology is superior in many respects to traditional Sanger sequencing technology, particularly in terms of cost and time (Martin and Wang, 2011). The expressed sequences produced using NGS technology are usually ten- to one-hundred-fold greater in number than those identified by traditional Sanger sequencing technologies (Christie et al., 2015). In a previous study, the Janus kinase/signal transducers and activators of transcription (JAK-STAT) signaling pathway was found to play an important role in crayfish intestinal antiviral immunity (Du et al., 2016). In the present study, we sequenced the lymph organ transcriptomes of normal and

WSSV-challenged crayfish in order to generate expression profiles and identify differentially expressed genes (DEGs) between normal crayfish's lymph organs and WSSV-challenged crayfish's lymph organs. The data obtained will provide an important resource for research on gene function, molecular events, and signaling pathways that are related to the invertebrate antiviral immune response.

## MATERIAL AND METHODS

### Ethics statement

The experimental procedures complied with the current applicable laws of China where they were performed. No specific permits were required for the research undertaken. Crayfish individuals were maintained under appropriate laboratory conditions to guarantee their welfare and responsiveness.

### Preparation of crayfish tissues and WSSV challenge

*P. clarkii* (weighing approximately 15-20 g) were bought from an aquaculture commercial market in Hangzhou, Zhejiang Province, China. The crayfish were initially cultured in water tanks at 26°-28°C for 10 days, and were fed twice a day with artificial food throughout the experiment (Li et al., 2012b). WSSV ( $3.2 \times 10^7$  copies per crayfish) was injected into the abdominal segment of each crayfish; 36 h later, the lymph organ was removed from at least 10 WSSV-challenged crayfish [Group WSSV-challenged (GW)]. The lymph organs were also removed from at least 10 normal crayfish and frozen immediately in liquid nitrogen [Group normal (GN)]. The two sets of samples were temporarily stored at -80°C for total RNA extraction (Du and Jin, 2015).

### RNA isolation and Illumina sequencing

The samples were delivered to the Beijing Genomics Institute-Shenzhen (Shenzhen, China) for total RNA extraction. According to the company's sequencing report, total RNA was extracted using TRIzol reagent (Invitrogen, USA) in accordance with the manufacturer protocol. The quality of RNA samples treated with DNase I (Invitrogen) was examined for the subsequent procedures, including mRNA purification, cDNA library construction, and transcriptome sequencing. In brief, about 5 µg DNase-treated total RNA was used to construct a cDNA library following the protocol of the Illumina TruSeq RNA Sample Preparation Kit (Illumina, USA). After some necessary quantifications and qualifications, the library was sequenced using an Illumina HiSeq™ 2000 instrument with 100-bp paired-end reads for GN and GW.

### De novo assembly and transcriptome analysis

*De novo* transcriptome assembly of the two sets of samples (GN and GW) was conducted using the RNA-Seq *de novo* assembly program, Trinity (Dedeine et al., 2015), with the default parameters set (Ali et al., 2015). In brief, raw reads generated by the Illumina HiSeq™ 2000 sequencer were trimmed by removing the adapter sequences. After low-quality reads (quality scores lower than 20) and short-length reads (below 10 bp) were removed, high-

quality clean reads were obtained. Three processes were then performed using Inchworm, Chrysalis, and Butterfly (He et al., 2015). Firstly, high-quality clean reads were processed by Inchworm to form longer fragments called contigs. The contigs were connected by Chrysalis to obtain unigenes that could not be extended at either end; the unigenes resulted in de Bruijn graphs. Finally, the de Bruijn graphs were processed using Butterfly in order to obtain transcripts (Li et al., 2013).

### **Transcriptome annotation and Gene Ontology (GO) analysis**

After the *de novo* transcriptome assembly was completed, the transcripts were used to conduct annotation (Mousavi et al., 2014), including protein functional annotation, Cluster of Orthologous Groups (COG) functional annotation, GO functional annotation ([www.geneontology.org](http://www.geneontology.org)), and pathway annotation. All of these were based on sequence similarity with known genes. The assembled contigs were annotated with sequences available in the National Center for Biotechnology Information (NCBI) database using the BLASTx and BLASTn algorithms (Leng et al., 2015). The unigenes were aligned by a BLASTx search against NCBI protein databases, including the non-redundant sequence database (Nr), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG), and COG (Li et al., 2015). None of the BLASTx hits were aligned by a BLASTn search against the NCBI non-redundant nucleotide database (Nt). The alignments were conducted in order to establish the homology of sequences with known genes with a cutoff E-value  $\leq 10^{-5}$  (Li et al., 2012a). The best alignment results were used to predict the sequence direction and the coding DNA sequence of the unigenes. Functional annotation was conducted using GO terms that were analyzed using the Blast2go software (<http://www.blast2go.com/b2ghome>). The complex biological behavior of the genes was analyzed by pathway annotation based on the KEGG database.

### **DEG identification**

To ascertain transcript expression levels in the crayfish lymph organ, clean reads were first mapped to all of the transcripts using the Bowtie software (Langmead and Salzberg, 2012). DEGs were then obtained on the basis of fragments per kilobase of exon per million fragments mapped (FPKM) of the genes, followed by a false discovery rate (FDR) control to correct the P values (Mortazavi et al., 2008). DEGs were identified using the EDGER software. For this analysis, the filtering threshold was set as an FDR control, 0.5.  $FDR \leq 0.001$  and the absolute value of  $\log_2 \text{Ratio} \geq 1$  were used as the filtering threshold to determine DEG significance (Banani et al., 2016). DEGs were identified between GW and GN by a comparative analysis of the data.

### **Quantitative real-time polymerase chain reaction (PCR) validation**

Quantitative real-time PCR (qRT-PCR) methods were used to determine the RNA levels of 10 randomly selected genes (Wang et al., 2016). cDNA templates from the two sets of samples (GN and GW) were diluted 30-fold in nuclease-free water and used as templates for the PCR. Gene-specific primer sequences were carefully designed using the Primer Premier 6 software, based on each gene sequence identified in the transcriptome library. The specific primers *Pc-18 S RNA*-qRT-F (5'-TCT TCT TAG AGG GAT TAG CCG-3') and *Pc-18 S RNA*-qRT-R (5'-AAG GGG ATT GAA CCG GTT A-3') were used to amplify *18 S RNA* as an inner

control. qRT-PCR was performed using SYBR<sup>®</sup> *Premix Ex Taq* (TaKaRa, Japan) and a real-time thermal cycler (CFX96 touch, Bio-Rad, USA) following the manufacturer instructions, in a total volume of 10  $\mu$ L that contained 5  $\mu$ L 2X *Premix Ex Taq*, 1  $\mu$ L 1:30 diluted cDNA, and 2  $\mu$ L (1  $\mu$ M) each of the forward and reverse primers. The amplification procedure comprised an initial denaturation step at 95°C for 3 min, followed by 40 cycles at 95°C for 15 s, 60°C for 40 s, and melting from 65° to 95°C. Three parallel experiments were performed to improve the integrity of the study. Differentially expressed levels of the target genes (between GW and GN) were calculated using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001). The data obtained were subjected to an unpaired-samples Student *t*-test. Significance was set at  $P < 0.05$ .

## RESULTS

### Illumina sequencing of the crayfish lymph organ

After cleaning and quality testing the GN samples, 44,264,442 clean reads were screened from 46,182,460 raw reads, which corresponded to 4,426,444,200 total nucleotides (nt). The Q20 (percentage of bases the quality of which was greater than 20 in clean reads), N (percentage of uncertain bases after filtering), and GC percentages were 97.92, 0.00, and 42.99%, respectively. For the GW samples, 50,263,784 clean reads were screened from 52,010,554 raw reads, which corresponded to 5,026,378,400 total nt. The Q20, N, and GC percentages were 97.86, 0.00, and 43.68%, respectively. All of these sequences were used for further analysis.

### *De novo* assembly of the transcriptome

For the GN samples, 44,264,442 high-quality clean reads were randomly assembled to produce 105,034 contigs with an N50 of 546 bp. The contigs were further assembled and clustered into 56,922 unigenes with a mean length of 564 nt, which were composed of 4886 distinct clusters and 52,036 distinct singletons. The N50 of the unigenes was 1032 bp. For the GW samples, 50,263,784 high-quality clean reads were randomly assembled to produce 137,886 contigs with an N50 of 611 bp. The contigs were further assembled and clustered into 74,732 unigenes with a mean length of 606 nt, which were composed of 7646 distinct clusters and 67,086 distinct singletons. The N50 of the unigenes was 1197 bp.

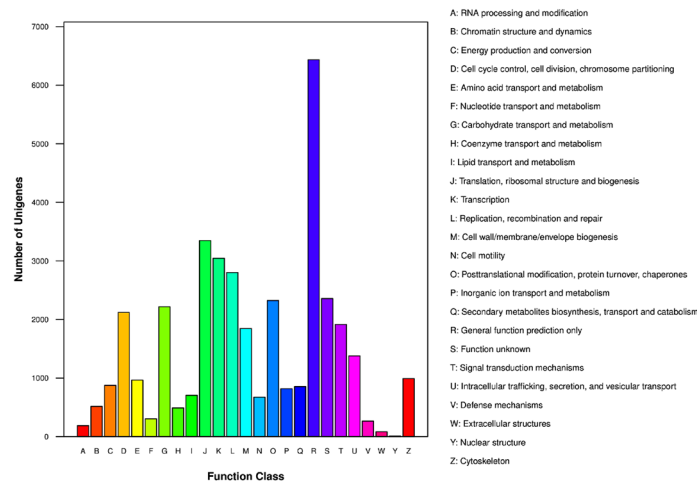
In the GN samples, the unigenes were mainly between 200 and 3000 nt long. The percentage of unigenes that were between 201 and 1000 bp long was 86.9% (49,456); between 1001 and 2000 bp long it was 8.0% (4547); between 2001 and 3000 bp long it was 2.7% (1538); and for over 3000 bp long it was 2.4% (1381). In the GW samples, the percentage of unigenes that were between 201 and 1000 bp long was 85.5% (63,916); between 1001 and 2000 bp long it was 8.5% (6337); between 2001 and 3000 bp long it was 3.1% (2303); and for over 3000 bp long it was 2.9% (2176).

### Functional annotation of predicted proteins

Sequence annotation was conducted based on the unigenes from the merged group. Subsequently, the putative functions of the unigenes were analyzed based on their GO and COG classifications (Young et al., 2010). GO classification is an international, standardized

gene function classification system, which provides a dynamically updated controlled vocabulary and an exactly defined conception to describe gene characteristics and their products in any organism (Di Lena et al., 2015). GO classification consists of three main categories: “biological process”, “cellular component”, and “molecular function” (Chen et al., 2015a). A total of 13,848 transcripts annotated in the GO database were categorized into 58 functional groups, including the three main GO categories. Among these functional groups, the “biological regulation”, “cellular process”, “metabolic process”, “cell”, “single-organism process”, “cell part”, “binding”, and “catalytic activity” terms were dominant.

COGs were delineated by comparing protein sequences encoded in complete genomes that represented major phylogenetic lineages (Tatusov et al., 2003). COG classification analysis is an important method of unigene functional annotation and evolutionary research (Lai and Lin, 2013). A total of 37,531 unigenes were mapped to 25 COG categories (Figure 1). The largest COG group was the R category, which represented “general function prediction only” (6435 unigenes), followed by the J category that represented “translation, ribosomal structure, and biogenesis” (3347 unigenes), the K category that represented “transcription” (3044 unigenes), and the L category that represented “replication, recombination, and repair” (2802 unigenes).



**Figure 1.** Cluster of Orthologous Groups (COG) classification of putative proteins.

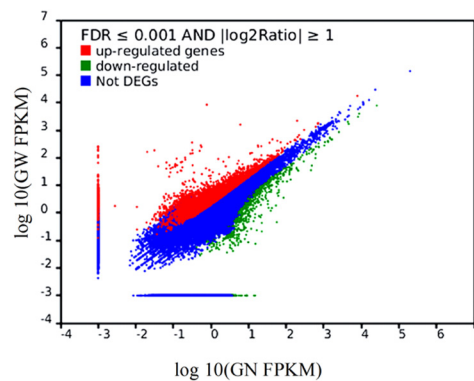
KEGG is a bioinformatic resource for linking genomes to life and the environment (Mao et al., 2005). The PATHWAY database record networks of molecular interactions in cells, and variants of them that are specific to particular organisms (Chen et al., 2015b). The genes from the merged groups (GN and GW) were categorized using the KEGG database in order to predict the unigenes’ functions (Feng et al., 2015). A total of 25,290 unigenes were classified into 257 KEGG pathways, the 25 most statistically significant of which are shown in Table 1. Important innate immunity-related pathways were predicted, including *Vibrio cholerae* infection (1092 sequences), focal adhesion (910 sequences), Epstein-Barr virus infection (860 sequences), HTLV-I infection (596 sequences), *Herpes simplex* infection (593 sequences), *Salmonella* infection (576 sequences), lysosome (610 sequences), and the MAPK signaling pathway (542 sequences). The mRNA surveillance pathway, endocytosis, phagosome, and tight junction were amongst the 25 most statistically significant KEGG classifications.

**Table 1.** Top 25 statistically significant Kyoto Encyclopedia of Genes and Genomes classifications.

No.	Pathway	Pathway definition	Number of sequences
1	path: ko01100	Metabolic pathways	3371 (13.33%)
2	path: ko05146	Amoebiasis	1148 (4.54%)
3	path: ko05110	<i>Vibrio cholerae</i> infection	1092 (4.32%)
4	path: ko05016	Huntington's disease	973 (3.85%)
5	path: ko04810	Regulation of actin cytoskeleton	950 (3.76%)
6	path: ko04510	Focal adhesion	910 (3.6%)
7	path: ko03040	Spliceosome	894 (3.53%)
8	path: ko05200	Pathways in cancer	879 (3.48%)
9	path: ko05169	Epstein-Barr virus infection	860 (3.4%)
10	path: ko03013	RNA transport	847 (3.35%)
11	path: ko00230	Purine metabolism	781 (3.09%)
12	path: ko04145	Phagosome	643 (2.54%)
13	path: ko04144	Endocytosis	638 (2.52%)
14	path: ko04270	Vascular smooth muscle contraction	633 (2.5%)
15	path: ko03010	Ribosome	632 (2.5%)
16	path: ko04530	Tight junction	616 (2.44%)
17	path: ko00240	Pyrimidine metabolism	616 (2.44%)
18	path: ko04142	Lysosome	610 (2.41%)
19	path: ko04141	Protein processing in endoplasmic reticulum	607 (2.4%)
20	path: ko05166	HTLV-I infection	596 (2.36%)
21	path: ko05168	Herpes simplex infection	593 (2.34%)
22	path: ko05132	Salmonella infection	576 (2.28%)
23	path: ko03015	mRNA surveillance pathway	563 (2.23%)
24	path: ko04120	Ubiquitin mediated proteolysis	557 (2.2%)
25	path: ko04010	MAPK signaling pathway	542 (2.14%)

## DEG analysis of the crayfish lymph organ after WSSV infection

As shown in Figure 2, 5933 DEGs were screened after a comparative analysis of the GN and GW samples; 4638 were differentially upregulated and 1295 were differentially downregulated by more than two-fold. Of the 5933 DEGs, 5603 were in both the GN and GW groups, including 4393 differentially upregulated and 1210 differentially downregulated genes. In addition, 245 genes were only found in the GW group and 85 were only found in the GN group.



**Figure 2.** Comparative analysis of gene expression levels for two transcript libraries of normal crayfish lymph organ (GN) and white spot syndrome virus-challenged crayfish lymph organ (GW). Red dots represent transcripts that were significantly upregulated in GW and green dots indicate transcripts that were significantly downregulated. The parameters “ $FDR \leq 0.001$ ” and “ $\log_2\text{Ratio} \geq 1$ ” were used as the criteria by which to judge the significance of the gene expression difference. DEG, differentially expressed gene; FPKM, fragments per kilobase of exon per million fragments mapped.

In order to confirm the biological functions of the DEGs in the GN and GW groups, GO classification and KEGG pathway analysis were conducted (Gupta et al., 2015). GO analysis of the 5933 DEGs revealed that 47 GO terms were significantly expressed (either up or down;  $P < 0.05$ ) and 36 were highly significantly expressed (either up or down;  $P < 0.01$ ). Among the 47 GO terms that had P values lower than 0.05, 31, 7, and 9 belonged to “cellular component”, “molecular function”, and “biological process”, respectively. Among the 36 GO terms that had P values lower than 0.01, 25, 5, and 6 belonged to “cellular component”, “molecular function”, and “biological process”, respectively. Gene clusters with significant expression differentiation were mainly related to “cellular component”.

The KEGG pathway analysis revealed that 35 pathways were significantly different ( $P < 0.05$ ) between the groups. Of these, 24 were highly significantly different between the groups ( $P < 0.01$ ). Some were related to crayfish innate immunity, including the JAK-STAT signaling pathway, insulin signaling pathway, mRNA surveillance pathway, Wnt signaling pathway, extracellular matrix-receptor interaction, cell adhesion molecules, peroxisome, lysosome, RNA degradation, and basal transcription factors (Table 2).

**Table 2.** Top 35 differentially expressed pathways between normal crayfish lymph organ (GN) and white spot syndrome virus-challenged crayfish lymph organ (GW).

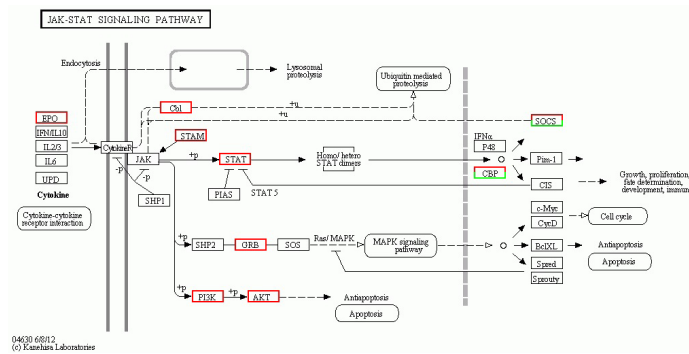
No.	Pathway	Number of DEGs	P value	Pathway ID
1	Glycolysis/Gluconeogenesis	27	3.582649e-13	ko00010
2	ECM-receptor interaction	76	2.300117e-11	ko04512
3	JAK-STAT signaling pathway	18	5.021604e-11	ko04630
4	Valine, leucine, and isoleucine degradation	6	1.128055e-08	ko00280
5	Insulin signaling pathway	38	3.260367e-08	ko04910
6	Pancreatic secretion	42	2.195456e-07	ko04972
7	Cell adhesion molecules	24	1.157608e-05	ko04514
8	Circadian rhythm-mammal	3	5.041055e-05	ko04710
9	Wnt signaling pathway	19	0.0001171871	ko04310
10	Peroxisome	24	0.000494032	ko04146
11	mRNA surveillance pathway	84	0.0007267959	ko03015
12	Huntington's disease	72	0.001115805	ko05016
13	Fat digestion and absorption	16	0.001116009	ko04975
14	Lysosome	61	0.001909225	ko04142
15	RNA degradation	27	0.002216912	ko03018
16	Osteoclast differentiation	20	0.002784256	ko04380
17	Other glycan degradation	7	0.00325044	ko00511
18	Basal transcription factors	26	0.003300297	ko03022
19	Alanine and glutamate metabolism	8	0.004215956	ko00250
20	Cytokine-cytokine receptor interaction	12	0.00645224	ko04060
21	Natural killer cell-mediated cytotoxicity	13	0.006618265	ko04650
22	Ribosome	8	0.007251795	ko03010
23	Nicotinate and nicotinamide metabolism	8	0.007983207	ko00760
24	Hepatitis C	15	0.008215165	ko05160
25	Antigen processing and presentation	32	0.01524374	ko04612
26	SNARE interactions in vesicular transport	1	0.01540097	ko04130
27	RIG-I-like receptor signaling pathway	8	0.02159715	ko04622
28	Ubiquitin mediated proteolysis	54	0.02664447	ko04120
29	Phototransduction- fly	22	0.02730456	ko04745
30	<i>Vibrio cholerae</i> infection	113	0.02743882	ko05110
31	beta-Alanine metabolism	4	0.03397021	ko00410
32	Shigellosis	42	0.03674726	ko05131
33	Linoleic acid metabolism	9	0.03837984	ko00591
34	Thiamine metabolism	1	0.04429496	ko00730
35	Tuberculosis	68	0.04874753	ko05152

DEG, differentially expressed gene; ECM, extracellular matrix.

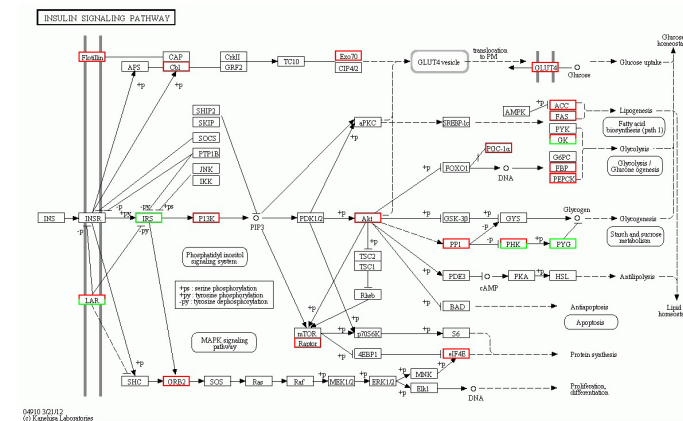


### In-depth analysis of DEGs involved in antiviral innate immunity signaling pathways

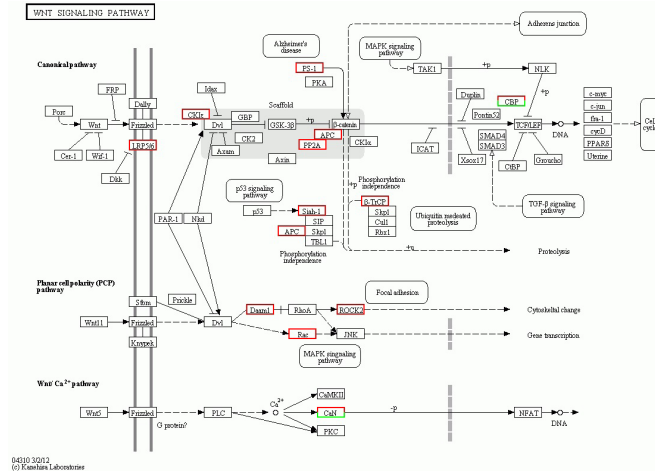
Based on the KEGG pathway analysis of the DEGs between GW and GN, some signaling pathways that were related to the invertebrate innate immune system were screened out, including the JAK-STAT, insulin, and Wnt signaling pathways. All three signaling pathways were significantly different between the groups ( $P < 0.01$ ). In the JAK-STAT signaling pathway, JAK activation stimulates cell proliferation, differentiation, migration, and apoptosis (Harel et al., 2015). Cell proliferation and apoptosis are involved in the antiviral immune response (Du et al., 2013). Eighteen innate immunity-related genes were significantly differentially expressed in the JAK-STAT signaling pathway, 15 of which were significantly upregulated and 3 significantly downregulated (Figure 3). In the insulin signaling pathway, 38 innate immunity-related genes were significantly differentially expressed, including 29 upregulated and 9 downregulated genes (Figure 4). In the Wnt signaling pathway, 19 innate immunity-related genes were significantly differentially expressed, including 15 upregulated and 4 downregulated genes (Figure 5).



**Figure 3.** Significantly differentially expressed genes identified by the Kyoto Encyclopedia of Genes and Genomes database as being involved in the JAK-STAT signaling pathway. Red boxes indicate significantly increased expression, green boxes indicate significantly decreased expression, and blue boxes indicate unchanged expression.



**Figure 4.** Significantly differentially expressed genes identified by the Kyoto Encyclopedia of Genes and Genomes database as being involved in the insulin signaling pathway. Red boxes indicate significantly increased expression, green boxes indicate significantly decreased expression, and blue boxes indicate unchanged expression.



**Figure 5.** Significantly differentially expressed genes identified by the Kyoto Encyclopedia of Genes and Genomes database as being involved in the Wnt signaling pathway. Red boxes indicate significantly increased expression, green boxes indicate significantly decreased expression, and blue boxes indicate unchanged expression.

## Validation of transcriptome data by qRT-PCR

We randomly chose 10 genes that were related to the immune system to evaluate their expression levels in the GW and GN samples using qRT-PCR (Gao et al., 2015). The results of the qRT-PCR were consistent with those obtained from the RNA-Seq data (Table 3) and had similar up/downregulation patterns, suggesting that the RNA-Seq data were reliable.

**Table 3.** Comparison of relative fold changes between normal crayfish lymph organ (GN) and white spot syndrome virus-challenged crayfish lymph organ (GW) based on RNA-Seq and quantitative real-time polymerase chain reaction (PCR) results.

Gene name	Protein identity	Fold variation (GW/GN)	
		Transcriptome	PCR
CL4171.contig3_All	Anti-lipopolysaccharide factor	10.7 (up)	7.5 (up)
CL3734.contig3_All	Serine proteinase inhibitor	6.2 (up)	4.7 (up)
CL450.contig1_All	Single WAP domain-containing protein	3.5 (up)	2.8 (up)
CL6168.contig1_All	Signal transducer and activator of transcription	3.4 (up)	2.9 (up)
CL2411.contig3_All	C-type lectin-like protein	3.3 (up)	4.8 (up)
CL4879.contig3_All	Hsp 70	27.5 (up)	18.4 (up)
CL2770.contig3_All	Caspase	8.6 (up)	5.3 (up)
CL1992.contig2_All	Toll like receptor-2	2.6 (up)	2.2 (up)
CL6040.contig1_All	Prophenoloxidase	3.7 (up)	4.2 (up)
CL3824.contig5_All	Peroxiectin	2.6 (up)	3.5 (up)

## DISCUSSION

Transcriptome sequencing is an important method of obtaining genomic information from non-model organisms when no reference genome is available. Genomic information is important for investigating the molecular mechanisms underlying an organism's biological traits. In the present study, *de novo*-assembled transcriptomes from the crayfish lymph organ were analyzed, and a large amount of sequence information was obtained. There were 5933

DEGs between GN and GW, including 4393 differentially upregulated and 1210 differentially downregulated genes. A total of 245 DEGs were only found in the GW group, and might be related to the crayfish anti-WSSV immune response.

Previous studies on the crayfish based on NGS technology have mainly focused on gonadal development, neuroendocrinology, and genetic markers (Shen et al., 2014), and very little research has been conducted on crayfish transcriptome expression levels after a viral challenge. We analyzed the DEGs for function annotation, clusters of orthologous proteins, and the annotation of signaling pathways that were related to immunity, in order to determine the underlying mechanisms involved in the crayfish anti-WSSV immune response. Based on the KEGG pathway analysis, some signaling pathways related to the invertebrate innate immune system were identified, e.g., the JAK-STAT, insulin, and Wnt signaling pathways.

The JAK-STAT signaling pathway is involved in the shrimp innate immune response (Sun et al., 2011), and has also been implicated in the insect antiviral immune response, which includes three main cellular components: the receptor Domeless, Janus kinase Hopscotch, and the STAT transcription factor (Agaisse and Perrimon, 2004). STAT transcription in the shrimp is significantly upregulated after WSSV infection (Sun et al., 2011), indicating that the JAK-STAT pathway plays an important role in shrimp antiviral immune responses. In the present study, some unigenes were annotated in the JAK-STAT signaling pathway, and their expression levels significantly differed after WSSV infection. This suggests that this pathway plays an important role in the crayfish antiviral innate immune response.

The insulin and Wnt signaling pathways have not been investigated in relation to crustacean innate immunity (Gurskaya et al., 2015). However, some important members of these pathways were predicted in the present study. In the insulin signaling pathway, 38 genes were significantly differentially expressed after WSSV infection, including 29 upregulated and 9 downregulated genes. In the Wnt signaling pathway, 19 innate immunity-related genes were significantly differentially expressed after WSSV infection. Of these, 15 were upregulated and 4 downregulated. All of the DEGs that differed significantly were related to crayfish anti-WSSV innate immunity. In conclusion, several genes and pathways that were related to innate immunity were modified after WSSV infection in the crayfish lymph organ. Our results provide new insights for future research into crayfish antiviral immunity, and could serve as an important theoretical basis for solving viral disease problems in crayfish breeding programs.

### Conflicts of interest

The authors declare no conflicts of interest.

### ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (Grant #31460698).

### REFERENCES

- Agaisse H and Perrimon N (2004). The roles of JAK/STAT signaling in *Drosophila* immune responses. *Immunol. Rev.* 198: 72-82. <http://dx.doi.org/10.1111/j.0105-2896.2004.0133.x>
- Ali MY, Pavasovic A, Mather PB and Prentis PJ (2015). Transcriptome analysis and characterisation of gill-expressed carbonic anhydrase and other key osmoregulatory genes in freshwater crayfish *Cherax quadricarinatus*. *Data Brief* 5: 187-193. <http://dx.doi.org/10.1016/j.dib.2015.08.018>

- Banani H, Marcet-Houben M, Ballester AR, Abbruscato P, et al. (2016). Genome sequencing and secondary metabolism of the postharvest pathogen *Penicillium griseofulvum*. *BMC Genomics* 17: 19. <http://dx.doi.org/10.1186/s12864-015-2347-x>
- Chen K, Li E, Li T, Xu C, et al. (2015a). Transcriptome and molecular pathway analysis of the hepatopancreas in the Pacific white shrimp *Litopenaeus vannamei* under chronic low-salinity stress. *PLoS One* 10: e0131503. <http://dx.doi.org/10.1371/journal.pone.0131503>
- Chen K, Li E, Xu Z, Li T, et al. (2015b). Comparative transcriptome analysis in the hepatopancreas tissue of Pacific white shrimp *Litopenaeus vannamei* fed different lipid sources at low salinity. *PLoS One* 10: e0144889. <http://dx.doi.org/10.1371/journal.pone.0144889>
- Christie AE, Chi M, Lameyer TJ, Pascual MG, et al. (2015). Neuropeptidergic signaling in the American lobster *Homarus americanus*: new insights from high-throughput nucleotide sequencing. *PLoS One* 10: e0145964. <http://dx.doi.org/10.1371/journal.pone.0145964>
- Dedeine F, Weinert LA, Bigot D, Josse T, et al. (2015). Comparative analysis of transcriptomes from secondary reproductives of three *Reticulitermes* termite species. *PLoS One* 10: e0145596. <http://dx.doi.org/10.1371/journal.pone.0145596>
- Di Lena P, Domeniconi G, Margara L and Moro G (2015). GOTA: GO term annotation of biomedical literature. *BMC Bioinformatics* 16: 346. <http://dx.doi.org/10.1186/s12859-015-0777-8>
- Du Z, Jin Y and Ren D (2016). In-depth comparative transcriptome analysis of intestines of red swamp crayfish, *Procambarus clarkii*, infected with WSSV. *Sci. Rep.* 6: 26780 <http://dx.doi.org/10.1038/srep26780>.
- Du ZQ and Jin YH (2015). Molecular characterization and antibacterial activity analysis of two novel penaeidin isoforms from Pacific white shrimp, *Litopenaeus vannamei*. *Appl. Biochem. Biotechnol.* 177: 1607-1620. <http://dx.doi.org/10.1007/s12010-015-1840-7>
- Du ZQ, Li XC, Wang ZH, Zhao XF, et al. (2010). A single WAP domain (SWD)-containing protein with antipathogenic relevance in red swamp crayfish, *Procambarus clarkii*. *Fish Shellfish Immunol.* 28: 134-142. <http://dx.doi.org/10.1016/j.fsi.2009.10.009>
- Du ZQ, Lan JF, Weng YD, Zhao XF, et al. (2013). BAX inhibitor-1 silencing suppresses white spot syndrome virus replication in red swamp crayfish, *Procambarus clarkii*. *Fish Shellfish Immunol.* 35: 46-53. <http://dx.doi.org/10.1016/j.fsi.2013.03.376>
- Ermolaeva MA and Schumacher B (2014). Insights from the worm: the *C. elegans* model for innate immunity. *Semin. Immunol.* 26: 303-309. <http://dx.doi.org/10.1016/j.smim.2014.04.005>
- Feng D, Li Q, Yu H, Zhao X, et al. (2015). Comparative transcriptome analysis of the Pacific oyster *Crassostrea gigas* characterized by shell colors: identification of genetic bases potentially involved in pigmentation. *PLoS One* 10: e0145257. <http://dx.doi.org/10.1371/journal.pone.0145257>
- Gao Y, Zhang X, Wei J, Sun X, et al. (2015). Whole transcriptome analysis provides insights into molecular mechanisms for molting in *Litopenaeus vannamei*. *PLoS One* 10: e0144350. <http://dx.doi.org/10.1371/journal.pone.0144350>
- Gupta P, Goel R, Agarwal AV, Asif MH, et al. (2015). Comparative transcriptome analysis of different chemotypes elucidates withanolide biosynthesis pathway from medicinal plant *Withania somnifera*. *Sci. Rep.* 5: 18611. <http://dx.doi.org/10.1038/srep18611>
- Gurskaya OY, Dobryakova YV and Markevich VA (2015). A role of the Wnt signaling in the regulation of brain function. *Zh. Vyssh. Nerv. Deiat. Im. I P Pavlova* 65: 387-399.
- Harel S, Higgins CA, Cerise JE, Dai Z, et al. (2015). Pharmacologic inhibition of JAK-STAT signaling promotes hair growth. *Sci. Adv.* 1: e1500973. <http://dx.doi.org/10.1126/sciadv.1500973>
- He W, Zhuang H, Fu Y, Guo L, et al. (2015). *De novo* transcriptome assembly of a Chinese locoweed (*Oxytropis ochrocephala*) species provides insights into genes associated with drought, salinity, and cold tolerance. *Front. Plant Sci.* 6: 1086. <http://dx.doi.org/10.3389/fpls.2015.01086>
- Jiang H, Xing Z, Lu W, Qian Z, et al. (2014). Transcriptome analysis of red swamp crawfish *Procambarus clarkii* reveals genes involved in gonadal development. *PLoS One* 9: e105122. <http://dx.doi.org/10.1371/journal.pone.0105122>
- Lai Z and Lin Y (2013). Analysis of the global transcriptome of longan (*Dimocarpus longan* Lour.) embryogenic callus using Illumina paired-end sequencing. *BMC Genomics* 14: 561. <http://dx.doi.org/10.1186/1471-2164-14-561>
- Langmead B and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9: 357-359. <http://dx.doi.org/10.1038/nmeth.1923>
- Leng X, Jia H, Sun X, Shangguan L, et al. (2015). Comparative transcriptome analysis of grapevine in response to copper stress. *Sci. Rep.* 5: 17749. <http://dx.doi.org/10.1038/srep17749>
- Li D, Liang Y, Wang X, Wang L, et al. (2015). Transcriptomic analysis of *Musca domestica* to reveal key genes of the prophenoloxidase-activating system. *G3 (Bethesda)* 5: 1827-1841. <http://dx.doi.org/10.1534/g3.115.016899>

- Li S, Zhang X, Sun Z, Li F, et al. (2013). Transcriptome analysis on Chinese shrimp *Fenneropenaeus chinensis* during WSSV acute infection. *PLoS One* 8: e58627. <http://dx.doi.org/10.1371/journal.pone.0058627>
- Li ST, Zhang P, Zhang M, Fu CH, et al. (2012a). Transcriptional profile of *Taxus chinensis* cells in response to methyl jasmonate. *BMC Genomics* 13: 295-305. <http://dx.doi.org/10.1186/1471-2164-13-295>
- Li Y, Deng W, Yang K and Wang W (2012b). The expression of prophenoloxidase mRNA in red swamp crayfish, *Procambarus clarkii*, when it was challenged. *Genomics* 99: 355-360. <http://dx.doi.org/10.1016/j.ygeno.2012.04.001>
- Livak KJ and Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-DDCt</sup> method. *Methods* 25: 402-408. <http://dx.doi.org/10.1006/meth.2001.1262>
- Manfrin C, Tom M, De Moro G, Gerdol M, et al. (2015). The eyestalk transcriptome of red swamp crayfish *Procambarus clarkii*. *Gene* 557: 28-34. <http://dx.doi.org/10.1016/j.gene.2014.12.001>
- Mao X, Cai T, Olyarchuk JG and Wei L (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21: 3787-3793. <http://dx.doi.org/10.1093/bioinformatics/bti430>
- Martin JA and Wang Z (2011). Next-generation transcriptome assembly. *Nat. Rev. Genet.* 12: 671-682. <http://dx.doi.org/10.1038/nrg3068>
- Mortazavi A, Williams BA, McCue K, Schaeffer L, et al. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5: 621-628. <http://dx.doi.org/10.1038/nmeth.1226>
- Mousavi S, Alisoltani A, Shiran B, Fallahi H, et al. (2014). *De novo* transcriptome assembly and comparative analysis of differentially expressed genes in *Prunus dulcis* Mill. in response to freezing stress. *PLoS One* 9: e104541. <http://dx.doi.org/10.1371/journal.pone.0104541>
- Shen H, Hu Y, Ma Y, Zhou X, et al. (2014). In-depth transcriptome analysis of the red swamp crayfish *Procambarus clarkii*. *PLoS One* 9: e110548. <http://dx.doi.org/10.1371/journal.pone.0110548>
- Sun C, Shao HL, Zhang XW, Zhao XF, et al. (2011). Molecular cloning and expression analysis of signal transducer and activator of transcription (STAT) from the Chinese white shrimp *Fenneropenaeus chinensis*. *Mol. Biol. Rep.* 38: 5313-5319. <http://dx.doi.org/10.1007/s11033-011-0681-x>
- Taffoni C and Pujol N (2015). Mechanisms of innate immunity in *C. elegans* epidermis. *Tissue Barriers* 3: e1078432. <http://dx.doi.org/10.1080/21688370.2015.1078432>
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, et al. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4: 41. <http://dx.doi.org/10.1186/1471-2105-4-41>
- Wang P, Wang J, Su YQ, Mao Y, et al. (2016). Transcriptome analysis of the *Larimichthys crocea* liver in response to *Cryptocaryon irritans*. *Fish Shellfish Immunol.* 48: 1-11. <http://dx.doi.org/10.1016/j.fsi.2015.11.011>
- Young MD, Wakefield MJ, Smyth GK and Oshlack A (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11: R14. <http://dx.doi.org/10.1186/gb-2010-11-2-r14>
- Zeng DG, Chen XL, Xie DX, Zhao YZ, et al. (2015). Identification of highly expressed host microRNAs that respond to white spot syndrome virus infection in the Pacific white shrimp *Litopenaeus vannamei* (Penaeidae). *Genet. Mol. Res.* 14: 4818-4828. <http://dx.doi.org/10.4238/2015.May.11.14>