BMC Biotechnology

Open Access

CrossMark

De novo characterization of venom apparatus transcriptome of *Pardosa pseudoannulata* and analysis of its gene expression in response to Bt protein

Rong Li[1,2], Zhenzhen Yan[1], Juan Wang[1], Qisheng Song[3] and Zhi Wang[1*]

Abstract

Background: *Pardosa pseudoannulata* is a prevailing spider species, and has been regarded as an important bio-control agent of insect pests in farmland of China. However, the available genomic and transcriptomic databases of *P. pseudoannulata* and their venom are limited, which severely hampers functional genomic analysis of *P. pseudoannulata*. Recently high-throughput sequencing technology has been proved to be an efficient tool for profiling the transcriptome of relevant non-target organisms exposed to *Bacillus thuringiensis* (Bt) protein through food webs.

Results: In this study, the transcriptome of the venom apparatus was analyzed. A total of 113,358 non-redundant unigenes were yielded, among which 34,041 unigenes with complete or various length encoding regions were assigned biological function annotations and annotated with gene ontology and karyotic orthologous group terms. In addition, 3726 unigenes involved in response to stimulus and 720 unigenes associated with immune-response pathways were identified. Furthermore, we investigated transcriptomic changes in the venom apparatus using tag-based DGE technique. A total of 1724 differentially expressed genes (DEGs) were detected, while 75 and 372 DEGs were functionally annotated with KEGG pathways and GO terms, respectively. qPCR analyses were performed to verify the DEGs directly or indirectly related to immune and stress responses, including genes encoding heat shock protein, toll-like receptor, GST and NADH dehydrogenase.

Conclusion: This is the first study conducted to specifically investigate the venom apparatus of *P. pseudoannulata* in response to Bt protein exposure through tritrophic chain. A substantial fraction of transcript sequences was generated by high-throughput sequencing of the venom apparatus of *P. pseudoannulata*. Then a comparative transcriptome analysis showing a large number of candidate genes involved in immune response were identified by the tag-based DGE technology. This transcriptome dataset will provide a comprehensive sequence resource for furture molecular genetic research of the venom apparatus of *P. pseudoannulata*.

Keywords: The venom apparatus, Transcriptome, *Pardosa pseudoannulata*, Bt protein, Immune system

* Correspondence: wangzhispider@aliyun.com
[1]College of Bioscience & Biotechnology, Hunan Agriculture University, Changsha 410128, China
Full list of author information is available at the end of the article

Li *et al. BMC Biotechnology* (2017) 17:73

Page 2 of 11

Background

Animals use the venom apparatus for defense and predation by injecting their venomous mixture into their preys, and venom also have important biological implications with biomedical and agricultural relevance [1]. Only certain animals synthesize venom, but their production are taxonomically widespread in invertebrates and vertebrates because they have evolved independently on numerous occasions [2]. Venom genes have already been previously reported in snakes, scorpions, stingrays and parasitic wasps [3–6]. Spider venoms recently gathered far more attention from worldwide research groups [7]. Because spider venoms are complex mixture, mainly containing disulfide peptides that typically have high affinity and specificity for particular subtypes of ion channels and receptors, they can be used as leads for developing therapeutics and bioinsecticides [8, 9].

However, our understanding of spiders and their venom is limited, due in large part to the lack of spider genome database and comprehensive transcriptome dataset [6]. With the exceedingly rapid development of next-generation sequencing technology (NGS) and drop in cost, the number of sequencing applications are increasing exponentially [10].

It is widely known that *Bacillus thuringiensis* (*Bt*) strains can produce Cry proteins that posses insecticidal activity and induce the formation of pores in midgut membranes of susceptible organisms. Therefore, Bt proteins have been employed as bioinsecticides against target organisms, such as lepidopteran pest larvae [11]. However, the possibility of adverse impacts on the non-target species remains a concern. In fact, many researchers are specialized in elucidating the mechanism of the effect of Bt proteins on non-target organisms via tritrophic food chain [12, 13]. Non-target organisms involved in these recent researches can be generally classified into two categories: model and non-model organisms. For example, Tang et al. conducted a research that when Sprague Dawley rats were fed with the rice line T1C-1 expressing Cry1C for 90 days, this Bt rice line had no influence on rat behavior, weight gain and blood biochemical parameters. At the same time, no histopathological changes were recorded [14]. In a related investigation of other model organism, zebrafish *Danio rerio,* it was found that Bt maize is just as safe and nutritious as non-Bt maize for two generations of zebrafish when they were fed with casein/gelatin-based diets containing 19% Bt maize [15]. With regard to the non-model organisms, on the one hand, studies usually centered on the effect of Bt protein on its ecological behavior, for example, food consumption of *Apis mellifera* [16] or growth and predation of *Hylyphantes graminicola* (Araneae: Linyphiidae) and *Coleosoma octomaculatum* (Araneae: Theridiidae) [17] or survival and fecundity of *Daphnia magna* exposed to the Bt maize [18]. On the other hand, some searchers have obtained massive amounts of sequence data for target organisms using NGS, but these profiles did not cover the effect of Bt protein on non-target organisms and their apparatus. For instance, transcriptomes of three orb-web spiders (*Gasteracantha arcuata, Nasoonaria sinensis* and *Gasteracantha hasselti*) were analyzed for revealing gene expression profiles associated with toxin and silk [19], but transcriptomic analysis of venom glands was performed only for a single fishing spider *Dolomedes mizhoanus* [20]. A series of studies of Bt effect on *P. pseudoannulata* have been performed in our lab. One of our studies showed that developmental time of the spiders from the 2nd to 8th instars was affected by Bt rice. Another study also showed that the activities of three key metabolic enzymes, acetylcholine esterase (AchE), glutathione peroxidase (GSH-Px), and superoxide dismutase (SOD) were significantly influenced in the spiders after feeding on Cry1Ab-containing fruit flies [21, 22]. Nevertheless, the Bt effect on the gene expression in venom glands of spiders has not been reported. In the study, we wanted to explore whether Bt protein has any impact on the venom apparatus of *P. pseudoannulata* through food webs using transcriptome analysis.

Up to now, only *Tetranychus urticae* (the two-spotted spider mite) genome and *Ixodes scapularis* (blacklegged tick) genome were sequenced in the class Arachnida, phylum Arthropoda [23]. Many transcriptomes of arachnida were analyzed using Illumina sequencing, including four suborders of harvestmen [24], *Tetranychus cinnabarinus* [25], *Theridion grallator, T. californicum* [26] and so on. In recent years, more studies were focused on the transcriptomes of apparatus, for example, venom glands of *Hadrurus gertschi* [27], *Latrodectus tredecimguttatus* [28], spinning glands of *Actinopus spp.* and *Gasteracantha cancriformis* [29], salivary glands of *Ixodes ricinus* [30], and ovaris of *Rhipicephalus (Boophilus) microplus* [31]. Comparing with conventional sequencing methods, the NGS provides a more ideal method for the analysis of transcriptome with high efficiency, and facilitates the studies on arachnidas' gene background.

In this paper, a global genetic and transcriptome-based analysis of the venom apparatus of *P. pseudoannulata* [30, 32] exposed to the Bt protein by tritrophic pathway was examined using high throughput next-generation RNA-sequencing technique coupled with bioinformatics. In addition, NGS technology was applied to analyze the differentially expressed genes (DEGs) in the venom apparatus of *P. pseudoannulata* following a tritrophic model: the wolf spiders preyed on brown planthoppers (BPHs) *Nilaparvata lugens*, which were reared on the Cry1Ab-expressing rice or non-transgenic parental commercial cultivar. DEGs were further validated by qPCR analyses to gain insight into the response of venom apparatus to Bt protein exposure and to identify the genes that may respond to environmental stress or immunity andcould serve as targets for improved bioinsecticides.

Li *et al. BMC Biotechnology* (2017) 17:73

Page 3 of 11

Methods

Insect rearing and sample preparation

BPHs were gathered from Bt and non-Bt control rice after feeding on rice stems for 15 days, and used as foods for spiders. Approximately 70,000 BPHs were collected from each group [33].

Forty healthy *P. pseudoannulata* adults of similar size were collected from traditional fields in the experimental farmland of Hunan Academy of Agriculture Sciences (113.08°E, 28.18°N) in 2014. Each spider was transferred to a glass tube (15 mm diameter, 100 mm height) with moisten cotton ball covered at the bottom for water. The spiders were starved for 48 h and then fed with 15 Cry1Ab-containing BPHs (the average amount of Cry1Ab protein in BHPs: 0.47 ng/mg) or 15 non-Cry1Ab-containing BPHs every day. The feeding process was conducted in a temperature-controlled chamber (25 ± 2 °C, 70 ± 10% RH and L:D 16:8 photoperiod, photophase from 06:00 to 22:00).

RNA isolation

This work was done by Oebiotech Enterprise, Shanghai. The venom apparatus (containing main gland, venom canal and fang) from a total of 19 control and 21 Bt-treated spiders, respectively, were carefully separated from other tissues, rinsed thoroughly in phosphate-buffer saline (PBS) solution (pH 7.2) to remove debris, and pooled. Total RNA of each sample was extracted immediately upon dissection using TRIzol reagent (Invitrogen, USA) and used to generate cDNA library. The procedures were performed according to the manufacturer's protocol. The quality and quantity of RNA were analyzed with electrophoresis in 1% TAE agarose gel and with a NanoDrop spectrophotometer (ThermoScientific, USA), respectively. The total RNAs with absorbance 260/280 nm ratio of 1.8~2.1 were chosen and treated with DNase I (Fermentas, Lithuania) prior to library construction. RNA samples were stored frozen at −80 °C until needed.

cDNA library construction

Poly-(A) mRNA was purified using oligo (dT) magnetic beads and used for the first-strand cDNA synthesis with random hexamer-primer and reverse transcriptase (Invitrogen, USA). The second-strand cDNA was synthesized using dNTPs, buffer, RNaseH (Invitrogen, USA) and DNA polymerase I (TaKaRa, Japan). The double-strand cDNA was end-repaired using T4 DNA polymerase, Klenow fragment, and T4 polynucleotide kinase (Thermo Scientific, USA) followed by a single (A) base addition using Klenow 3′ to 5′ exo-ploymerase, and further ligated with PE Adapter Oligo Mix supplied by the mRNA-Seq Sample Preparation Kit (Illumina, USA) using T4 DNA ligase (Thermo Scientific, USA). The resulting cDNA with adaptor-ligated was then subjected to PCR amplification for about 15 cycles.

Bioinformatics analysis of RNA-seq results

The cDNA library was sequenced on the Illmina Hiseq2000 platform. Sequencing-received raw image data were transformed by base calling into sequence data, which were called raw reads [34]. The raw reads as the direct product from Illmina sequencing were transformed into clean reads by removing low quality reads, adaptor sequence, fragments with bases N and duplication sequence. Then clean reads were de novo assembled with the short reads assembling program Trinity, which is a very favorable tool, especially in the absence of a reference genome [35]. The resulting unigenes with a minimal length of 200 bp were used for sequence similarity searches in the Nr/NCBI (non-redundant, http://www.ncbi.nlm.nih.gov), Swiss-Prot/UniProtKB (http://www.uniprot.org), KOG/NCBI (http://www.ncbi.nlm.nih.gov), KEGG (http://www.genome.jp/kegg/) and GO (http://geneontology.org/) with a siginificant threshold of E-value $< 10^{-5}$.

Analysis of differentially expressed genes (DEGs)

Gene expression of all unigenes in the venom apparatus of spiders fed with Bt BPHs and non-Bt BPHs were estimated by calculating read density as 'reads per kilobase of exon model per million mapped reads' (FPKM) [36]. In the present study, a Perl script was used to calculate FPKM for each assembled cDNA sequence with parsing alignment data using the software Bowtie 2 [37]. And we used the false discovery rate (FDR) method mentioned by Audic and Claverie to determine the *P*-value threshold for multiple testing by controlling the FDR value [38]. A threshold value of FDR <0.05 and absolute value of log2 fold change > 2 were used to identify DEGs.

Validation of gene expression by qPCR

qPCR analysis was conducted to validate key venom apparatus genes of interest, which was identified either by their exhibition of relatively large expression fold-changes from DEG analyses, or by scientific literature about potential gene functions related to immunity. Total RNA of each sample was isolated with TRIzol (Invitrogen Corp, USA) and treated with DNase I (Fermentas, Lithuania) according to the manufacturer's protocol. The cDNA was synthesized using a RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas, Lithuania) and the qRT-PCR was performed using ABI 7900HT (ABI, USA) with a volume of 25 μL, containing 1 μL of 1:10 cDNA diluted with ddH$_2$O, 12.5 μL 2xSYBR Green Master Mix (ABI, USA) and 200 nM primer pair. Reactions were performed in triplicate to ensure consistent technical replication and run in 96-well plates under the following conditions: 94 °C for 3 min to activate the polymerase, then followed by 40 cycles for 94 °C for 30 s, 59 °C for 30 s, 72 °C for 45 s. Cycling ended at 72 °

Li *et al. BMC Biotechnology* (2017) 17:73

Page 4 of 11

C for 5 min. Relative gene expression was evaluated with a DataAssist Software version 3.0 (Applied Biosystems/ Life Technologies), using 18 s rRNA (primers: 5′- AGATGCCCTTAGATGTCCGG-3′ and 5′-AAGGG- CAGGGACGTAATCAA-3′) as an endogenous control for RNA load and gene expression in analyses. The relative quantitative method ($\triangle\triangle C_T$) was used to calculate the fold change of target genes [39].

Results

Illumina paired-end sequencing and reads assembly

Sequences of mRNAs pooled from the whole body of spiderlings were analyzed by Illumina Hiseq 2000 (Illumina, USA) and generated 63 million raw reads (consisting of 6,369,225,236 bp). After removing low quality reads and adaptors, reliable reads were de novo assembled with Trinity software. A total of 159,816 transcripts were assembled ranging from 200 to 17,468 bp, with an average length size of 605 bp. Among these transcripts, 105,564 (66.05%) were smaller than 500 bp, 31,228 (19.54%) were between 500 bp and 1000 bp, 23,024 (14.41%) were longer than 1000 bp. These transcripts were further assembled and clustered. The longest sequences in each cluster were retained and designated as unigenes. Taken together, a total of 113,358 unigenes were assembled with a mean length size of 509 bp. The length distribution of assembled unigenes was shown in Fig. 1. The size distribution indicated that there were 102,465 unigenes (90.3% of the total) with the length of less than1000 bp.

Annotation of all assembled unigenes

The annotated unigenes were subjected to BLAST searching against Nr, Swiss-Prot, GO, KOG and KEGG databases with an E-value threshold of e < $1e^{-5}$. Of these, 33,509 (29.6% of all distinct sequences) unigenes could be matched Nr database (Additional file 1: Table S1), 29,042

Fig. 1 Venom apparatus transcriptome. The x-axis indicates the unigene size and Y-axis indicates the number of unigenes of each size

(25.6%) to Swiss-Prot database (Additional file 2: Table S2), 28,600 (25.2%) to GO, 26,684 (23.5%) to KOG, 5521(4.9%) to KEGG. The species distribution of the top BLAST hits in Nr database for each unique sequence is shown in Fig. 2 and the summary statistics of BLAST assignment were shown in Table 1.

Gene ontology (GO) is dynamic ontology-based resource that provides computationally tractable and excellent broad coverage of available information about molecular biology [40]. It covers several major domains of the overall GO, which are represented by the "cellular component", "biological process" and "molecular function". Based on the annotation in Nr database, the annotated unigenes against GO database were collected by Blast2GO (Additional file 3: Table S3). We used WEGO software to perform GO functional classification for all the GO annotated unigenes to comprehend the distribution of gene functions of this species at the macro level (Fig. 3). Total 9973 GO term annotations corresponding to 28,600 unigenes were produced and assigned into 52 functional groups consisting of three domains: biological process (83,861), molecular function (47,010) and cellular component (27,496) (Additional file 4: Table S4).
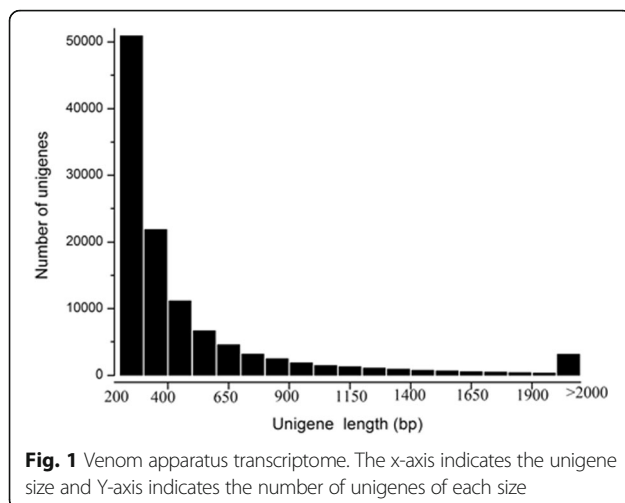
To evaluate the completeness of the transcriptome library and the efficiency of annotation process, all unigenes were aligned to eukaryotic orthologous groups (KOG) database to predict and classify possible functions. From a total of 113,358 Nr hits, 26,684 unigenes (32.36% of the total) were annotated and formed 25 KOG classifications (Additional file 5: Table S5 & Fig. 4).
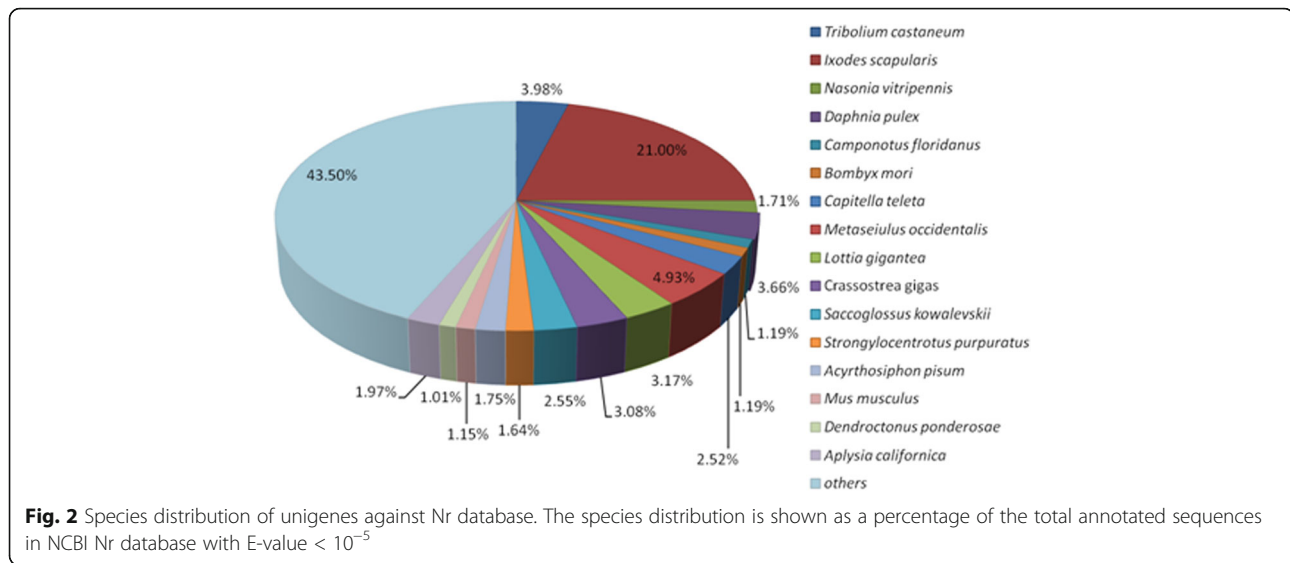
To identify the metabolic pathways populated by these unigenes, total 5221 annotated unigenes were grouped into six categories and mapped to 326 known metabolic or signaling pathways (Fig. 5). Among these pathways, the most representative pathways were signal transduction (1692 unigenes), infectious disease (1510 unigenes), cancers (1340 unigenes) and endocrine system (905 unigenes).

Functional genes in response to environmental stress and immunity

In this study, the sequence and annotation information from GO and KEGG provided valuable gene resources for research on the molecular biology of these hallmarks in venom apparatus of *P. pseudoannulata*.

According to KEGG analysis, 720 unigenes were involved in 15 immune-response pathways (Fig. 5), such as 'Chemokine signaling pathway', 'Platelet activation', 'Fc gamma R-mediated phagocytosis', 'Toll-like receptor signaling pathway', 'Complement and coagulation cascades' and 'NOD-like receptor signaling pathway'. A total of 1782 unigenes belong to 24 signal transduction pathways (Fig. 5). Besides the KEGG analyses, the GO annotation identified 1009 and 3726 unigenes in respond to immune system process (GO: 2,220,682) and stimulus

Li *et al. BMC Biotechnology* (2017) 17:73

Page 5 of 11

Fig. 2 Species distribution of unigenes against Nr database. The species distribution is shown as a percentage of the total annotated sequences in NCBI Nr database with E-value < 10^{-5}

(GO: 0050896), respectively (Additional file 6: Table S6). Heat shock protein (HSP) family performs crucial functions in correct protein folding and unfolding, translocation of proteins as well as assembly and disassembly of protein complexes. They protect hosts from stress caused by infection, inflammation or similar events [41]. As for GO annotation, the HSP families, including HSP70 and HSP90, were detected.

Changes in gene expression in the venom apparatus

Venom, which has special function for prey capture, defense and competition deterrence, consists of a complex mixture of proteins, peptides and some small molecules. To date, little is known about how venom gland gene expression varies among species because of paucity of spider genome data. Some researchers analyzed the venom gland genome of spiders, and found the evolutionary shifts of genes among species [42, 43].

To calculate the expression abundance of each gene between control and Bt-treated group, the MARS model in the DEGseq package (http://www.bioconductor.org/packages/2.6/bioc/html/DEGseq.html) was used [44]. A total of 1724 significantly changed gene entries were observed

including 1654 up-regulated and 70 down-regulated genes. To identify more realistic DEGs, negative binomial distribution was introduced [45] for testifying the significant difference of reads and the value of 'basemean' which was used to estimate the expression abundance (Additional file 7: Figure S1). A total of 152 DEGs were observed between the two groups corresponding to 137 up-regulated and 15 down-regulated genes (Additional file 8: Table S7).

Function annotation of DEGs

Different genes usually cooperate with each other to execute their biological functions. To comprehend the function of DEGs, we mapped all the genes with *P*-value < 0.01 to terms in the KEGG database for identifying genes involved in metabolic or signaling pathways and compared this with the whole transcriptome background. Among all the genes with KEGG annotation, 76 KEGG pathways were identified corresponding to 75 annotated genes (Additional file 9: Table S8). The main enriched metabolism pathways centered on the oxidative phosphorylation. Genetic information processing including ribosome was also enriched.

In addition, all the DEGs were also mapped to the GO terms to thoroughly describe the properties of genes and their production (Additional file 10: Table S9). A total pg 334 GO term annotations corresponding to 372 DEGs were produced and assigned into 47 functional groups and three categories. Among which 951 were assigned in biological process category, 488 in molecular function category and 826 in cellular component category (Additional file 11: Figure S2). In the category of biological process, 6 DEGs were directly associated with 'immune system process' (GO: 0002376) and 38 DEGs were mapped to 'response to stimulus' (GO: 0050896). There were annotations in several other biological processes that may indirectly participate in immune

Table 1 Annotation of unigenes in venom apparatus

Database	Number of annotated unigenes	Percentage of annotated unigenes
Nr	33,509	29.6%
Swissprot	29,042	25.6%
KEGG	5521	4.9%
KOG	26,684	23.5%
GO	28,600	25.2%

All 113,358 unigenes were annotated against Nr (non-redundant), Swiss-prot, KEGG, KOG and GO databases

Li et al. BMC Biotechnology (2017) 17:73
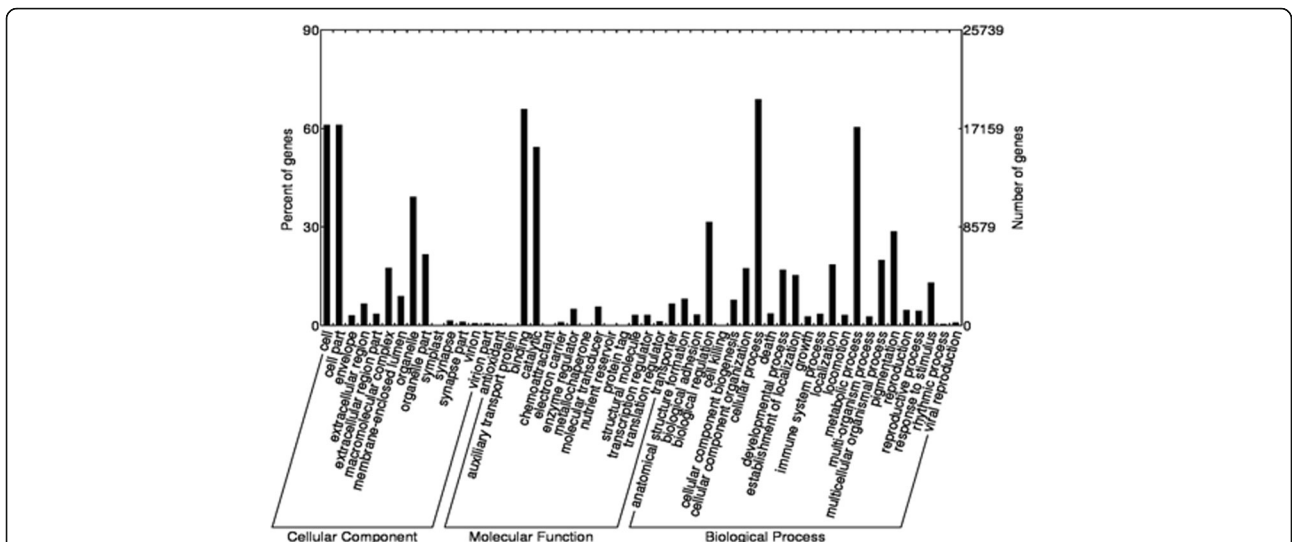
Page 6 of 11

Fig. 3 GO categories of the unigenes. 28,600 unigenes were assigned 544 GO terms, which were divided into three categories: cellular component, molecular function and biological process. The left and right Y-axis denote separately the percent and number of genes in the category

response, such as the cell cycle; DNA replication, transcription, translation; metabolism of carbohydrates, amino acid; ATPase family members, cytochrome oxidase and NADH dehydrogenase. As we discussed previously, HSP families are crucial factors in response to the environmental stress. The present study found that HSP 90 was significantly up-regulated when stressed. Considering the enriched DEGs in the venom apparatus, several important pathways and genes were explored. ATPase, cytochrome oxidase and NADH dehydrogenase were over transcribed in DEGs, indicating that a compensating of partial uncoupling oxidative phosphorylation and key role in energy metabolism.

Validation of DEGs using qPCR

To further validate RNA-seq results, three DEGs in the venom apparatus were selected randomly for analysis by qPCR. Three DEGs were detected differentially expressed, implying that the genes expression in the venom gland of spiders is indeed affected by Bt protein.

Unigenes comp49659_c0_ seq1, involved in defense response to bacterium, was up-regulated with 1.32-fold. Unigene comp59183_c0_seq1, participated in hormone-mediated signaling pathway, was up-regulated by 1.27-fold. Unigenes comp60183_c0_seq1, associated with organism presynaptic membrane cystine knot toxin, was up-regulated with 1.37-fold (Fig. 6). In conclusion, the validity of the sequencing data was confirmed.
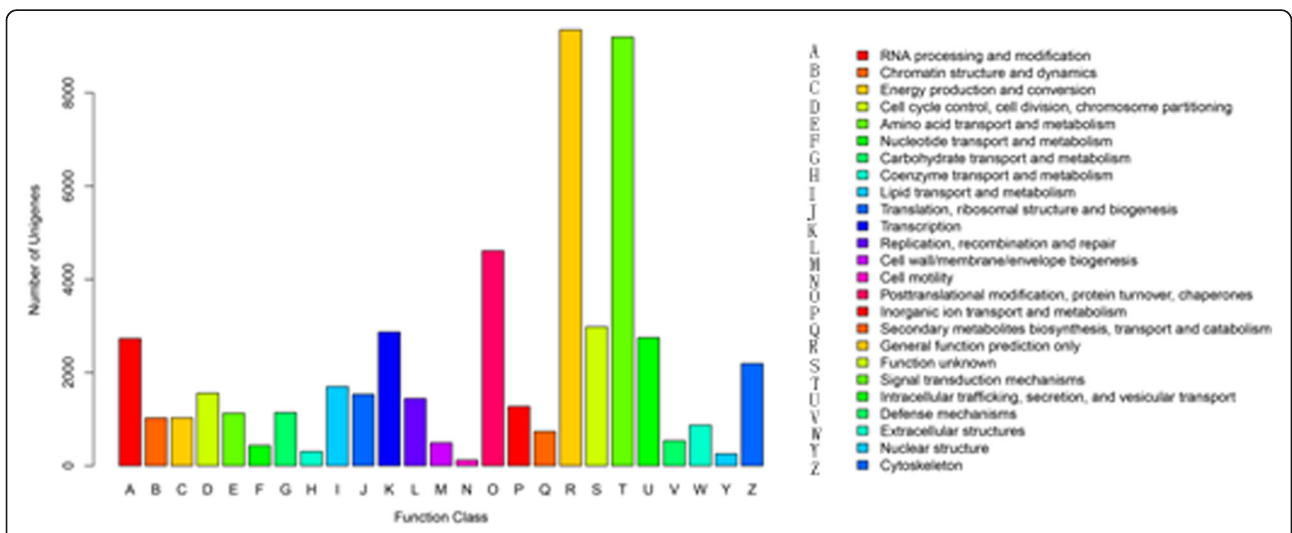

Fig. 4 KOG classification of the unigenes. All the unigenes were aligned to the KOG database and can be classified functionally into 25 categories

Li *et al. BMC Biotechnology* (2017) 17:73
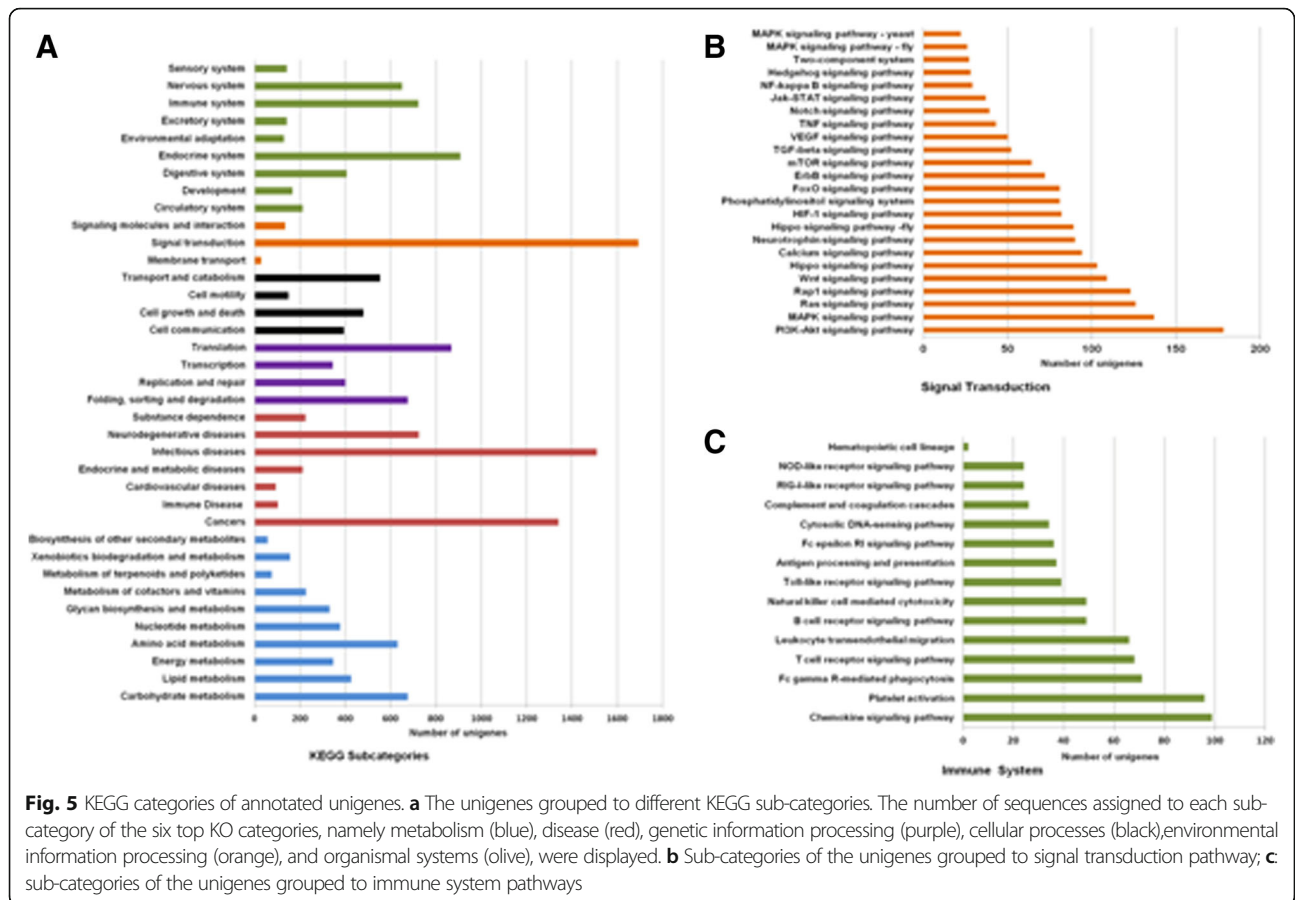
Page 7 of 11

Fig. 5 KEGG categories of annotated unigenes. **a** The unigenes grouped to different KEGG sub-categories. The number of sequences assigned to each sub-category of the six top KO categories, namely metabolism (blue), disease (red), genetic information processing (purple), cellular processes (black), environmental information processing (orange), and organismal systems (olive), were displayed. **b** Sub-categories of the unigenes grouped to signal transduction pathway; **c**: sub-categories of the unigenes grouped to immune system pathways

Discussion

Our previous study showed that Cry1Ab toxin expressed by transgenic rice can be trasferred and accumulated in spiders via the food web [33]. In the present study, next generation sequencing was carried out and produced 113,358 unigenes. To annotate these unigenes, distinct sequences

Fig. 6 Validation of differentially expressed genes by q-pCR. Bars in each panel represent the mean ± standard error ($n = 3$), * $P < 0.05$

were firstly searched against five public databases, Swiss prot, Nr, COG, GO, and KEGG with an E-value threshold of $e < 1e-5$. In total, 33,509 (29.6% of all distinct sequences) unigenes matched to known genes in Nr database; 29,042 (25.6% of all distinct sequences) unigenes were annotated in Swiss-Prot database. Additionally, the unigenes matched 14 of total 16 sequences from *P. pseudoannulata* and 244 of total 558 sequences from *Lycosa singoriensis* (family: Lycosidae) deposited in NCBI Nr database. The remaining unmatched unigenes (>200 bp) were further analyzed using ESTscan software and the predicted CDS were translated into peptide sequences [46].

Surprisingly, BLASTx annotation of the venom apparatus transcriptome sequences of *P. pseudoannulata* revealed the highest similarity with *I. scapularis* of the order Ixodida, and a lower identity match (4.93%) with *M. occidentalis*, a representative of the order Trombidiformes. The two orders belong to the class Arachnida, which also encompasses *P. pseudoannulata.* To expound this similarity, we compared the number of protein sequences from *I. scapularis* (5832) and *M. occidentalis* (21) (chromosome in mitochondria, 11,717 of unplaced). Clearly, the available protein sequences for *M. occidentalis* are fewer than that of *I. scapularis*, due to unavailable
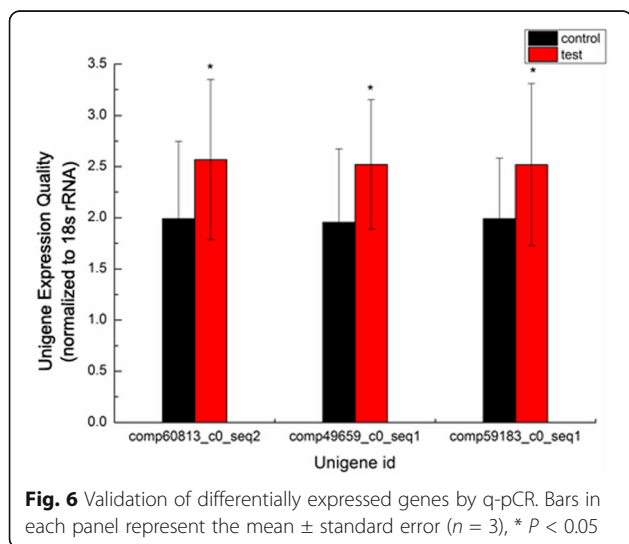
Li *et al. BMC Biotechnology* (2017) 17:73

Page 8 of 11

genome sequences for *M. occidentalis* while the genome of *I. scapularis* has already been sequenced. However, the relationship between species calls for further investigation.

Possible functions of the annotated unigenes were analyzed by matching to GO, KOG and KEGG databases. Although only a small part of unigenes were functionally annotated, the results of these searching help us to learn more properties about the venom apparatus of *P. pseudoannulata*. Interestingly, no cancer to date has ever been reported in the arthropod animal, but our results demonstrated that 1340 unigenes were matched to the 'cancer' KEGG pathway. Cancer is a disease of aberrant multicellularity, which is found in a multitude of living animals (from molluscs to mammals) [47, 48]. It has been reported that the core 'multicellularity' genes of *Amphimedon queenslandica* were implicated in cancer, indicating the ancient origins of cancer [49]. Although some unigenes of the venom apparatus of *P. pseudoannulata* were matched to the 'cancer' KEGG pathway, yet whether it is related to evolution of the multicellularity requires further investigation.

Invertebrate animals can be found in almost every habitat in the world. Because many of them live in environments where microorganisms thrive and proliferate, their widespread distribution and survival are largely due to the successful responses to the environmental stress and changes in immunity [50]. In addition, due to the lack of immunoglobulins, invertebrates only possess innate immunity, which is considered to be an ancient defense mechanism [51, 52]. However, the exact molecular and cellular basis of immune system remains poorly understood. In this study, the sequence and annotation information from GO and KEGG provided valuable gene resources for research on the molecular biology of these hallmarks in the venom apparatus of *P. pseudoannulata.*

Although significant progress has recently been made in studying the immune mechanism at the gene level in insects, for instance, *Apis mellifera, Bombyx mori, Drosophila melanogaster* and *Tribolium castaneum* [53–56], yet the information of immune related genes in spiders is still scarce and fragmentary. Margaret et al. isolated a 34-residue orally active insecticidal peptide (OAIP-1) from the venom of the *Australian tarantula*, which is likely to be synergized by the gut-lytic activity of Bt expressed in insect-resistant transgenic crops [57]. KEGG enrichment analyses indicated that 720 unigenes were involved in 15 immune-response pathways, including 'Toll-like receptor signaling pathway', 'Complement and coagulation cascades' and 'NOD-like receptor signaling pathway'. The complement and coagulation systems have a fundamentally clinical role in injury and inflammation [58]. C3 (complement component 3) and C1q are the central components in complement system and were detected in this annotated pathway. Studies in

the horseshoe crab have provided a breakthrough in our understanding of the coagulation pathway in arthropods [59]. It also has been reported that *I. scapularis* genome encodes for at least 11 genes that may be part of the coagulation pathway [60]. Therefore, this defense mechanism has been strongly present in arthropods. Platelets are the major regulators of coagulation and have been ascribed to an emerging role in the immune-response [61]. It has been reported that enzymes encoded by genes from the venom structures of the caterpillar *Lonomia obliqua* induced the platelet aggregation [62]. Toll like receptor signaling pathway plays a critical role in innate immunity and is highly conserved in structure and function from insects to mammals [63]. TRAF3 was detected in the database, a highly versatile regulator mediating certain innate immune receptor and cytokine receptor signals [64].

Accordingly, there were 1782 unigenes involved in 24 signal transduction pathways (Fig. 5). Among these pathways, 'Jak-STAT signaling pathway', 'NF-kappa B signaling pathway' and 'TNF signaling pathway' were identified and regarded as the main pathways regulating the immune response in insects [45]. SOCS (suppressor of cytokine signaling), a negative regulator of the JAK/STST pathway, was detected and exerted significant effect on regulating the immune response in *Drosophila* [65]. Nuclear factor kappa B (NF-κB) transcription factors are critical to the control of response to cellular stress and are also involved in the regulation of cell-cycle/growth, survival, apoptosis, inflammation and immunity [66]. In addition, the GO annotation analyses indicated that GO terms associated with immune system process and stimulus. Among them, Heat shock protein (HSP) family performs crucial functions in correct protein folding and unfolding, translocation of proteins as well as assembly and disassembly of protein complexes. They protect hosts from stress caused by infection, inflammation or similar events [41].

Consequently, the involvement of these genes in metabolic and signaling pathways provides the basis for further identification of the biological functions of candidate genes in response to environmental stress and immunity.

Comparative transcriptome analysis indicated that some DEGs may indirectly participate in immune response, such as the cell cycle; DNA replication, transcription, translation; metabolism of carbohydrates, amino acid; ATPase family members, cytochrome oxidase and NADH dehydrogenase. As we discussed previously, HSP families are crucial factors in response to the environmental stress. The present study found that HSP 90 was significantly up-regulated when stressed. Considering the enriched DEGs in the venom apparatus, several important pathways and genes were explored. ATPase, cytochrome oxidase and NADH dehydrogenase were over transcribed in DEGs,

Li et al. BMC Biotechnology (2017) 17:73

Page 9 of 11

indicating that a compensating of partial uncoupling oxidative phosphorylation and key role in energy metabolism. Besides these direct or indirect mechanisms associated with immune and stress responses, detoxification is important adaptation that allow insects to overcome the chemical defenses of plants and animals they feed on [67]. Genes related to detoxification have been identified in many species, including *A. gambiae*, *A. aegypti*, *D. melanogaster* and *A. mellifera* [68–71]. Insects considerably rely on three families of enzymes to disarm toxic xenobiotics, such as esterases (ESTs), cytochrome P450 monooxygenases (P450s) and glutathione-S-transferases (GSTs) [72]. In light of this, our study found that genes encoding GST were detected over-expressed in DEGs. In this study, we focused on the functional genes in the venom apparatus of *P. pseudoannulata* in response to environmental stress and immunity. The launching of transcriptomics study of *P. pseudoannulata*, as a non-model animal, will lay the foundation for future functional genomic research.

Conclusions

In this study, we used high-throughput sequencing technology to identify the transcriptome of the venom apparatus of *P. pseudoannulata*, a species for which little genome data is available. The research showed that a number of candidate genes involved in immune response were identified. This transcriptome dataset will provide a valuable resource for genetic and genomic studies in *P. pseudoannulata*.

Additional files

Additional file 1: Table S1. The unigenes annotated by BLASTx against NCBI Nr database. (XLSX 59937 kb)

Additional file 2: Table S2. The unigenes annotated by BLASTx against Swiss-Prot database. (XLSX 54995 kb)

Additional file 3: Table S3. GO annotation of unigene. (XLSX 2004 kb)

Additional file 4: Table S4. The number of annotated unigenes mapped to GO database. (XLSX 1027 kb)

Additional file 5: Table S5. KOG annotation of unigene. Annotated unigenes were categorized into 25 subgroups and listed in each sheet. (XLSX 1508 kb)

Additional file 6: Table S6. Genes involved in environmental stress and immunity based on GO annotation. (XLSX 4058 kb)

Additional file 7: Figure S1. DEGs from Bt-treated and control venom apparatus. Differential gene expression was analyzed using the DESeq package and plotted as an MA plot of log_2 fold change versus the averages of the normalized counts. Each point represents a gene (circle) or a novel transcribed unit (triangle). Genes marked in red were detected as differentially expressed at a 1% FDR with more than a 2-fold change. (PDF 598 kb)

Additional file 8: Table S7. A list of DEGs based on negative binomial distribution. *P*-value (<0.01) and absolute value of log_2FoldChange (>2) were used as a cut-off. (XLSX 44 kb)

Additional file 9: Table S8. Pathway enrichment analysis of DEGs. (XLSX 22 kb)

Additional file 10: Table S9. A list of DEGs annotated in GO database. (XLSX 135 kb)

Additional file 11: Figure S2. DEGs annotated in GO database by WEGO. (PDF 6 kb)

Abbreviations
BLAST: Basic local alignment search tool; EST: Expressed sequence tag; KEGG: Kyoto encyclopedia of genes and genomes; NCBI: National center for biotechnology information; qPCR: Quantitative real-time polymerase chain reaction

Availability of data and materials
The data sets supporting the results of this article are included within the article and its additional files.

Authors' contributions
This work was done at College of Bioscience & Biotechnology, Hunan Agriculture University, Changsha, China. ZW conceived and designed the experiments. JW prepared the samples, extracted the total RNA and performed the qRT-PCR. RL and ZZY performed the de novo assembly and data analysis. RL, QSS, and ZW wrote and revised the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate
All experimental protocols and spider handling procedures were conducted in accordance with the guidelines and procedures approved by Institutional Animal Care and Use Committee of Hunan Agricultural University, Changsha of China.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]College of Bioscience & Biotechnology, Hunan Agriculture University, Changsha 410128, China. [2]Department of Biosciences, Hunan University of Arts and Science, Changde 415000, China. [3]Division of Plant Sciences, University of Missouri, Columbia, MO 65211, USA.

References
1. Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JD, King GF, et al. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annu Rev Genomics Hum Genet. 2009;10:483–511.
2. Casewell NR, Wuster W, Vonk FJ, Harrison RA, Fry BG. Complex cocktails: the evolutionary novelty of venoms. Trends Ecol Evol. 2013;28(4):219–29.
3. Pla D, Sanz L, Whiteley G, Wagstaff SC, Harrison RA, Casewell NR, Calvete JJ. What killed Karl Patterson Schmidt? Combined venom gland transcriptomic, venomic and antivenomic analysis of the south African green tree snake (the boomslang), *Dispholidus typus*. Biochim Biophys Acta. 1861;2017:814–23.
4. Carlos ES, Jimena IC, Cesar VFB, Ernesto O, Lourival DP. Venom gland transcriptomic and proteomic analyses of the enigmatic scorpion *Superstitionia donensis* (Scorpiones: Superstitioniidae), with insights on the evolution of its venom components. Toxins. 2016;8:367. doi: 10.3390/toxins8120367.

Li *et al. BMC Biotechnology* (2017) 17:73

Page 10 of 11

5. Nelson GOJ, Gabriel RF, Marlon HC, Fabrício FC, Elizabete SC, Domingos GN, Márcia RM, Elisabeth FS, Octávio LF, AAl S. Venom gland transcriptome analyses of two freshwater stingrays (Myliobatiformes: Potamotrygonidae) from Brazil. Sci Rep. 2016. doi:10.1038/srep21935.

6. Andre DS, David W. The venom gland transcriptome of the parasitoid wasp *Nasonia vitripennis* highlights the importance of novel genes in venom function. BMC Genomics. 2016;17:571–87.

7. Chaim OM, Trevisan-Silva D, Chaves-Moreira D, Wille ACM, Ferrer VP, Matsubara FH, et al. Brown spider (*Loxosceles genus*) venom toxins: tools for biological purposes. Toxins. 2011;3:309–44.

8. Saez NJ, Senff S, Jensen JE, Er SY, Herzig V, Rash LD, et al. Spider-venom peptides as therapeutics. Toxins (Basel) 2010; 2: 2851-2871.

9. Windley MJ, Herzig V, Dziemborowicz SA, Hardy MC, King GF, Nicholson GM. Spider-venom peptides as bioinsecticides. Toxins. 2012;4:191–227.

10. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, et al. Library construction for next-generation sequencing: overviews and challenges. Biotechniques. 2014; 56(2): 61-64, 66, 68, passim.

11. Höfte H, Whiteley HR. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol Rev. 1989;53(2):242–55.

12. Tian JC, Chen Y, Li ZL, Li K, Chen M, Peng YF, et al. Transgenic Cry1Ab rice does not impact ecological fitness and predation of a generalist spider. PLoS One. 2012;7(4):e35164.

13. Sparks ME, Blackburn MB, Kuhar D, GundersenRindal DE. Transcriptome of the *Lymantria dispar* (gypsy moth) larval midgut in response to infection by *Bacillus thuringiensis*. PLoS One. 2013;8(5):43–51.

14. Tang X, Han F, Zhao K, Xu Y, Wu X, Wang J, et al. A 90-day dietary toxicity study of genetically modified rice T1C-1 expressing Cry1C protein in Sprague Dawley rats. PLoS One. 2012;7(12):e52507.

15. Sanden M, Ornsrud R, Sissener NH, Jorgensen S, Gu J, Bakke AM, et al. Cross-generational feeding of Bt (*Bacillus thuringiensis*)-maize to zebrafish (Danio Rerio) showed no adverse effects on the parental or offspring generations. Br J Nutr. 2013;110(12):2222–33.

16. Ramirez-Romero R, Desneux N, Decourtye A, Chaffiol A, Pham-Delegue MH. Does Cry1Ab protein affect learning performances of the honey bee *Apis mellifera* L. (hymenoptera, Apidae)? Ecotoxicol Environ Saf. 2008;70(2):327–33.

17. Liu J, Li M. Bt cotton impacts on the growth and predation behavior of spiders. Acta Ecol Sin. 2006;26(3):945–9. (in Chinese)

18. Raybould A, Vlachos D. Non-target organism effects tests on Vip3A and their application to the ecological risk assessment for cultivation of MIR162 maize. Transgenic Res. 2011;20(3):599–611.

19. Zhao YJ, Zeng Y, Chen L, Dong Y, Wang W. Analysis of transcriptomes of three orb-web spider species reveals gene profiles involved in silk and toxin. Insect Sci. 2014;21(6):687–98.

20. Jiang LP, Liu CJ, Duan ZG, Deng MC, Tang X, Liang SP. Transcriptome analysis of venom glands from a single fishing spider *Dolomedes mizhoanus*. Toxicon. 73:23–32.

21. Zhou J, Xiao KF, Wei BY, Wang Z, Tian Y, Tian YX, Song QS. Bioaccumulation of Cry1Ab protein from an herbivore reduces anti-oxidant enzyme activities in two spider species. PLoS One. 2014;9:e84724.

22. Wang J, Peng YD, Xiao KF, Wei BY, JL H, Wang Z, Song QS, Zhou XG. Transcriptomic response of wolf spider, *Pardosa pseudoannulata*, to transgenic rice expressing *Bacillus thuringiensis* Cry1Ab protein. BMC Biotechnol. 2017;17:7–14.

23. Grbic M, Van Leeuwen T, Clark RM, Rombauts S, Rouze P, Grbic V, et al. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature. 2011;479(7374):487–92.

24. Hedin M, Starrett J, Akhter S, Schönhofer AL, Shultz JW. Phylogenomic resolution of paleozoic divergences in harvestmen (Arachnida, Opiliones) via analysis of next-generation transcriptome data. PLoS One. 2012;7(8):e42888.

25. Xu Z, Zhu W, Liu Y, Liu X, Chen Q, Peng M, et al. Analysis of insecticide resistance-related genes of the carmine spider mite *Tetranychus cinnabarinus* based on a de novo assembled transcriptome. PLoS One. 2014; 9(5):e94779.

26. Croucher PJP, Brewer MS, Winchell CJ, Oxford GS, Gillespie RG. De novo characterization of the gene-rich transcriptomes of two color-polymorphic spiders, *Theridion grallator* and *T. californicum* (Araneae: Theridiidae), with special reference to pigment genes. BMC Genomics. 2013;14:862.

27. Schwartz EF, Diego-Garcia E, Rodriguez de la Vega RC, Possani LD. Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus gertschi* (Arachnida: Scorpiones). BMC Genomics. 2007;8:119.

28. He Q, Duan Z, Yu Y, Liu Z, Liu Z, Liang S. The venom gland Transcriptome of *Latrodectus tredecimguttatus* revealed by deep sequencing and cDNA library analysis. PLoS One. 2013;8(11):e81357.

29. Prosdocimi F, Bittencourt D, da Silva FR, Kirst M, Motta PC, Rech EL. Spinning gland transcriptomics from two main clades of spiders (order: Araneae)-insights on their molecular, anatomical and behavioral evolution. PLoS One. 2011;6(6):e21634.

30. Zhao WC, Zhang WJ. Evaluation of the control effects of *Pardosa pseudoannulata* on *Nilaparvata lugens* (Sta°l) with amonoclonal antibody. Acta Ecol Sin. 2005;25(1):78–82. (in Chinese)

31. Schwarz A, von Reumont Bö M, Erhart J, Chagas AC, Ribeiro Jé MC, Kotsyfakis M. De novo Ixodes ricinus salivary gland transcriptome analysis using two next-generation sequencing methodologies. FASEB J. 2013;27(12):4745–56.

32. Sauberer N, Zulka KP, Abensperg-Traun M, Berg H-M, Bieringer G, Milasowszky N. Surrogate taxa for biodiversity in agricultural landscapes of eastern Austria. Biol Conserv. 2004;117(2):181–91.

33. Tian YX, Zhou Y, Xiao KF, Wang Z, CJ J. Effect of Cry1Ab protein on hemocytes of the wolf spider *Pardosa pseudoannulata*. Biocontrol Sci Tech. 2013;23:423–32.

34. Li SW, Yang H, Liu YF, Liao QR, Du J, Jin DC. Transcriptome and gene expression analysis of the rice leaf folder, *Cnaphalocrosis medinalis*. PLoS One. 2012;7(11):e47401.

35. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644–52.

36. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with trinity. Nat Protoc. 2013;8(8):1494–512.

37. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9(4):357–9.

38. Audic S, Claverie JM. The significance of digital gene expression profiles. Genome Res. 1997;7(10):986–95.

39. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) method. Methods. 2001;25(4):402–8.

40. Blake JA. Ten quick tips for using the gene ontology. PLoS Comput Biol. 2013;9(11):e1003343.

41. Kaul G, Thippeswamy H. Role of heat shock proteins in diseases and their therapeutic potential. Indian J Microbiol. 2011;51(2):124–31.

42. Gendreau KL, Haney RA, Schwager EE, Wierschin T, Stanke M, Richards S, Garb JE. House spider genome uncovers evolutionary shifts in the diversity and expression of black widow venom proteins associated with extreme toxicity. BMC Genomics. 2017;18:178–92.

43. Haney RA, Clarke TH, Gadgil R, Fitzpatrick R, Hayashi CY, Ayoub NA, Garb JE. Effects of gene duplication, positive selection, and shifts in gene expression on the evolution of the venom gland transcriptome in widow spiders. Genome Biol Evol. 2016;8(1):228–42.

44. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics. 2010;26(1):136–8.

45. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11:R106.

46. Iseli C, Jongeneel CV, Bucher P. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. Proc Int Conf Intell Syst Mol Biol. 1999:138–48.

47. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100:57–70.

48. Lichtenstein AV. Cancer: evolutionary, genetic and epigenetic aspects. Clin Epigenetics. 2010;1(3):85–100.

49. Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier ME, Mitros T, et al. The *Amphimedon queenslandica* genome and the evolution of animal complexity. Nature. 2010;466(7307):720–6.

50. Soderhall K, Cerenius L. Role of the prophenoloxidase-activating system in invertebrate immunity. Curr Opin Immunol. 1998;10(1):23–8.

51. Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RA. Phylogenetic perspectives in innate immunity. Science. 1999;284(5418):1313–8.

52. Iwanaga S, Lee BL. Recent advances in the innate immunity of invertebrate animals. J Biochem Mol Biol. 2005;38(2):128–50.

53. Evans JD, Aronstein K, Chen YP, Hetru C, Imler JL, Jiang H, et al. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect Mol Biol. 2006;15(5):645–56.

54. Tanaka H, Ishibashi J, Fujita K, Nakajima Y, Sagisaka A, Tomimoto K, et al. A genome-wide analysis of genes and gene families involved in innate immunity of *Bombyx mori*. Insect Biochem Mol Biol. 2008;38(12):1087–110.

55. Romeo Y, Lemaitre B. *Drosophila* immunity. Humana Press. 2008;415:379–94.

56. Zou Z, Evans JD, Lu Z, Zhao P, Williams M, Sumathipala N, et al. Comparative genomic analysis of the *Tribolium* immune system. Genome Biol. 2007;8(8):R177.

57. Margaret CH, Norelle LD, Mehdi M, Rodrigo AVM, Glenn FK. Isolation of an orally active insecticidal toxin from the venom of an Australian tarantula. PLoS One. 2013;8:e73136.

58. Amara U, Flierl MA, Rittirsch D, Klos A, Chen H, Acker B, et al. Molecular intercommunication between the complement and coagulation systems. J Immunol. 2010;185(9):5628–36.

59. Theopold U, Schmidt O, Soderhall K, Dushay MS. Coagulation in arthropods: defence, wound closure and healing. Trends Immunol. 2004;25(6):289–94.

60. Smith AA, Pal U. Immunity-related genes in *Ixodes scapularis*-perspectives from genome information. Front Cell Infect Microbiol. 2014;4:116.

61. Semple JW, Italiano JE, Freedman J. Platelets and the immune continuum. Nat Rev Immunol. 2011;11(4):264–74.

62. Veiga ABG, Ribeiro Jé MC, Guimarães JA, Francischetti IMBA. Catalog for the transcripts from the venomous structures of the caterpillar *Lonomia obliqua*: identification of the proteins potentially involved in the coagulation disorder and hemorrhagic syndrome. Gene. 2005;355:11–27.

63. Song X, Jin P, Qin S, Chen L, Ma F. The evolution and origin of animal toll-like receptor signaling pathway revealed by network-level molecular evolutionary analyses. PLoS One. 2012;7(12):e51657.

64. Hildebrand JM, Yi Z, Buchta CM, Poovassery J, Stunz LL, Bishop GA. Roles of tumor necrosis factor receptor associated factor 3 (TRAF3) and TRAF5 in immune cell functions. Immunol Rev. 2011;244(1):55–74.

65. Baeg GH, Zhou R, Perrimon N. Genome-wide RNAi analysis of JAK/STAT signaling components in *Drosophila*. Genes Dev. 2005;19(16):1861–70.

66. Bonizzi G, Karin M. The two NF-kappa B activation pathways and their role in innate and adaptive immunity. Trends Immunol. 2004;25(6):280–8.

67. Xu J, Strange JP, Welker DL, James RR. Detoxification and stress response genes expressed in a western north American bumble bee, *Bombus huntii* (Hymenoptera: Apidae). BMC Genomics. 2013;14:874.

68. David JP, Strode C, Vontas J, Nikou D, Vaughan A, Pignatelli PM, et al. The *Anopheles gambiae* detoxification chip: a highly specific microarray to study metabolic-based insecticide resistance in malaria vectors. Proc Natl Acad Sci U S A. 2005;102(11):4080–4.

69. Claudianos C, Ranson H, Johnson RM, Biswas S, Schuler MA, Berenbaum MR, et al. A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. Insect Mol Biol. 2006;15(5):615–36.

70. Willoughby L, Chung H, Lumb C, Robin C, Batterham P, Daborn PJA. Comparison of *Drosophila melanogaster* detoxification gene induction responses for six insecticides, caffeine and phenobarbital. Insect Biochem Mol Biol. 2006;36(12):934–42.

71. Poupardin R, Reynaud S, Strode C, Ranson H, Vontas J, David JP. Cross-induction of detoxification genes by environmental xenobiotics and insecticides in the mosquito *Aedes aegypti*: impact on larval tolerance to chemical insecticides. Insect Biochem Mol Biol. 2008;38(5):540–51.

72. Oakeshott JG, Johnson RM, Berenbaum MR, Ranson H, Cristino AS, Claudianos C. Metabolic enzymes associated with xenobiotic and chemosensory responses in *Nasonia vitripennis*. Insect Mol Biol. 2010;19(Suppl 1):147–63.