

Differential Gene Expression Analysis of the *Epacromius coerulipes* (Orthoptera: Acrididae) Transcriptome

Yongling Jin,^{1,2,3} Bin Cong,¹ Liyan Wang,² Yugang Gao,² Haiyan Zhang,² Hui Dong,¹ and Zhiwei Lin²

¹Shenyang Agriculture University, Shenyang of Liaoning province, China, 110061 (gygjyl08@163.com; bin1956@163.com; Biocontrol@163.com), ²Heilongjiang Bayi Agricultural University, Daqing of Heilongjiang province, China, 163319 ((byndwly@126.com; nongdaggy@163.com; zhangyh51@126.com; lzwsdm@sohu.com), and ³Corresponding author, e-mail: gygjyl08@163.com

Subject Editor: Swale Daniel

Received 12 December 2015; Accepted 5 February 2016

Abstract

Epacromius coerulipes (Ivanov) is one of the most widely distributed locusts. To date, the main methods to kill locusts still rely on chemical controls, which can result in the selection of locusts with resistance to chemical pesticides. Butene-fipronil is a new pesticide that was discovered by the structural modification of fipronil. This pesticide has been used to control various agricultural pests and has become an important pesticide product to control pests that exhibit resistance to other pesticides, including locusts. To extend its useful half-life, studies of the initiation and progression of resistance to this pesticide are needed. Herein, two *E. coerulipes* strains, a pesticide-sensitive (PS) and a pesticide-resistant (PR) strain, were chosen to undergo de novo assembly by paired-end transcriptome Illumina sequencing. Overall, 63,033 unigenes were detected; the average gene length was 772 bp and the N50 was 1,589 bp. Among these unigenes, ~25,132 (39.87% of the total) could be identified as known proteins in bioinformatic databases from national centers. A comparison of the PR and PS strains revealed that 2,568 genes were differentially expressed, including 1,646 and 922 genes that were up- and down-regulated, respectively. According to the Gene Ontology (GO) database, among biological processes the metabolic process group was the largest group (6,900 genes, 22.47%) and contained a high frequency of differentially expressed genes (544 genes, 27.54%). According to the Clusters of Orthologous Groups (COG) categories, 28 genes, representing 2.98% of all genes, belonged to the group of genes involved in the biosynthesis, transportation, and catabolism of secondary metabolites. The differentially expressed genes that we identified are involved in 50 metabolic pathways. Among these pathways, the metabolism pathway was the most represented. After enrichment analysis of differential gene expression pathways, six pathways—ribosome; starch, and sucrose metabolism; ascorbate and aldarate metabolism; drug metabolism-cytochrome P450; metabolism of xenobiotics by cytochrome P450; and glutathione metabolism—showed a high degree of enrichment. Among these pathways, drug metabolism-cytochrome P450, metabolism of xenobiotics by cytochrome P450, and glutathione metabolism have been associated with pesticide metabolism. Furthermore, 316 unigenes in the *E. coerulipes* transcriptome encode detoxifying enzymes and 76 unigenes encode target proteins of pesticides. Among these genes, 23 genes that encode detoxifying enzymes in the resistance group were found to be up-regulated. The transcriptome sequencing results of *E. coerulipes* established a genomics database of *E. coerulipes* for the first time. This study also establishes a molecular basis for gene function analysis of *E. coerulipes*. Moreover, it provides a theoretical resource for mechanistic studies on pesticide resistance through the screening and investigation of resistance genes.

Key words: pesticide resistance, *Epacromius coerulipes* (Ivanov), transcriptome, differential gene expression analysis

Heilongjiang is one of the 10 provinces that contain grasslands in China, and it contains 4.33 million hm² of grassland. However, ~1.33 million hm² of grassland are attacked by locusts annually and 0.66–0.87 million hm² of grassland become disaster areas (Zhou et al. 2010). Recently, changes in the climate and cropland environment in northern China have aggravated the regional occurrence of

locusts and resulted in an increasing trend for various locust together hazards (Hong et al. 2014). Previous studies of the prairie locust in Heilongjiang reported that the annual growth of dominant locust species in Heilongjiang prairies, such as *Epacromius coerulipes* (Ivanov) and *Oxya chinensis* Thunberg, can reach a population density of 300–500 insects/m² and greater. The occurrence rate of

Epacromius coerulipes accounts for up to 30% of the total, and exceeds 50% in some regions, with serious and frequent occurrence each year and occasional outbursts (Liu et al. 2008, Zhou et al. 2011, Wang et al. 2014).

E. coerulipes, an insect belonging to Orthoptera, Acridoidea, Edipodidae, is one of the locusts that is widely distributed in the grasslands of China and in areas of both farming and herding. It not only harms grasslands, as they have many hosts, and particularly thrive on gramineous forage grass. They can grow in vegetation-sparse wasteland and saline-alkali soil, degenerate grassland, and aggravate salt and alkali conditions (Tian et al. 2009); this can establish a vicious circle of insect proliferation and damage to the ecological balance in the prairie. *E. coerulipes* also damages gramineous crops and broad-leaved crops, such as soybeans and alfalfa (Tian et al. 2003, Tian et al. 2010, Xu et al. 2013). Usually, hazard to crops begins from the edge of a field, then expands to the middle and even damages all crops in serious occurrences. Meanwhile, *E. coerulipes* can fly short distances in cases of food shortage and cause widespread and severe economic damage to agriculture and husbandry.

Because locust control in China focuses on chemical pesticides, locusts have begun to acquire chemical pesticide resistance and the efficiency of locust control decreases throughout the continuous occurrence of locust plagues. Consequently, it is important to study the molecular mechanisms of pesticides on locusts to best use chemical pesticides and extend their life span to sustainably control locusts. However, there are few studies on pesticide resistance genes in locusts. Studies of pesticide resistance, aided by transcriptome technology, have been restricted to migratory locusts rather than *E. coerulipes*. Therefore, high-flux Illumina sequencing techniques have been applied to discover pesticide resistance genes by assembly, functional annotation, and analysis, which has greatly improved our knowledge of *E. coerulipes* pesticide resistance from a molecular perspective by studying resistance mechanisms to provide a theoretical basis for mixing and rotationally using chemical pesticides. In addition, such efforts may allow the frequency of resistance genes to be limited as a consequence of improved chemical prevention technology for locusts. Meanwhile, such types of databases can provide useful data for the study of the genes and genome of *E. coerulipes*.

Materials and Methods

Sampling and Feeding of Insects

Insects

E. coerulipes are collected from the natural prairie in Honggang District, Daqing, Heilongjiang, with sensitivity to butene-fipronil of $LD_{50} = 0.0303 \mu\text{g}/\text{larva}$. They are subcultured in insect rearing cages (length, height, and width of 1 by 1 by 1 m; cages were covered with screens) under conditions of $28 \pm 2^\circ\text{C}$, $65 \pm 6\%$ relative humidity, and a photoperiod of 14:10 (L:D) h and fed fresh wheat seedlings.

Sensitive Strain Selection

The purpose was to develop sensitive strains by nonpesticide treatments. After subculturing generations F5 and F7 in insect rearing rooms, sensitivity to butene-fipronil reached $LD_{50} = 0.0298 \mu\text{g}/\text{larva}$ and $LD_{50} = 0.0299 \mu\text{g}/\text{larva}$, respectively. If an *E. coerulipes* strain was highly sensitive to butene-fipronil, it is classified as a sensitive strain for the following experiment, hereinafter referred as PS.

Butene-Fipronil Resistance Strain Selection

If insects were sprayed with butene-fipronil and fed wheat seedlings, their death rate ranged from 30 to 70% for resistance strain selection. Compared with sensitive strains, generation F5 $LD_{50} = 0.1773 \mu\text{g}/\text{larva}$ yielded a resistance ratio of 5.95, while the F7 $LD_{50} = 0.3499 \mu\text{g}/\text{larva}$ with a resistance ratio of 11.74 after toxicity tests were conducted for subculture selection in insect rearing rooms. In such cases, F7 will be applied as the resistance strain for later experiments, hereinafter referred as PR.

Extraction and Testing of RNA

Fifth-instar larvae female nymphs of the above two strains (PS and PR strain) were used for RNA extraction. Total RNA was extracted following the manufacturer's instructions (Invitrogen Trizol Reagent of RNA extraction kit 15596018). This was carried out to ensure that the samples for transcriptome sequencing were of a sufficient quality. The purity, concentration, and integrity of RNA samples were tested according to the manufacturer's methods for the gel and Nanodrop 2000 (Thermo Fisher Scientific) and Agilent Bioanalyzer 2100 (Agilent).

cDNA Library Construction and Sequencing

When the quality of RNA samples was confirmed, using a NEB kit for library preparation, a cDNA library was constructed as follows: magnetic beads coupled with oligo(dT) were applied to enrich mRNA in eukaryotes. The mRNA transcripts were randomly broken by bivalent cations under high temperature conditions (94°C). The first cDNA chain was synthesized using a target fragment of mRNA as a template and randomly synthesized hexamers as primers. Then, buffer solution, dNTPs, RNase H, and DNA polymerase I were added to synthesize the second cDNA chain. The cDNA was next purified with AMPure XP beads. Purified double-stranded cDNA was repaired at the end, and then an A-tail was added and ligated with a sequencing joint. The fragment size (the actual size was 150–250 bp without a joint) was selected using AMPure XP beads (cDNA purification with 1.8 times the magnetic beads, i.e., 100 μl reaction system and 180 μl magnetic beads). The cDNA library was obtained after PCR enrichment (PCR conditions: 98°C 10 s, 65°C 30 s, 72°C 30 s, 72°C 5 min, 12 cycles). The prepared DNA library was subjected to sequencing using an Illumina HiSeq2500 (Beijing Biomarker Technologies Co.) with a read length of PE125. Two samples are sequenced in the same RUN, but a different index was added to distinguish them before sequencing.

The Transcriptome Assembly

Sample sequences were assembled to indirectly deepen the sequencing results and obtain a more complete transcriptome dataset. To obtain high-quality clean data, raw sequencing data were processed to remove sequencing joints, poly-N regions, and low-quality reads. Meanwhile, the Q20, Q30, GC content, and repetitive sequences were calculated and all downstream analyses were carried out using high-quality clean data (Q30 > 85%). High-quality sequencing data were assembled using Trinity software (Grabherr et al. 2011). The sequencing reads were first broken into smaller fragments (K-mer), and then these fragments were extended to generate a longer fragment (Contig). The fragment collections (Component) were obtained based on overlaps of these fragments. Transcript sequences were then recognized in each fragment collection using a De Bruijn graph-based method and sequencing reads data (Grabherr et al. 2011, Haas et al. 2013).

Functional Annotation of Genes

Unigene sequences were compared in the following databases using BLAST software (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) to obtain annotation information for unigenes. Based on BLAST parameters, those unigenes with an E-value less than 10^{-5} were selected. Sequences of selected unigenes were aligned within databases, including the NCBI nonredundant protein sequences (Nr: <ftp://ftp.ncbi.nih.gov/blast/db/>), Clusters of Orthologous Groups (COG: <http://www.ncbi.nlm.nih.gov/COG/>), a manually annotated and reviewed protein sequence database (Swiss-Prot: <http://www.uniprot.org/>), the Kyoto Encyclopedia of Genes and Genomes (KEGG: <http://www.genome.jp/kegg/>), and Gene Ontology (GO: <http://www.geneontology.org/>) databases.

Calculating the Expression Level of Unigenes

Sequencing reads of each sample were compared with the unigene database using the Bowtie method (Langmead et al. 2009). The comparison results and RSEM (Li et al. 2011) (RSEM is software used to calculate levels) were used to estimate the expression levels with the FPKM value (FPKM represents the expression level of genes) representing the gene expression level.

$$\text{FPKM} = \frac{\text{cDNA Fragments}}{\text{Mapped Fragments (Millions)} \times \text{Transcript Length (kb)}}$$

In the above formula, cDNA fragments represent the number of aligned fragments within certain transcripts and indicate the number of double reads. Mapped fragments (Millions) represent the total number of aligned fragments within transcripts with a unit of 10^6 . Transcript length (kb) is the length of transcripts with a unit of 10^3 bases. Then, information was obtained for each gene.

Detection and Cluster Analysis of Differentially Expressed Genes

Differentially expressed gene sets for the two samples were acquired by differential expression analysis using EBSeq (Leng et al. 2013). In the differential gene expression analysis, significant p-values based on the original hypothesis was corrected using the Benjamini-Hochberg method (Anders et al. 2010). The corrected p-value, i.e., the False Discovery Rate (FDR), was used as a key index to screen differentially expressed genes to minimize the false-positive results caused by independent statistical hypothesis tests for high-value gene expression values. During this screening process, $\text{FDR} < 0.01$ and Fold Change (FC) ≥ 2 were used as screening criteria. FC represented the ratio of expression values between two samples. The selected differentially expressed genes were analyzed for hierarchical clustering (Murtagh et al. 2014). Identically or similarly expressed genes were clustered to reveal the various expression patterns under different experimental conditions.

Enrichment Analysis of Differentially Expressed Genes

After be identified by screening, the differentially expressed genes were analyzed by GO function enrichment, COG classification, pathway annotation analysis, and KEGG pathway enrichment analysis.

The differentially expressed genes identified by screening were used to do GO function enrichment and draw histogram: extract GO annotation of differential expression gene, map GO function to the corresponding secondary features on the background of all unigene's GO annotation (Ye et al. 2006), and then obtain accurate results of GO function enrichment by using the Fisher's exact test.

COG classification was used to extract the Cog annotation results of differential expression gene, and then drawing histogram to select (<http://www.ncbi.nlm.nih.gov/COG/>).

KEGG pathway enrichment analysis by KOBAS2.0 (<http://kobas.cbi.pku.edu.cn/home.do>) (Xie et al. 2011) extracted the pathway information of differential expression gene, and then on the background of all unigene's pathway information, calculated the enrichment factor and detected significance by the Fisher's exact test. When enrichment factors were smaller and the absolute Q value was higher, the enrichment level of a differentially expressed gene was more significant and reliable. The calculation formula for the Enrichment Factor was as follows:

$$\text{Enrichment Factor} = \frac{\text{Total number of differentially expressed genes} / \text{Total gene number in the KEGG}}{\text{Number of differentially expressed genes in a pathway} / \text{Total gene number in a pathway}}$$

Results

Sequencing Analysis and Assembly

The extracted cDNA samples from *E. coeruleipes* were subjected to sequencing using an Illumina sequencing platform. The data shown in Table 1 indicate that the sequencing results were highly accurate and could be utilized for future studies. The sequencing assembly was performed using Trinity software (Grabherr et al. 2011), and the parameters are shown in Table 2. These findings suggest that the sequencing assembly was ideal.

Gene Functional Annotation

As shown in Table 3, 25,132 unigenes were successfully annotated, which represents 39.87% of all genes. Among these genes, there were 24,841 Nr database annotated unigenes, which represents 39.41% of all genes. The Swissprot database contained 16,490 annotated unigenes, which represents 26.16% of total genes. The COG database had 8,031 annotated unigenes, which represents 12.71% of all genes. The GO database included 11,558 annotated unigenes, which represents 18.34% of total genes. The KEGG database had 7,218 annotated unigenes, which represents 11.46% of all genes. However, 37,901 unigenes, which represents 60.13% of all genes, were unannotated. These findings indicate that these unigenes can yield short reads by sequencing technology, and can be subjected to various types of analysis (Hou et al. 2011).

Annotation and Clustering Analysis of Differentially Expressed Genes

Statistical analyses and annotations for differentially expressed genes between the PS and PR strains of *E. coeruleipes* are presented in Table 4, among which total 2,568 (4.07%) differentially expressed genes were acquired, and proportion of total annotations in each database reached 7.15%. Up-regulated genes accounted for

Table 1. Summary of the raw reads for the two samples

Samples	Read number	Clean Data	GC content	Cycle Q20%	% \geq Q30
PR (T01)	26,869,615	6,766,002,460	50.71%	100.00	90.63%
PS (T02)	26,428,337	6,653,605,161	50.81%	100.00	90.09%

Note: GC content: Clean Data G and C percentage of the total bases; Cycle Q20: Average quality score is greater than or equal to 20% of the cycle, Q30: Quality Score of base is greater than or equal to 30% of the total bases.

64.10% of all differentially expressed genes. The blue and red zones, respectively, represent differences with low and high levels of significance (Fig. 1).

Functional Enrichment Analysis of Differentially Expressed Genes

The GO database contains three main categories of annotation: molecular function, cellular component, and biological process (Fig. 2). In the cellular component group, the 3,000 unigenes in a cell or cellular organelle were the largest category. The number of differentially expressed genes was also the greatest, and both the number of unigenes in a cell or cellular organelle were greater than 270. In the molecular function group, the total number of catalytic activity and differentially expressed genes from them were greatest and represented the highest proportion of all genes. The proportion of differentially expressed genes increased in certain subgroups, including structural molecular activity, electron carrier activity, antioxidant activity, and channel regulator activity. In the biological processes group, the number of genes involved in metabolic processes was the greatest (6,870, representing 27.34% of all genes). The number of differentially expressed genes was also the greatest and had the highest proportion (544, 30.25%) of all genes. The proportion of

Table 2. Summary of Illumina transcriptome assembly for *E. coeruleipes*

Length range	Contig	Transcript	Unigene
200–300	6,232,095 (99.35%)	30,249 (31.46%)	26,218 (41.59%)
300–500	15,971 (0.25%)	18,617 (19.36%)	14,150 (22.45%)
500–1,000	10,968 (0.17%)	16,364 (17.02%)	9,772 (15.50%)
1,000–2,000	7,402 (0.12%)	14,974 (15.57%)	6,855 (10.88%)
2,000+	6,195 (0.10%)	15,947 (16.59%)	6,038 (9.58%)
Total number	6,272,631	96,151	63,033
Total length	306,171,970	105,999,072	48,663,063
N50 length	47	2,269	1,589
Mean length	48.81	1,102.42	772.03

Note: The sequencing assembly obtained 6,272,631 Contigs and the N50 length was 47.96, 151 transcripts were obtained. The total length was 105.999 Mb, the average length was 1,102 bp, and the N50 length was 2,269. Among them, the number of transcripts with a length over 1 kb was 30,921, which was 32.16% of the total number. Additionally, 15,947 of transcripts with a length over 2 kb were identified, which was 16.59% of all transcripts. After clustering and assembly analysis of the transcripts, 63,033 unigenes were obtained, with a total length of 48.663 Mb and an average length of 772 bp, and a N50 length of 1,589 bp. Among them, 12,893 unigenes with a length over 1 kb were identified, which represents 20.45% of the total, and 6,083 unigenes with a length over 2 kb were identified, which represents 9.58% of the total.

Table 3. Functional annotation of the *E. coeruleipes* transcriptome

Annotated database	Annotated unigenes number	300 ≤ length < 1,000	Length ≥ 1,000
COG_Annotation	8,013 (12.71)	2,603	3,955
GO_Annotation	11,558 (18.34%)	3,863	5,064
KEGG_Annotation	7,218 (11.46%)	2,260	3,694
Swissprot_Annotation	16,490 (26.16%)	5,676	7,889
nr_Annotation	24,841 (39.41%)	9,079	9,881
All_Annotated	25,132 (39.87%)	9,185	9,895

Note: Nr database annotated 24,841 unigenes, which represents 39.41% of the total. The Swissprot database annotated 16,490 unigenes, which represents 26.16% of the total. The COG database annotated 8,031 unigenes, which represents 12.71% of the total. The GO database annotated 11,558 unigenes, which represents 18.34% of the total. The KEGG database annotated 7,218 unigenes, which represents 11.46% of the total.

differentially expressed genes (132, 0.43%; 10, 0.51%) also increased in the immune system processes subgroup.

In Fig. 3, general function prediction alone was the largest class group; the number of differentially expressed genes was 133, which represents 14.16% of all genes. The number of differentially expressed genes involved in carbohydrate transport and metabolism represents 12.89%; in translation, ribosomal structure, and biogenesis represents 12.14%; the number of differentially expressed genes in other subgroups represents <10%. Because secondary metabolites in insects significantly affect insecticides, secondary metabolite biosynthesis, transport, and catabolism is an important class group among the COG categories. The number of differentially expressed genes was 28, which represents 2.98% of the total genes.

As shown in Fig. 4, the KEGG annotation results of differentially expressed genes were classified according to KEGG pathway classification. The differentially expressed genes were distributed across 50 metabolic pathways. Among them, the metabolism pathway included the greatest number of genes. The number of differentially expressed genes involved in the metabolism of xenobiotics by cytochrome P450, drug metabolism of cytochrome P450, drug metabolism of other enzymes, and glutathione metabolism, were 14, 15, 7, and 15, respectively. These metabolic pathways are all related to the degradation of pesticides in insects.

KEGG enrichment analysis of differentially expressed genes is shown in Fig. 5, and 20 selected metabolic pathways were analyzed. This analysis indicated that six metabolic pathways are of the highest concentration when diagrams of the top 20 metabolic pathways with high concentration were analyzed. These six pathways—ribosome, starch and sucrose metabolism; ascorbate and aldarate metabolism; drug metabolism-cytochrome P450; metabolism of xenobiotics by cytochrome P450; and glutathione metabolism—showed the highest enrichment scores. Among these pathways, three pathways—drug metabolism-cytochrome P450; metabolism of xenobiotics by cytochrome P450; and glutathione metabolism—are related to pesticide metabolism.

Table 4. Number of differentially expressed genes and gene annotation

DEGs number	Annotated	COG	GO	KEGG	Swiss-Prot	nr
All DEG	2,568	1,798	760	815	584	1,424
Up-regulated	1,646	1,177	537	579	453	956
Down-regulated	922	621	223	236	131	468

Note: The 2,568 differentially expressed genes were obtained and 1,798 of them were annotated to each database. Up-regulated genes included 1,646 and 1,177 genes from those that were annotated; down-regulated genes included 922 and 621 genes from those that were annotated.

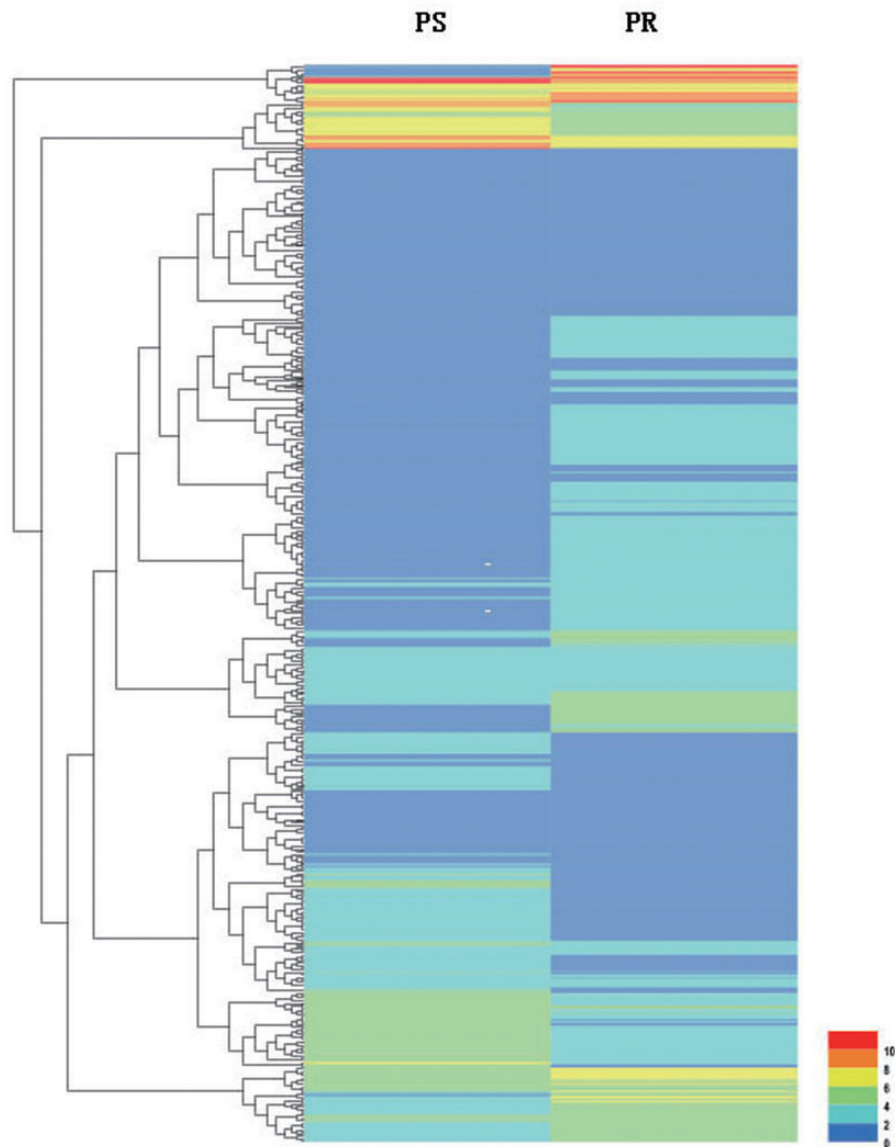


Fig. 1. Clustering gene expression pattern of differentially expressed genes. Note: Each column represents one sample, each row represents one gene, and colors represent the expression level, FPKM with \log_2 in sample genes. The blue area represents a low significant difference and the red area represents a highly significant difference.

Annotation and Analysis of Pesticide Resistance Genes

In the *E. coerulea* transcriptome, many gene sequences that encoded pesticide detoxification enzymes and target proteins were annotated. In this present study, 316 genes that encoded pesticide detoxification enzymes were annotated, including 43 glutathione transferase (GST) genes, 90 carboxylesterase (CarEs) genes, and 183 cytochrome P450 genes. There were 74 unigenes that encoded target proteins of pesticides, including 10 γ -aminobutyric acid receptor (GABA), 15 nicotinic acetylcholine receptor (nAChRs), 3 ryanodine receptor, 38 AChE enzyme, and 10 voltage-gated sodium channel (VGSC) genes. Additionally, 23 genes, including 5 GST, 8 CarEs, and 10 cytochrome P450s, were up-regulated in the PR strain compared with the PS strain (Table 5).

Discussion

As a consequence of increasing research efforts and the completion of insect genomes and transcriptomes, the breadth and depth of

molecular knowledge on insects has significantly expanded. However, locust genetics received relatively less attention. Badisco et al. (2011a) studied the central nervous system (CNS) of the desert locust (*Schistocerca gregaria gregaria*) based on expressed sequence tags (EST) reads and obtained 34,672 original EST sequences from the CNS. Those ESTs were assembled into 12,709 transcript sequences; one-third of them were annotated. Based on the EST database of *S. gregaria gregaria*, Badisco et al. (2011b) compared the CNS genes between solitary and gregarious desert locusts, and they identified 214 differentially expressed genes. Chen et al. (2010) studied the transcriptomes of the migratory locust (*Locusta migratoria*) by de novo assembly. This study yielded a deeper understanding of the genetics of incomplete metamorphosis in insects and the origins of their deformation. Furthermore, understanding the genes and pathways involved in the development of locusts and the phase transitions in locusts could be helpful for controlling these insects. Jiang et al. (2012) first analyzed retroelements in the migratory locust transcriptome and identified 105 retroelements; this finding allowed

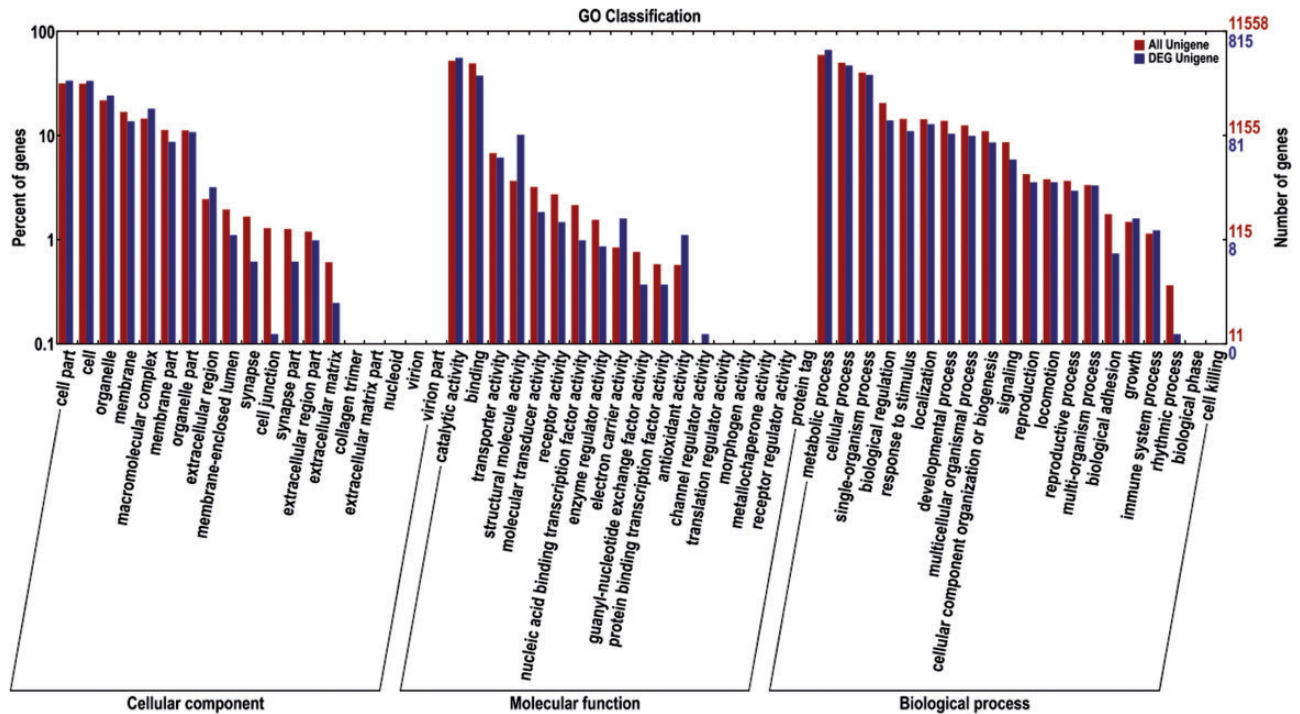


Fig. 2. GO classification. Annotation statistics of differentially expressed genes in the secondary node of GO. Note: The horizontal axis shows secondary nodes of three categories in GO. The vertical axis displays the percentage of annotated genes versus the total gene number. The red columns display annotation information of the total genes and the blue columns represent annotation information of the differentially expressed genes.

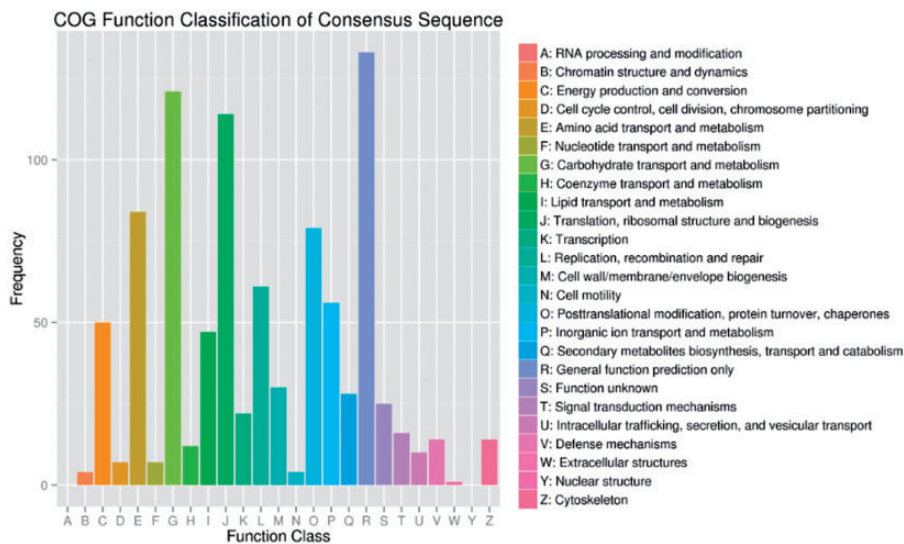


Fig. 3. COG function classification of consensus sequences. Note: The categories of the COG are shown on the horizontal axis and gene numbers are plotted on the vertical axis. The number of differentially expressed genes involved in general function prediction is 133; in carbohydrate transport and metabolism is 121; in translation, ribosomal structure, and biogenesis is 114; in amino acid transport and metabolism is 84; in posttranslational modification, protein turnover, and chaperones is 79; in replication, recombination, and repair is 61; in inorganic ion transport and metabolism is 56; in energy production and conversion is 5; in lipid transport and metabolism is 47; and in cell wall/membrane/envelope biogenesis is 30; the number of differentially expressed genes in other subgroups is <30.

us to better understand the retrotransposon distribution in the migratory locust transcriptome. Moreover, these findings showed that non-long terminal repeat (LTR) retrotransposons in migratory locust transcriptomes were extremely rich and exhibited great diversity. Lv (2012) and Yang (2013) analyzed the transcriptomes of nymphs and adults of *O. chinensis* and those of three insect strains of *Atractomorpha sinensis*, respectively. These studies enhanced the transcriptome database of orthopteran insects. Transcriptome

sequencing was used to investigate the resistance in insects, such as *L. migratoria*, the silverleaf whitefly (*Bemisia tabaci*), the brown planthopper (*Nilaparvata lugens*), and the greenhouse whitefly (*Trialeurodes vaporariorum*) (Wang 2009, Wang et al. 2010, Xue et al. 2010, Karatolos et al. 2012). These studies greatly enhanced the efficiency and quantity of gene annotation (Ansoerge. 2009) and improved research into insect resistance to insecticides. These findings initially resulted in the application of the transcriptome

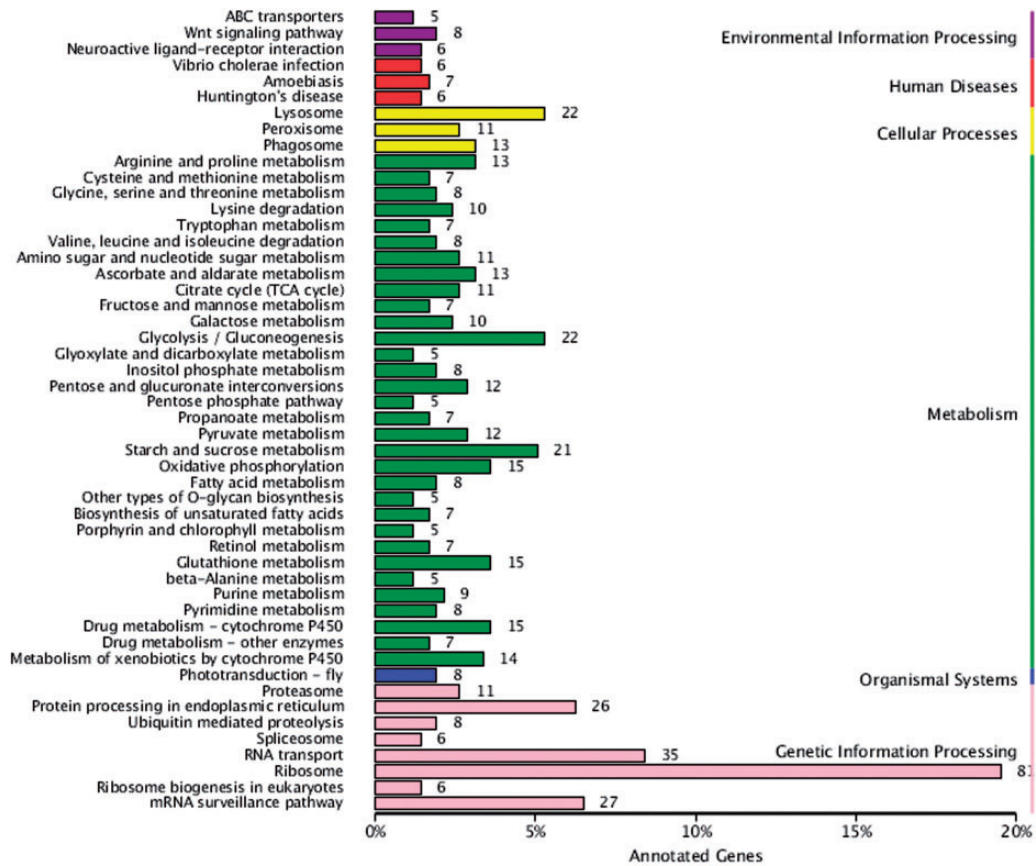


Fig. 4. KEGG categories of differentially expressed genes. Note: The vertical axis lists the names of the metabolic pathways in the KEGG, and the horizontal axis shows the proportion of annotated genes in each pathway versus the total number of annotated genes. The differentially expressed genes are distributed in 50 metabolic pathways, the pathway of metabolism holds the highest all number of genes.

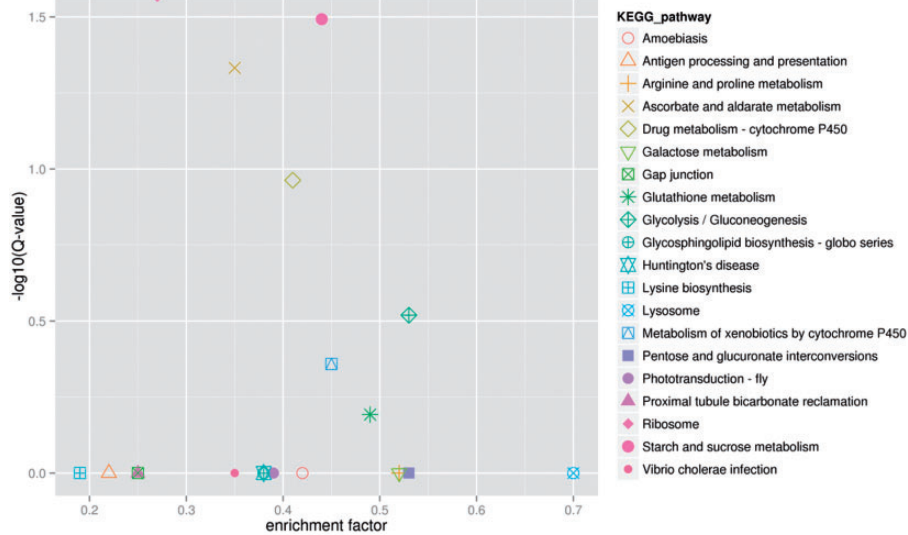


Fig. 5. Scatterplot of differentially expressed genes from the KEGG pathway enrichment analysis. Note: Every graph in the figure represents one KEGG pathway and the name of the pathway is shown on the right side of the legend. The horizontal axis shows the enrichment factor, which represents the proportion of the annotated differentially expressed genes in each pathway compared to the total annotated gene number in that pathway. The smaller enrichment factor indicates a higher enrichment score of differentially expressed genes in the pathway. The vertical axis displays an absolute Q value, which is the p-value tested by multiple hypothesis and calibrations. Thus, a larger value on the vertical axis indicates a more reliable enrichment score of differentially expressed genes in a pathway.

Table 5. Unique transcripts associated with insecticide resistance in *E. coerulipes*

Enzymes or target	Annotated unigenes number	Up-regulated number (PS-vs.-PR)
GSTs	43	5
CarEs	90	8
CYP450	183	10
GABA	10	–
AchR	15	–
RyR	3	–
AchE	38	–
CGSC	10	–

Note: GSTs, glutathione S-transferases; CarEs, carboxylesterase; CYP450, cytochrome P450; GABA, γ -aminobutyric acid receptor; nAchR, nicotinic acetylcholine receptor; RyR, ryanodine receptor; AchE, acetylcholinesterase; VGSC, voltage-gated sodium channel; Up-regulated number (PS-vs.-PR), GSTs: c8747.graph_c0, c34411.graph_c0, c33067.graph_c0, c40846.graph_c0, c31380.graph_c0; CarEs : c33871.graph_c0, c30386.graph_c0, c33452.graph_c0, c22727.graph_c0, c27991.graph_c0, c34612.graph_c0, c35381.graph_c0, c30559.graph_c0; CYP450: c23816.graph_c0, c37409.graph_c0, c29735.graph_c0, c34866.graph_c0, c34963.graph_c0, c34617.graph_c0, c39389.graph_c0, c37409.graph_c0, c22350.graph_c0, c39389.graph_c0.

technique to study *E. coerulipes* resistance to butane-fipronil. After butane-fipronil resistance selection, *E. coerulipes* can acquire 2,568 (1,178 are annotated) differentially expressed genes, which are distributed across 50 metabolic pathways; the largest category of these pathways was metabolic pathways. Additionally, it has been reported that the drug metabolism-cytochrome P450, metabolism of xenobiotics by cytochrome P450, and glutathione metabolism pathways are abundantly expressed and are closely linked to pesticide metabolism by insects, which provides a theoretical basis for tracking metabolic resistance in detoxifying enzymes.

Pesticide detoxification enzymes in insects, i.e., cytochrome P450s, CarEs, and GSTs, participate in the metabolism of allogenic materials, plant secondary substances, and pesticides (Strode et al. 2008, Wang 2011). A total of 316 unigenes among the *E. coerulipes* transcriptomes encode detoxifying enzymes. Different types and quantities of pesticide detoxification enzyme genes have been reported in previous studies based on the analysis of other insect transcriptomes, such as 35 P450s, 6 CarEs, and 7 GST genes in *B. tabaci*; 53 P450s, 49 CarEs, and 15 GST genes in *Dialeurodes citri*; and 102 P450s genes in *Cimex hemipterus* (Wang 2011, Mamidala et al. 2012, Chen et al. 2013). Furthermore, based on an analysis of insect genome sequencing results, 85, 106, 164, and 83 P450s genes have been found in *Drosophila melanogaster*, *Anopheles gambiae*, *Aedes aegypti*, and *Acyrtosiphon pisum* (Harris), respectively (Adams et al. 2000, Holt et al. 2002, Strode et al. 2008, Ramsey et al. 2010).

Many studies have established that pesticide resistance in insects can be achieved by the overexpression of various types of detoxification enzymes. For example, the activity of P450 mono-oxygenase was greatly increased in fipronil-resistant *Sogatella furcifera* (Jian et al. 2010). Fipronil resistance in *Bemisia tabaci* and *Laodelphax striatellus* is derived from increasing activity of p450 mono-oxygenase (Kang et al. 2006, Wang et al. 2010). Expression changes in the carboxylesterase genes *pxae22* and *pxae31* can influence the sensitivity of *Plutella xylostella* to fipronil (Ren et al. 2015). The fipronil resistance of *Chilo suppressalis* is related to the activity and expression levels of carboxylesterase and p450 mono-oxygenase (Li et al.

2007). Because butene-fipronil and fipronil are the same kind of insecticide, they may have the same resistance mechanism. The above analysis indicates that 23 newly discovered up-regulated detoxifying enzymes may contribute to the metabolism of butane-fipronil and show certain connections with the onset of resistance. Consequently, it is necessary to verify those genes and their expression levels to identify the exact relationship between *E. coerulipes* resistance to butene-fipronil and detoxifying enzymes in order to control resistance development by detoxifying enzyme inhibitor for the efficient use of chemical pesticides to achieve locust control.

The molecular targets of pesticides include the GABA receptor, nicotinic acetylcholine receptor, ryanodine receptor, AchE enzyme, and voltage-gated sodium channel (VGSC). These target proteins have been reported to be associated with pesticide resistance. It has also been reported that mutations in these genes encoding these target proteins lead to varying degrees of insensitivity to pesticides in insects. Mutations at M918V, L925I, and T929V in the VGSC gene have been associated with pyrethroid resistance (Chung et al. 2011). Single-site mutations in a conserved position, Y151S, in two nicotinic acetylcholine receptor subunit genes, Nla1 and Nla3, have been linked to imidacloprid resistance in the brown planthopper (Liu et al. 2005). The four mutated sites in the gene encoding the ryanodine receptor have been associated with chlorantraniliprole resistance in *Plutella xylostella* (Guo et al. 2014). Endosulfan resistance in *B. tabaci* (Houndete et al. 2010) has been associated with mutations in the GABA receptor subunit genes. The GABA receptor is also the molecular target of phenyl pyrazole pesticides (butene-fipronil and fipronil). Rdl gene mutation (Rdl1a→Rdl1s) of the *P. xylostella* GABA receptor is linked to their strong resistance to fipronil. The fipronil-resistant *Laodelphax striatellus* (Chen. 2012) may be related to A2' mutations N in the GABA receptor gene. Our study found that 76 unigenes encode target proteins of pesticides, and the annotation of genes that encode target proteins in this present study may represent the foundation for next-generation screening of mutation sites.

Transcriptome research on *E. coerulipes* has rapidly and effectively screened out resistance-related detoxifying enzymes and target site genes, in addition to providing valuable data for further study of the *E. coerulipes* resistance mechanism and control strategies.

Acknowledgments

This study was supported by a grant from the National Public Welfare Industry (Agriculture) Scientific Research Concentration (grant 201003079), the Key Laboratory of Cold Crop Germplasm Improvement and Cultivation in Heilongjiang Province of Heilongjiang Bayi Agricultural University, and the Insect and Pest Control Laboratory of the Plant Protection Institute of Shenyang Agricultural University.

References Cited

- Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287: 2185–2195.
- Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11: R106.
- Anson, W. J. 2009. Next-generation DNA sequencing techniques. *New Biotechnol.* 25: 195–203.
- Badisco, L., J. Huybrechts, G. Simonet, H. Verlinden, E. Marchal, R. Huybrechts, L. Schoofs, A. De Loof, and J. V. Broeck. 2011a. Transcriptome analysis of the desert locust central nervous system:

- Production and annotation of a schistocerca gregaria EST database. *PLoS ONE* 6: e17274.
- Badisco, L., S. R. Ott, S. M. Rogers, T. Matheson, D. Knapen, L. Vergauwen, H. Verlinden, E. Marchal, M.R.J. Sheehy, M. Burrows, et al. 2011b. Microarray-based transcriptomic analysis of differences between long-term gregarious and solitary desert locusts. *PLoS ONE* 6: e28110.
- Chen, S., P. C. Yang, F. Jiang, Y. Y. Wei, Z. Y. Ma, and L. Kang. 2010. De Novo analysis of transcriptome dynamics in the migratory locust during the development of phase traits. *PLoS ONE* 5: e15633.
- Chen, Y. 2012. The mechanisms of fipronil resistance in *Laodelphax striatell*. Master Dissertation, Nanjing Agricultural University, Nanjing, Jiangsu.
- Chen, E. H., D. D. Wei, G. M. Shen, G. R. Yuan, P. P. Bai, and J. J. Wang. 2013. De novo characterization of the *Dialeurodes citri* transcriptome: Mining genes involved in stress resistance and simple sequence repeats (SSRs) discovery. *Insect Mol. Biol.* 1–14.
- Chung, I. H., S. Kang, Y. R. Kim, J. H. Kim, J. W. Jung, S. Lee, S. H. Lee, and S. Y. Hwang. 2011. Development of a low-density DNA microarray for diagnosis of target-site mutations of pyrethroid and organophosphate resistance mutations in the whitefly *Bemisia tabaci*. *Pest Manag. Sci.* 67: 1541–1548.
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, et al. 2011. Full length transcriptome assembly from RNA Seq data without a reference genome. *Nature Biotechnol.* 29: 644–652.
- Guo, L., Y. Wang, X. Zhou, Z. Li, S. Liu, L. Liang Pei, and X. Gao. 2014. Functional analysis of a point mutation in the ryanodine receptor of *Plutella xylostella* (L.) associated with resistance to chlorantraniliprole. *Pest Manag. Sci.* 70: 1083–1089.
- Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protocol.* 8: 1494–512.
- Holt, R. A., G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab, and D. R. Nusskern. 2002. The genome sequence of the Malaria Mosquito *Anopheles gambiae*. *Science* 298: 129–149.
- Hong, J., Y. F. Ning, G. L. Du, and X. J. Yun. 2014. Current situation and cause analysis of grassland pests on grassland in china. *Pratacult. Sci.* 31: 1374–1379.
- Hou, R., Z. Bao, S. Wang, H. L. Su, Y. Li, H. X. Du, J. J. Hu, S. Wang, X. L. Hu. 2011. Transcriptome sequencing and de novo analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX. *PLoS ONE* 6: e21560.
- Houndete, T. A., G. K. Ketoh, O.S.A. Hema, T. Brevault, I. A. Glitho, and T. Martin. 2010. Insecticide resistance in field populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) in West Africa. *Pest Manag. Sci.* 66: 1181–1185.
- Jian, T., J. Li, Y. Shao, B. J. Yang, and Z. W. Liu. 2010. Fipronil resistance in the whitebacked planthopper (*Sogatella furcifera*): possible resistance mechanisms and cross-resistance. *Pest Manag. Sci.* 66: 121–125.
- Jiang, F., M. Yang, W. Guo, X. Wang, and L. Kang. 2012. Large-scale transcriptome analysis of retroelements in the Migratory Locust, *Locusta migratoria*. *PLoS ONE* 7: e40532.
- Kang, C. Y., G. Wu, and T. Miyata. 2006. Synergism of enzyme inhibitors and mechanisms of insecticide resistance in *Bemisia tabaci*. *J. Appl. Entomol.* 130: 377–385.
- Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10: R25.
- Leng, N., J. A. Dawson, J. A. Thomson, V. Ruotti, A. I. Rissman, B. M. Smits, J. D. Haag, M. N. Gould, R. M. Stewart, and C. Kendziorski. 2013. EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29: 1035–1043.
- Li, A. G., Y. H. Yang, S. Wu, C. Li, and Y. Wu. 2006. Investigation of resistance mechanisms to fipronil in diamondback moth. *J. Econ. Entomol.* 99: 914–919.
- Li, B., and N. D. Colin. 2011. RSEM: Accurate transcript quantification from RNA Seq data with or without a reference genome. *BMC Bioinformatics* 12: 323.
- Li, X., Q. Huang, J. Yuan, and Z. Tang. 2007. Fipronil resistance mechanisms in the rice stem borer. *Chilo suppressalis*. *Pestic. Biochem. Physiol.* 89: 169–174.
- Liu, Y. G., and Y. F. Liu. 2008. The locusts occurrence regularity and control measures in the Fuyu County. *Modern Agric. Sci. Technol.* 22: 142–143.
- Liu, Z. W., M. S. Williamson, S. J. Lansdell, I. Denholm, Z. J. Han, and N. S. Millar. 2005. A nicotinic acetylcholine receptor mutation conferring target-site resistance to imidacloprid in *Nilaparvata lugens* (brown planthopper). *Proc. Natl. Acad. Sci. USA.* 102: 8420–8425.
- Ly, H. J. 2012. Comparative transcriptome of larva and adult *Oxya chinensis* and mitochondrial transcriptome mapping. Master Dissertation, Shaanxi Normal University, Xian, Shanxi.
- Mamidala, P., A. J. Wijeratne, S. Wijeratne, K. Kornacker, B. Sudhamalla, L. J. Rivera-Vega, A. Hoelmer, T. Meulia, S. C. Jones, and O. Mittapalli. 2012. RNA-Seq and molecular docking reveal multi-level pesticide resistance in the bed bug. *BMC Genomics* 13: 6–16.
- Murtagh, F., and P. Legendre. 2014. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *J. Classif.* 31: 274–295.
- Ramsey, J. S., D. S. Rider, T. K. Walsh, M. De Vos, K.H.P. Gordon, L. Ponnala, S. L. Macmil, B. A. Roe, G. ans Jander. 2010. Comparative analysis of detoxification enzymes in *Acyrtosiphon pisum* and *Myzus persicae*. *Insect Mol. Biol.* 19: 155–164.
- Ren, N. N., M. Xie, Y. C. You, J.-Y. Li, W.-J. Chen, X.-M. Cheng, and M.-S. You. 2015. Fipronil-resistance mediated by carboxylesterases in the diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). *Acta Entomol. Sin.* 58: 288–296.
- Strode, C., C. S. Wondji, J. P. David, N. J. Hawkes, N. Lumjuan, D. R. Nelson, D. R. Drane, S.H.P. Parakrama Karunaratne, J. Hemingway, W. C. Black IV, et al. 2008. Genomic analysis of detoxification genes in the mosquito *Aedes aegypti*. *Insect Mol. Biol.* 38: 113–123.
- Tian, F. W., B. Z. Wang, Z. Y. Wang, and G. F. Wei. 2003. Investigation and analysis of locust species in alfalfa field binzhou city of Shandong. *Pratacult. Sci.* 120: 63–64.
- Tian, F. W., J. Z. Li, and Z. H. Wu. 2009. Preliminary study on relationship between occurrence of *Epacromius coeruleipes* (Ivan.) and environmental factors in north Shan dong. *J. Anhui Agric. Sci.* 37: 15895–15896.
- Tian, F. W., X. A. Zhang, Q. W. Wang, and S. H. Men. 2010. Preliminary research of damage and control index of *Epacromius coeruleipes*. *Serves Agric. Technol.* 2007: 1161–1162.
- Wang, J. 2009. Construction and analysis of cDNA library of locusta migratoria manilensis immune tissues and central nervous system. Master Dissertation, University of Chongqing, Chongqing, Sichuan.
- Wang, X. W., J. B. Luan, J. M. Li, Y. Y. Bao, C. X. Zhang, S. S. Liu, and X. W. Wang. 2010. De novo characterization of a whitefly transcriptome and analysis of its gene expression during development. *BMC Genomics* 11: 400–410.
- Wang, Y. Z., and Z. J. Han. 2010. Fipronil resistance in small brown planthopper *Laodelphax striatellus* (Fallen) and mechanisms. *J. Nanjing Agric. Univ.* 33: 49–54.
- Wang, X. W., X. Han, L. X. Zhang, J.-S. Fan, Y. Sun, S.-X. Wang, Y.-L. Zhou, Z.-H. Chen, K. Zhang, and G.-Q. Wang. 2015. Research on community dynamics of rangeland grasshopper in western Heilongjiang Province. *J. Eng. Heilongjiang Univ.* 6: 1–9.
- Wang, Z. Y. 2011. Monitoring and mechanisms of imidacloprid resistance in the whitefly *Bemisia tabaci*. PhD Dissertation, Nanjing Agricultural University, Nanjing, Jiangsu.
- Xie, C., X. Mao, J. Huang, Y. Ding, J. Wu, S. Dong, L. Kong, G. Gao, C. Y. Li, and L. Wei. 2011. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39: 316–322.
- Xu, Y. X., B. P. Pang, and J. X. Zhao. 2013. Comparison of growth and development of five grasshoppers in inner mongolia grasslands. *J. Inner Mongolia Agric. Univ.* 34: 76–79.
- Xue, J. A., Y. Y. Bao, B. L. Li, Y. B. Cheng, Z. Y. Peng, H. Liu, H. J. Xu, Z. R. Zhu, Y. G. Lou, J. A. Cheng, and C. X. Zhang. 2010. Transcriptome analysis of the brown planthopper *nilaparvata lugens*. *PLoS ONE* 5: e14233.

- Yang, J. 2013. Comparative transcriptome of *Atractomorpha sinensis* three stages and mitochondrial transcript mapping. Master Dissertation, Shaanxi Normal University, Xian, Shanxi.
- Ye, J., L. Fang, H. Zheng, Y. Zhang, J. Chen, Z. Zhang, J. Wang, S. Li, R. Li, L. Bolund, and J. Wang. 2006. WEGO: A web tool for plotting GO annotations. *Nucleic Acids Res.* 34: 293–297.
- Zhou, Y., G. Wang, and G. Li. 2010. Study on main Species of grasshoppers and their Integrated control in grassland of west Heilongjiang Province[J]. *Chin. Agric. Sci. Bull.* 27: 382–386.
- Zhou, Y. L., G. Q. Wang, and G. Z. Li. 2011. Study on main species of grasshoppers and their integrated control in grassland of west Heilongjiang Province. *Chin. Agric. Sci. Bull.* 27: 382–386.