Accepted Manuscript

Comparative antennal transcriptome of Apis cerana cerana from four developmental stages

Huiting Zhao, Zhu Peng, Yali Du, Kai Xu, Lina Guo, Shuang Yang, Weihua Ma, Yusuo Jiang

Please cite this article as: Huiting Zhao, Zhu Peng, Yali Du, Kai Xu, Lina Guo, Shuang Yang, Weihua Ma, Yusuo Jiang , Comparative antennal transcriptome of Apis cerana cerana from four developmental stages. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Gene(2017), doi:10.1016/j.gene.2018.03.068

Comparativeantennal transcriptome of *Apis cerana cerana* from four developmental stages

Huiting Zhao[a], Zhu Peng[a], Yali Du[b], Kai Xu[b], Lina Guo[b], Shuang Yang[b], Weihua Ma[c],

Yusuo Jiang[b*]

[a]College of Life Sciences, Shanxi Agricultural University, Taigu, Shanxi, China

[b]College of Animal Science and Veterinary Medicine, Shanxi Agricultural University, Taigu, Shanxi, China

[c]Institute of Horticulture, Shanxi Academy of Agricultural Science, Taiyuan, Shanxi 030031, China

*Correspondence: jiangys-001@163.com (Y. Jiang)

ABSTRACT

Apis cerana cerana, an important endemic honey bee species in China, possesses valuable characteristics such as a sensitive olfactory system, good foraging ability, and strong resistance to parasitic mites. Here, we performed transcriptome sequencing of the antenna, the major chemosensory organ of the bee, using an Illumina sequencer, to identify typical differentially expressed genes (DEGs) in adult worker bees of different ages, namely, T1(1 day); T2 (10 days); T3(15 days); and T4 (25 days). Surprisingly, the expression levels of DEGs changed significantly between the T1 period and the other three periods. All the DEGs were classified into 26 expression profiles by trend analysis. Selected trend clusters were analyzed, and valuable information on gene expression patterns was obtained. We found that the expression levels of genes encoding cuticle proteins declined after eclosion, while those of immunity-related genes increased. In addition, genes encoding venom proteins and major royal jelly proteins were enriched at the T2 stage; small heat shock proteins showed significantly higher expression at the T3 stage; and some metabolism-related genes were more highly expressed at the T4 stage. The DEGs identified in this study may serve as a valuable resource for the characterization of expression patterns of antennal genes in *A. cerana. cerana*. Furthermore, this study provides insights into the relationship between labor division in social bees and gene function.

Keywords: *Apis cerana cerana*; Antenna; Developmental stages; Transcriptome; Differentially expressed genes.

1. Introduction

Insects currently account for more than two-thirds of all animal species, and new insect species continue to be discovered. During evolution, insect species have developed several characteristics that have enabled them to adapt to a wide range of environments. One of these characteristics is the ability to detect low concentration levels of chemical compounds in the environment. The antenna is the main chemosensory organ of insects, which can perceive specific volatile organic compounds at levels far below the detection thresholds of current analytical devices (Schott et al., 2013). This exceptional sensing ability plays essential roles in locating hosts, foraging, searching for mates, aggregation, selecting oviposition sites, choosing habitats, and avoiding predators. Moreover, the characteristics of insect antennae have been emulated in biosensor applications, such as the fire detection or robotic olfactory searches (Schott et al., 2013).

Insect antennae have three parts, namely, the scape, pedicel, and flagellum (Nichols, 1989). The scape and pedicel each have a single segment, while the flagellum has a variable number of segments. Various types of sensory sensilla are distributed on the antenna surface (Ai et al., 2010). The number, morphology, type, and distribution of sensilla vary among insect species. Environmental volatile compounds are recognized and discriminated by different sensilla (Schneider, 1964; Kaissling, 1971).The antenna of the honeybee, as in all insects, is the main chemosensory structure. In honey bees, the flagellum has 10 segments in the female queen and workers, and 11 segments in drones. The morphologies of various sensilla types were first described in *Apis mellifera* by Slifer and Sekhon (1961).

Honey bees are social insects, female workers at different developmental stages show complex social behaviors such as feeding the brood and queen, cleaning the comb, foraging, and guarding the colony; many of these behaviors are mediated via pheromone communication using the antennae (Slessor et al., 2005; Le Conte and Hefetz, 2008). During the different phases of behavioral and physiological development in worker bees, many genes show differential

expression patterns. Although numerous studies have been carried out on gene expression in the past decade, they were largely restricted to specific genes, rather than undertaking large-scale testing, owing to limitations of available technologies. However, development of next-generation sequencing has effectively overcome this constraint, and large-scale identification of gene expression profiles in the antenna can now be easily accomplished.

In an earlier study, we have reported de novo transcriptome sequencing of antennal genes in the Chinese honeybee *Apis cerana cerana*. When analyzing the obtained sequencing data, we focused on five chemosensory gene families, among which109 olfactory-related genes were identified. Of these, 19 were differentially expressed genes (DEGs) at different developmental stages (Zhao et al., 2016). Following the publication of genome information for *A. cerana* (Park et al., 2015), we reanalyzed the antennal raw RNA sequencing (RNA-Seq) data based on the reference genome. In this study, all DEGs were identified and their expression characteristics were analyzed. This documentation can provide valuable information on molecular markers for the development and labor division of worker bees. Meanwhile, the study provides insights into functional analysis of DEGs, especially those that show high levels and similar patterns of expression.

2. Materials and methods

2.1 Insects and samples for RNA-Seq

Honey bees (*A. cerana. cerana*) were reared in hives under natural conditions. Frames with bee pupae close to eclosion were taken from two hives and incubated in an environmental chamber at a constant temperature of 33°C and 80% humidity. After eclosion, the workers were marked and returned to their hives until sampling. The antennae of approximately 200 worker bees from two hives were dissected at 1 (T1), 10 (T2), 15 (T3), and 25 (T4) days of age. Samples for the RNA-Seq experiment were prepared as described in a previous report (Zhao et al., 2016).

2.2 Mapping of reads to the reference genome

Raw reads of the eight samples were filtered by removing reads containing adaptors or those

that had more than 10% of unknown nucleotides. In addition, reads with more than 50% of low-quality bases (Q-value≤20) were removed. High-quality clean reads from each sample were mapped to the genome of the Asian honeybee (http://mnbldb.snu.ac.kr/genome_seq.php) usingTopHat2 (Kim et al., 2013).

2.3Analysis of DEGs

The gene expression level was normalized using the fragments per kilobase of transcript per million mapped reads (FPKM) method (Trapnell et al., 2010). The edgeR package (http://www.r-project.org/) was used to identify DEGs across samples. Genes with an absolute value of log2 (expression fold change) of ≥1 and a false discovery rate (FDR) of <0.05 in a comparison were identified as significant DEGs. DEGs were then subjected to trend and enrichment analyses for GO functions and KEGG pathways, with a Q-value of ≤0.05. The trend analysis was performed using the Short Time-series Expression Miner (STEM, version 1.2.2b) software (Ernst and Bar-Joseph, 2006).

2.4 Confirmation of RNA-Seq results by quantitative reverse transcription PCR

To confirm the results of the transcriptomic analysis, 10genes were chosen from the sequencing data and subjected to quantitative reverse transcription PCR (qPCR). *Arp1* was chosen as the internal reference gene. The qPCR was carried out on anMx3000P qPCR system (Stratagene, La Jolla, CA,USA) using a SYBR premix Ex Taq kit (TaKaRa). The cycling conditions were as follows: denaturation at 95°C for 20 s, followed by 40 cycles at 95°C for 15 s and 60°C for 20 s. Melting curve analysis was performed at 95°C for 20 s, 60°C for 30 s, and 95°C for 30 s to confirm the specificity of the PCR products. Each qPCR analysis was performed in triplicate. Relative gene expression levels were calculated using the comparative $2^{-\Delta\Delta Ct}$ method (Livak et al. 2001). The primers used for qPCR are listed in Table S1.

3. Results

3.1 Illumina sequencing and transcriptome assembly

To determine gene expression profiles in the antennae of worker bees at four different developmental stages, eight transcriptome libraries (two repeats for each stage) were constructed using Illumina HiSeq 2500. The stages (samples) were named T1-1, T1-2, T2-1, T2-2, T3-1, T3-2, T4-1, and T4-2, respectively. After filtering the raw reads, an average of 42 million clean reads were generated per sample. The major sequencing assembly information for the eight libraries is summarized in Table 1. In each library, over 81% of clean reads were mapped to the reference genome. In total, we detected 11,462 genes by sequencing analysis, of which 10,608(98.9%) were matched to the reference genome, and the other 854 were new genes. A correlation analysis of the two repeat samples (biological replicates) at each stage produced R^2 coefficients of >0.98, demonstrating the reliability and operational stability of the RNA-Seq results. The Illumina sequencing data have been submitted to the SRA of NCBI (accession number: SRR3180625).

Table 1. Major characteristics of total raw reads for each library generated by RNA-Seq.

Sample	Number of clean reads	Q30(%)	Genome mapping ratio (%)	Number of known genes	Number of new genes	Pearson correlation coefficient
T1-1	49262762	91.96	86.98	10256	846	0.9999
T1-2	42553352	92.57	87.28	10222	841	
T2-1	40276220	92.03	87.71	10197	841	0.9812
T2-2	46011244	91.93	81.72	10178	839	
T3-1	41243348	93.00	84.72	10223	840	0.9992
T3-2	41437034	93.47	87.85	10225	846	
T4-1	38971782	93.09	86.53	10224	848	0.9993
T4-2	36494586	92.99	87.08	10204	840	

T1-1, T1-2: biological replicates for 1-day-old workers.

T2-1, T2-2: biological replicates for 10-day-old workers.

T3-1, T3-2: biological replicates for of 15-day-old workers.

T4-1, T4-2: biological replicates for 25-day-old workers.

3.2 Identification of DEGs and trend analysis

The abundance of each gene mapped to the reference genome was quantified and normalized

using the FPKM method. On the basis of the gene expression values, a venn diagram was constructed (S1 Fig.). The diagram showed that most of the genes were generally expressed in all stages. Even though there were few specific genes in each period, the expression value was very low for each of these genes. This may indicate that no genes were uniquely expressed in any of the stages, while some DEGs were identified among the samples. Compared with the T1stage, 928, 1225, and 1155 DEGs were identified in the T2, T3, and T4 stage samples, respectively, and the number of downregulated genes was greater than those of upregulated genes. Compared with the T2 stage, 113 DEGs were upregulated and 59 were downregulated at the T3 stage, while 61 DEGs were upregulated and 79 were downregulated at the T4 stage. Comparison of the T3 and T4 samples resulted in the identification of 142 DEGs, and the numbers of upregulated and downregulated genes were similar (Fig.1).
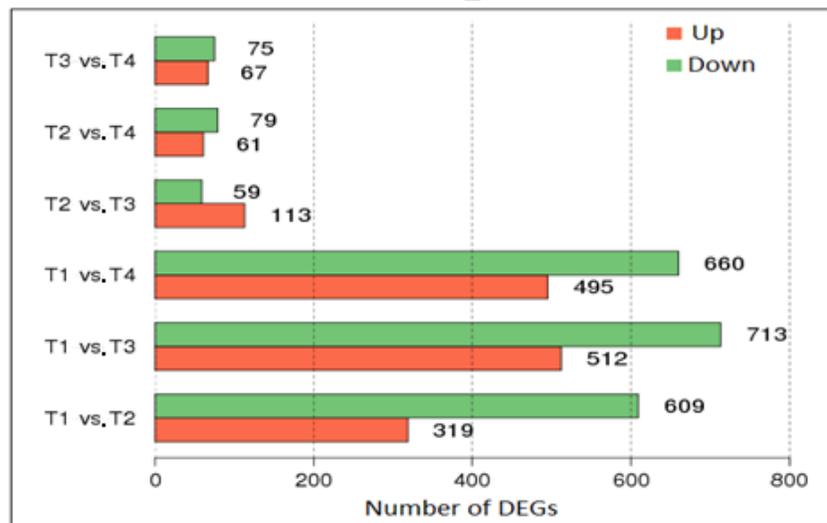

Fig. 1. Statistics of differentially expressed genes among samples.

Trend analysis and clustering of the DEGs identified in the samples from the four developmental stages were performed using the STEM software, and 26 profiles were obtained (Fig. 2). We then categorized these profiles and analyzed several typical profiles, namely, profiles 0, 4, 13, 14, 18, 19, 21, 22, and 25. The genes with the most significantly changed expression identified by this focused trend analysis are shown and annotated in Table 2.
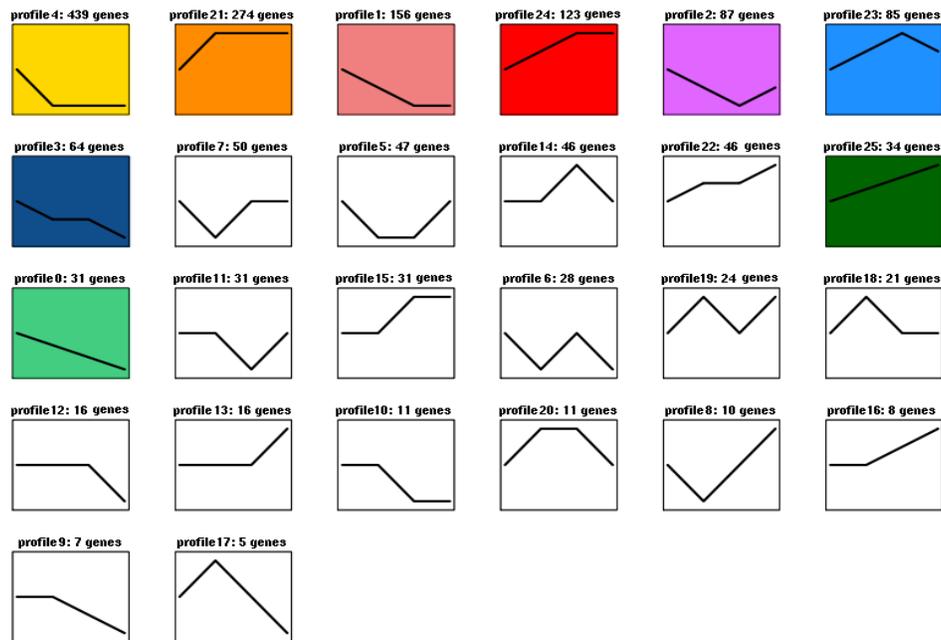
Fig. 2. Trend profiles generated by STEM across the four phases. These profiles are ordered by the numbers of genes assigned. Colored profiles indicate genes that gene expression were significantly enriched in this pattern, with a Q value of ≤0.05. The number of genes belonging to each pattern is indicated above the profile.

Profiles 4 and 21 had the largest numbers of DEGs (439 and 274, respectively). Profile 4 included genes showing relatively higher expression in T1 than in the other three stages, which had no obvious variation in the expression level in these stages. In this profile, the most significantly changed DEGs encoded structural proteins, such as cuticle protein, keratin, and chitin deacetylase. By contrast, in profile21, the expression levels of DEGs were relatively lower in the T1 stage than in the other three stages, and genes encoding immunity-related proteins, such as cytochrome P450, defensin, and trypsin inhibitor, were found in this profile.

Profile 0 included genes that showed decreased expression with as the age increased, while profile 25 included genes expressed at increasing levels with the increase of the age. Genes showing the greatest changes in expression included those for an RNA-binding protein, the pupal

cuticle protein, and blue-sensitive opsin in profile 0, as well as that for the organic cation transporter protein in profile 25.

The genes included in profiles18 and 19 showed relatively higher levels of expression at the T2 stage; the genes showing the greatest change in expression were those encoding venom-associated proteins, e.g., melittin, apamin, secapin, phospholipase A2, and venom carboxylesterase 6. In addition, genes encoding major royal jelly protein (MRJP) family members, MRJP1, 2, 3, 4, 5, 7,and 8, were also included.

Profile 14 contained genes whose expression levels were relatively higher at the T3 stage. The greatest changed genes was observed for encoding a small heat shock protein (sHSP) and pheromone-processing carboxypeptidase.

Profiles 13 and 22 included genes enriched at the T4 stage. These genes mostly participated in metabolic processes, such as an ATP-binding cassette sub-family E member, small subunit processome component, glucose dehydrogenase, acyl-CoA Delta (11) desaturase, glycine-rich RNA-binding protein 3, chondroitin proteoglycan 2, and chorion peroxidase.

Table 2. Expression and annotation information for DEGs changed significantly in eight selected profiles.

Gene ID	Profile	FPKM values				Gene description
		T1	T2	T3	T4	
ncbi_107993596	0	134.87	22.96	1.23	0.05	RNA-binding protein 12-like [*Apis dorsata*]
ncbi_108001081	0	1.56	0.27	0.15	0.06	pupal cuticle protein C1B-like [*Apis dorsata*]
ncbi_107999889	0	1.42	0.56	0.24	0.16	blue sensitive opsin [*Apis cerana*]
ncbi_107992678	4	637.11	4.75	6.75	6.36	larval/pupal cuticle protein H1C-like [*Linepithema humile*]
ncbi_107993465	4	110.82	0.36	0.10	0.03	glycine-rich RNA-binding protein 7-like [*Apis cerana*]
ncbi_107994956	4	16.04	0.39	0.18	0.28	cuticular protein precursor [*Apis mellifera*]
ncbi_107995141	4	90.79	5.92	4.48	5.22	low-density lipoprotein receptor-related protein, partial [*Cerapachys biroi*]
ncbi_107995161	4	104.07	14.42	7.55	8.27	chitin deacetylase 1 precursor [*Lasius niger*]
ncbi_107995288	4	72.64	0.68	0.65	0.41	cuticle protein 19-like isoform X1 [*Apis dorsata*]
ncbi_107995724	4	29.85	0.48	0.27	0.53	pupal cuticle protein C1B-like [*Pogonomyrmex barbatus*]
ncbi_107996964	4	48.95	4.14	2.94	2.86	endocuticle structural glycoprotein SgAbd-8-like isoform X1 [*Apis mellifera*]
ncbi_107998402	4	2451.34	193.93	111.97	113.08	keratin-associated protein 19-2-like isoform X2 [*Apis dorsata*]
ncbi_107999058	4	29.44	4.35	2.84	2.75	cuticular protein 27 precursor [*Apis mellifera*]
ncbi_107999765	4	401.65	62.19	36.49	30.32	collagen alpha-5(IV) chain [*Apis mellifera*]
ncbi_107999808	4	5.87	0.13	0.12	0.09	zinc finger protein 512B-like [*Apis mellifera*]
ncbi_108001920	4	158.62	4.53	2.04	2.33	cuticular protein analogous to peritrophins 3-C precursor [*Apis mellifera*]
ncbi_108002591	4	22.01	0.71	0.63	0.57	cuticle protein 18.7-like isoform X1 [*Apis dorsata*]
ncbi_108002599	4	8.37	0.55	0.46	0.30	pupal cuticle protein-like [*Fopius arisanus*]
ncbi_108003544	4	84.44	12.63	7.70	5.90	protein Skeletor, isoforms B/C isoform X1 [*Apis florea*]
ncbi_108003841	4	2000.38	223.86	166.16	97.09	keratin-associated protein 19-2-like [*Apis cerana*]
ncbi_108003854	4	18.49	1.84	1.43	0.81	cadherin-23-like [*Apis mellifera*]
ncbi_107994630	13	0.001	0.001	0.001	0.93	ATP-binding cassette sub-family E member 1 [*Apis florea*]
ncbi_108003158	13	0.001	0.001	0.001	0.05	small subunit processome component 20 homolog, partial [*Apis dorsata*]
XLOC_002064	13	0.14	0.27	0.125	0.64	keratin, type I cytoskeletal 9-like [*Apis florea*]

ncbi_107994729	13	2.18	3.27	3.08	10.75	glucose dehydrogenase [FAD, quinone] isoform X1 [*Apis mellifera*]
ncbi_107994022	14	16.47	20.57	634.28	22.37	small heat shock protein 24.2a (sHSP24.2a) gene [*Apis cerana cerana*]
ncbi_107996023	14	78.68	162.53	2061.81	181.53	small heat shock protein 21.7a (sHSP21.7a) gene [*Apis cerana cerana*]
ncbi_107996026	14	330.37	518.61	3726.35	440.66	small heat shock protein 22.6 (sHSP22.6) gene [*Apis cerana cerana*]
ncbi_107996027	14	133.32	272.45	3720.97	221.64	small heat shock protein 23 (sHSP23) gene [*Apis cerana cerana*]
ncbi_107998816	14	1.62	3.02	9.73	0.84	pheromone-processing carboxypeptidase KEX1-like isoform X3 [*Apis florea*]
ncbi_108000807	18	17.74	10925.46	24.19	6.50	RecName: Full=Melittin; Flags: Precursor [*Vespa magnifica*]
ncbi_107998818	18	0.89	433.39	1.17	0.34	phospholipase A2 precursor [*Apis mellifera*]
ncbi_107993904	18	0.09	49.78	0.18	0.19	RecName: Full=Apamin; Flags: Precursor [*Apis cerana cerana*]
ncbi_108001335	18	0.37	21.54	0.125	0.14	RecName: Full=Secapin; Flags: Precursor [*Vespula maculifrons*]
ncbi_107998239	18	0.001	0.74	0.02	0.03	venom carboxylesterase-6 precursor [*Apis mellifera*]
ncbi_107994025	19	0.67	420.79	6.12	112.97	major royal jelly protein 2 [*Apis cerana cerana*]
ncbi_107997158	19	1.63	1228.05	16.03	298.86	major royal jelly protein 1 [*Apis cerana*]
ncbi_107997171	19	2.32	552.93	8.71	142.30	major royal jelly protein 3 [*Apis cerana*]
ncbi_107997172	19	2.06	1194.82	18.4	253.71	major royal jelly protein 4 [*Apis cerana*]
ncbi_107997173	19	1.25	87.20	4.17	17.64	major royal jelly protein 7 [*Apis mellifera*]
ncbi_107997178	19	0.13	81.91	1.81	16.38	major royal jelly protein 5 [*Apis cerana*]
ncbi_107997822	19	3.93	181.92	8.73	56.1	glucose oxidase [*Apis mellifera*]
ncbi_108003250	19	4.82	1507.53	18.10	354.44	apisimin [*Apis cerana cerana*]
ncbi_107993198	21	1.75	20.96	14.93	33.45	aminopeptidase N-like [Apis dorsata]
ncbi_107993803	21	8.92	100.43	56.87	140.47	defensin precursor [*Apis cerana cerana*]
ncbi_107995017	21	49.99	452.17	528.49	570.66	dehydrogenase/reductase SDR family member 11-like [*Apis dorsata*]
ncbi_107996362	21	0.56	97.52	11.89	30.24	trypsin inhibitor-like [*Apis dorsata*]
ncbi_107999588	21	1.69	13.96	17.07	23.25	fibrillin-1-like [*Apis mellifera*]
ncbi_108002368	21	0.36	17.7	18.62	22.52	pancreatic triacylglycerol lipase-like isoform X2 [*Apis mellifera*]
ncbi_108004123	21	4.80	68.21	46.29	65.60	cytochrome P450 9e2 [*Apis mellifera*]
XLOC_002595	21	1.16	10.63	6.56	5.47	major royal jelly protein 8 precursor [*Apis mellifera*]
ncbi_108000067	22	0.12	2.67	3.60	15.11	acyl-CoA Delta(11) desaturase-like [*Apis florea*]

ncbi_107993442	22	0.60	1.52	1.61	2.64	glycine-rich RNA-binding protein 3, mitochondrial-like [*Apis dorsata*]
ncbi_107995250	22	5.975	17.75	17.88	27.16	venom acid phosphatase Acph-1-like [*Apis dorsata*]
ncbi_107996385	22	5.08	16.995	12.235	21.65	chondroitin proteoglycan 2-like [*Apis florea*]
ncbi_107997475	22	0.29	1.20	0.77	1.54	genetic suppressor element 1-like [*Apis florea*]
ncbi_107998608	22	0.53	1.51	0.93	2.19	chorion peroxidase isoform X2 [*Apis florea*]
ncbi_107995432	25	16.49	50.47	95.7	225.8	organic cation transporter protein-like isoform 1 [*Apis mellifera*]
ncbi_107992806	25	0.81	1.51	2.74	3.78	myosin-9 isoform X1 [*Athalia rosae*]
ncbi_108000543	25	0.83	1.28	2.11	3.46	gustatory receptor 68a-like [*Monomorium pharaonis*]

3.3 *validation by quantitative real-time PCR*

To validate the accuracy and reproducibility of the transcriptome data, 10 DEGs were selected and quantified using qRT-PCR. The results showed that the expression patterns of nine of the 10 selected genes were consistent with those obtained by RNA-Seq analysis (Fig. 3). Thus, the qPCR results confirmed that the RNA-Seq data were reliable.
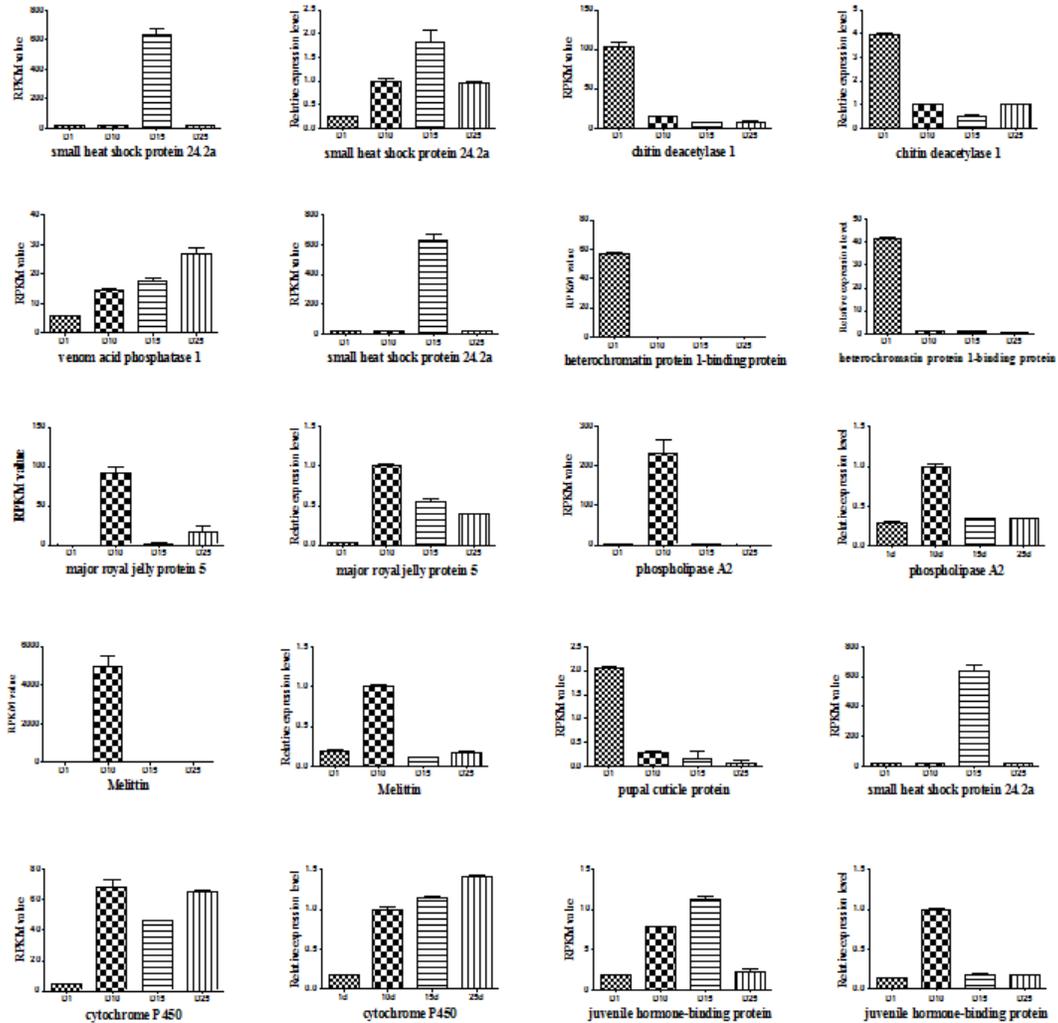

Fig. 3. Comparison of the FPKM values from RNA-Seq (right) and qPCR results (left) for 10 DEGs.

3.4 GO-term and KEGG pathway enrichment analysis of DEGs

The GO functional enrichment analysis showed that DEGs in all comparison groups fell into

three groups: (ⅰ)"cellular component," with 317 genes; (ⅱ) "biological process," with 566 genes; and (ⅲ)"molecular function," with 681 genes. Of all the GO terms, 59 showed particularly high enrichment, with Q values≤0.05. Among these, 32 terms were for molecular function, 23 for biological process, and four for cellular component (Table S2). In the biological process terms, the upregulated genes were significantly associated with negative regulation of cellular communication, signal transduction, proteolysis, and metabolic processes, while the downregulated genes mainly participated in aminoglycan metabolic processes and carbohydrate derivative metabolic processes. In the molecular function terms, the upregulated genes were associated with enzyme inhibitor activity, as well as with phospholipase and lipase activity; the downregulated genes were mainly associated with transmembrane and neurotransmitter transporter activity.

The pathway enrichment analysis revealed 270 DEGs could be annotated, and these were assigned to 125 KEGG pathways. We filtered out the pathways that were significantly enriched and had Q-values of ≤ 0.05. As a result, 26 of the 125 KEGG pathways were selected; these are listed in Table S3, and a scatter plot for these enriched pathway is shown in S2 Fig.

When comparing T1 with the other three stages, the significantly enriched pathways included those for neuroactive ligand–receptor interaction; tricarboxylic acid cycle and lysosome; extracellular matrix receptor interaction; retinol metabolism; the longevity-regulating pathway; and carbon, pyruvate, drug, tyrosine, starch, and sucrose metabolism pathways. When comparing T2 versus the T3 stage, the significantly enriched pathways involved protein processing in the endoplasmic reticulum and the longevity-regulating pathway. When comparing T2 versus the T4 stage, the AGE-RAGE signaling pathway in diabetic complications was significantly enriched. When comparing T3 versus theT4 stage, the significantly enriched pathways were mainly linked to protein processing in the endoplasmic reticulum; the longevity-regulating pathway, and glycine, serine, and threonine metabolism.

4. Discussion

The antennae of insects are essential for sensing diverse odors and for transmitting sensory information, which is integrated in the brain, and ultimately induces behavioral responses. Honey bees have a pair of geniculate antennae arising anterior to the eyes. After emergence from brood cells, the soft body of an adult worker becomes stronger, enabling it to carry out different tasks at different ages. At the same time, variations in gene expression and internal physiological functions also cause morphological changes. In this study, we analyzed antennal transcriptomes by RNA-Seq at four different developmental phases in worker bees. Quality assessment of the assembly results showed that clean reads were mapped well to the reference genome, and there was a high correlation between replicate samples. The high quality of the data enabled us to perform further analyses of antenna-related genes.

Regardless of differences in the developmental stage, tissue, sex, species, or treatment, identification of DEGs in an RNA-Seq experiment can yield insights into the mechanisms of physiological and behavioral changes. In this study, we a particular feature identified via analysis of DEGs was that the T1 stage had a larger number of DEGs than the other three stages did (e.g., profiles 4 and 21). This feature might be attributable to dramatic morphological and physiological changes during eclosion transition from the dormant pupal stage to the active adult. After then, changes were relatively steady as the insects grew and most genes were seemingly expressed stably.

Of the DEGs found in profiles 4 and 21, the most significantly downregulated genes were those encoding cuticle proteins. The insect cuticle serves as a barrier to the external environment and is made of many cuticle proteins. During growth and differentiation from the larval to adult stages, the exoskeleton or cuticle periodically degrades and regenerates (Charles, 2010). Some studies have reported that insect cuticle protein genes are expressed at higher levels in the pupal phase (Togawa et al., 2007; Soares et al., 2011; Jasrapuria et al., 2012). Since we did not collect samples at the pupal stage, the result showing higher expression of the genes involved in the cuticle protein synthesis in the newly emerged bees cannot be easily compared with those of the

previous reports.

By contrast, the most significantly upregulated genes were those for immunity-related proteins, such as cytochrome P450, defensin, and trypsin inhibitor.

In insects, cytochrome P450 isoforms play an important role in detoxifying xenobiotic compounds such as insecticides and plant toxins (Schuler, 2011; Feyereisen, 2012). They are also important for the biosynthesis and degradation pathways of endogenous compounds such as pheromones, hydroxyecdysone, and juvenile hormone (Reed et al., 1994; Niwa et al., 2004; Liu et al., 2015), which perform important functions in insect growth and development. An earlier study has reported that cytochrome P450 gene expression was low in the antennae of in newly emerged workers, relatively high in nurses, and the highest in foragers (Mao et al., 2015). Our current findings are consistent with the data of this report. Insect defensins are a class of small, conserved antimicrobial peptides with known functions after bacterial infection (Blandin et al., 2002). Trypsin inhibitor can inhibit proteinase activity, and thus plays important roles in various biological processes, such as immune responses and anticoagulation (Bhattacharyya et al., 2007; Lingaraju and Gowda, 2008).The body of a newly emerged bee is not fully developed, and its immune system is imperfect. Upregulation of the genes encoding immunity-related proteins might be beneficial because of their detoxification and anti-infection abilities.

Honey bees are eusocial insects, demonstrating division of labor, which varies depending on the insect age. Generally, young workers perform tasks such as nursing and cleaning within the colony during the first 1 to 3 weeks; subsequently, they leave the nest to forage for 1 to 2 weeks before dying (Winston, 1987; Seeley, 1995). Juvenile nurse bees(6 - 12 days old) secrete royal jelly from glands in the hypopharynx, which is fed to the queen and larvae (Kubo et al., 1996). Later, after the development of a wax gland, the hypopharyngeal gland degrades. The MRJP family consists of nine members(MRJP1 to 9), each with a function in nutrition (Albertová et al., 2005) or even beyond nutrition (Corby-Harris, 2016). In this study, we identified seven MRJPs in the antennal transcriptome; the missing proteins were MRJP6 and MRJP9. Of the seven *MRJP*

genes, six showed similar expression patterns and were most abundant in the T2 stage (10 days old). They were also slightly elevated in the T4 stage (25 days old), but the expression levels were low in the T1 and T3 stages. Of the six genes, expression of *MRJP1* and *MRJP4* was particularly enriched in the T2 stage (FPKM value > 1000). This expression pattern is consistent with the feeding behavior of worker bees. It is noteworthy that the expression of *MRJP8* (profile 21) was markedly lower than that of the other six genes (profile19) and was not enhanced at the T2 stage. Previous studies have reported that the expression of *MRJP1*, *2*, *3*, *4*, *7*, and *8* was elevated in the head or hypopharyngeal gland of nursing worker bees at approximately 9 days of age but was low in foragers (Liu et al., 2015) or in other adult stages (Li et al., 2016). Our findings here are similar to these earlier reports, except for the case of forager stage, which may suggest that *MRJP* genes might have other biological functions in addition to nutrition.

The honey bee venom is synthesized in the venom glands of workers and queens. The venom is a complex mixture of proteins, peptides, and low-molecular-weight components (Sliva et al., 2017). Among these constituents, the phospholipase A2 and hyaluronidase enzymes are the main proteins, and melittin is the most abundant peptide, comprising approximately 50% of the dry weight of bee venom (Park et al., 2011). To the best of our knowledge, research on bee venom has mainly focused on medicinal properties, such as anti-autoimmune disease (Lee et al., 2016), anti-inflammatory (Terra et al., 2007), and blood pressure-lowering effects, as well as the activation of the immune system (Arruda et al., 2007; Mingomataj et al., 2014). There are comparatively few reports on the expression patterns of genes of bee venom. It is worth noting that the bee venom genes in our study such as melittin, apamin, secapin, phospholipase A2, and carboxylesterase-6, showed significantly high expression levels at the T2 stage (profile18). In addition, the FPKM levels of melittin and phospholipase A2 were comparatively higher. This particular pattern of expression of venom-related genes might be associated with the defensive duties of worker bees. In view of the potential applications and benefits of bee venom, it will be worthwhile to investigate these proteins to possibly upscale the venom production in the future.

HSPs are known stress proteins that help insects cope with stressors such as heat, cold, crowding, and anoxia (King and MacRae, 2015; Huang, et al., 2017). We found here that the expression level of some *sHSP* genes markedly increased in 15-day-old bees (T3 stage, profile 14). This effect was likely induced by temperature stimuli. Since the bees for our analysis were sampled at the time when ample nectar and pollen resources were available outside, we speculate that the transition from in-hive activities to foraging activities may have stimulated the upregulation of the *sHSP* genes. An increase in the *sHSP* gene expression may provide a self-defense mechanism for adaptation to the external environment. The energy body demands increase as the labor intensity of foraging increase; therefore, expression of the genes related to energy metabolism might have increased accordingly at a later stage (such as the T4 stage).

5. Conclusions

By analyzing the RNA-Seq data for the antennae of honey bees, we detected 11,462 genes, of which 10,608 were matched to the reference genome. Our results indicated that the gene expression levels considerably differed between the T1 stage and the other developmental stages (T2, T3, and T4). Overall, more downregulated DEGs were detected. The number of DEGs between the T2, T3, and T4 stages was limited. Unexpectedly, the expression level of genes for MRJPs and the venom were enriched in the antennae of worker bees. This effect might be associated with the worker bee role in feeding larvae and in defending the hive against invaders. Furthermore, 59 functional GO terms and 26 KEGG pathways were significantly enriched with Q-values of ≤0.05. This study provides a better understanding of the gene expression characteristics of the antennae of *A. cerana. cerana*.

Acknowledgements

Conflicts of interest

None.

Captions

Fig. S1. Venn diagrams of the genes expressed at four different developmental stages.

Fig. S2. Top 20 of pathway enriched among samples.

Table S1. Primers used for qRT-PCR analysis of DEGs.

Table S2. GO classification of the enriched DEGs associated with clusters.

Table S3. KEGG pathway of the enriched DEGs among samples.

References

Schott, M., Wehrenfennig, C., Gasch, T., Vilcinskas, A., 2013. Insect antenna-based biosensors

for in situ detection of volatiles.Adv. Biochem. Eng. Biotechnol. 136, 101-122. doi:

10.1007/10_2013_210.

Martinez, D., Arhidi, L., Demondion, E., Masson, J.B., Lucas, P., 2014.

Using insect electroantennogram sensors on autonomous robots for olfactory searches. J. Vis.

Exp. 90, e51704. doi: 10.3791/51704.

Ai, H., Yoshida, A., Yokohari, F., 2010. Vibration receptive sensilla on the wing margins of the

silkworm moth Bombyx mori.J. Insect. Physiol. 56, 236-46. doi: 10.1016/j.jinsphys.2009.10.007.

Schneider, D., 1964. Insect antennae. Ann. Rev. Ent. 9, 103-122.

Kaissling, K.E., 1971. Insect olfaction. In: Beidler LM, editor. Handbook of sensory physiology

IV Chemical senses Part 1. Berlin: Springer-Verlag. 351-431pp.

Bawin, T., Collard, F., De Backer, L., Yarou, B.B., Compère, P., Francis, F., Verheggen, F.J., 2017.

Structure and distribution of the sensilla on the antennae of Tuta absoluta (Lepidoptera:

Gelechiidae). Micron. 96, 16-28. doi: 10.1016/j.micron.2017.01.008.

Slifer, E.H., Sekhon, S.S., 1961. Fine structure of the sense organs on the flagellum of the honey

bee, *Apis mellifera* Linnaeus. J. morphol.109, 351-381.

Slessor, K.N., Winston, M.L., Le Conte, Y., 2005. Pheromone communication in the honeybee (Apis mellifera L.).J. Chem. Ecol., 31, 2731-2745.

Le Conte., Y. and Hefetz, A., 2008. Primer pheromones in social hymenoptera. Annu. Rev. Entomol. 53, 523-542.

Zhao, H., Du, Y., Gao, P., Wang, S., Pan, J., Jiang, Y., 2016. Antennal Transcriptome and Differential Expression Analysis of Five Chemosensory Gene Families from the Asian Honeybee *Apis cerana cerana*. Plos One. 11, e0165374. doi: 10.1371/journal.pone.0165374. eCollection 2016.

Park, D., Jung, J.W., Choi, B.S., Jayakodi, M., Lee, J., Lim, J., Yu, Y., Choi, Y.S. et al., 2015. Uncovering the novel characteristics of Asian honey bee, *Apis cerana*, by whole genome sequencing. BMC genomics. 16, 1. Epub 2015/01/03. doi: 10.1186/1471-2164-16-1.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biol.14, R36. doi: 10.1186/gb-2013-14-4-r36.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg S.L., Wold, B.J, Pachter L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28, 511-515. doi: 10.1038/nbt.1621.

Ernst, J. and Bar-Joseph, Z., 2006. STEM, a tool for the analysis of short time series gene expression data. BMC Bioinformatics. 7, 191. doi: 10.1186/1471-2105-7-191.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T) Method. Methods. 25, 402-408. doi: 10.1006/meth.2001.1262.

Charles, J.P., 2010. The regulation of expression of insect cuticle protein genes. Insect Biochem. Mol. Biol.40, 205-213. doi: 10.1016/j.ibmb.2009.12.005.

Togawa, T., Dunn, W.A., Emmons, A.C., Willis, J.H., 2007. CPF and CPFL, two related gene families encoding cuticular proteins of Anopheles gambiae and other insects. Insect. Biochem. Mol. Biol. 37, 675-688.

Soares, M.P., Silva-Torres, F.A., Elias-Neto, M., Nunes, F.M., Simões, Z.L., Bitondi, M.M., 2011. Ecdysteroid-dependent expression of the tweedle and peroxidase genes during adult cuticle formation in the honey bee, *Apis mellifera*. PLoS One. 6, e20513. doi: 10.1371/journal.pone.0020513.

Jasrapuria, S., Specht, C.A., Kramer, K.J., Beeman, R.W., Muthukrishnan, S., 2012. Gene families of cuticular proteins analogous to peritrophins (CPAPs) in Tribolium castaneum have diverse functions. PLoS One. 11,e49844.doi: 10.1371/journal.pone.0049844.

Feyereisen, R., 2012. Insect CYP genes and P450 enzymes, in: LI. Gilbert (Ed.). Insect Molecular Biology and Biochemistry, Elsevier Academic Press, Oxford, 236-316 pp.

Schuler, M.A., 2011. P450s in plant-in sect interactions. Biochem. Biophys. Acta. 1814, 36-45.

Reed, J.R., Vanderwel, D., Choi, S., Pomonis, J.G., Reitz, R.C., Blomqusit, G.J., 1994. Unusual mechanism of hydrocarbon formation in the housefly:cytochrome P450 converts aldehyde to the sex pheromone component (Z)-9-tricosene and CO_2, Proc. Natl. Acad. Sci. U.S.A. 91, 10000-10004.

Niwa, R., Matsuda, T., Yoshiyama, T., Namiki, T., Mita, K., Fujimoto, Y., et al. CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid biosynthesis in the prothoracic glands of Bombyx and Drosophila. J. Biol. Chem. 279, 35942-35949.

Liu, N., Li, M., Gong, Y., Liu, F., Li, T., 2015. Cytochrome P450s--Their expression, regulation, and role in insecticide resistance. Pestic. Biochem. Phys. 120, 77-81.

Blandin, S., Moita, L.F., Köcher, T., Wilm, M., Kafatos, F., and Levashina, E.A., 2002. Reverse genetics in the mosquito Anopheles gambiae: targeted disruption of the defensin gene. EMBO Rep. 3, 852-856. doi: 10.1093/embo-reports/kvf180.

Bhattacharyya, A., Mazumdar Leighton S., Babu, C.R., 2007. Bioinsecticidal activity of Archidendron ellipticum trypsin inhibitor on growth and serine digestive enzymes during larval development of Spodoptera litura.Comp. Biochem. Physiol. C. Toxicol. Pharmacol. 145, 669-677. doi: 10.1016/j.cbpc.2007.03.003.

Lingaraju, M.H., Gowda, L.R., 2008.A Kunitz trypsin inhibitor of Entada scandens seeds: another member with single disulfide bridge.Biochim. Biophys. Acta. 1784, 850-855. doi: 10.1016/j.bbapap.2008.02.013.

Mao, W., Schuler, M.A., Berenbaum, M.R., 2015.Task-related differential expression of four cytochrome P450 genes in honeybee appendages.Insect Mol. Biol.24, 582-588. doi: 10.1111/imb.12183.

Winston, M.L., 1987. The biology of the honey bee. Harvard University Press Cambridge, London.

Kubo, T., Sasaki, M., Nakamura, J., Sasagawa, H., Ohashi, K., Takeuchi, H., Natori, S., 1996.Change in the expression of hypopharyngeal-gland proteins of the worker honeybees (Apis mellifera L.) with age and/or role.J. Biochem. 119, 291-295.

Buttstedt, A., Moritz, R.F., Erler, S., 2014. Origin and function of the major royal jelly proteins of the honeybee (Apis mellifera) as members of the yellow gene family. Biol. Rev. Camb. Philos. Soc. 89, 255-269. doi: 10.1111/brv.12052.

Albertová, V., Su, S., Brockmann, A., Gadau, J., Albert, S., 2005. Organization and potential function of the mrjp3 locus in four honeybee species. J. Agric. Food. Chem. 20, 8075-8081.

Corby-Harris, V., Meador, C.A., Snyder, L.A., Schwan, M.R., Maes, P., Jones, B.M., Walton, A., Anderson, K.E., 2016. Transcriptional, translational, and physiological signatures of undernourished honey bees (Apis mellifera) suggest a role for hormonal factors in hypopharyngeal gland degradation.J. Insect. Physiol., 85, 65-75. doi: 10.1016/j.jinsphys.2015.11.016.

Liu, F., Zong, C., Yu, L.S., Su S.K., 2015. Analysis of differentially expressed genes associated

with the behavioral transition between nurses and foragers in Apis mellifera ligustica, Chinese J. Appl. Entomol. 52, 300-307.

Li, X.Y., Liang, Q.H., Yang, W.J., Liang, Q., Li, J.H., 2016. Expression of the major royal jelly protein 7 gene in honeybee (*Apis mellifera* ligustica Spinola). J. Environ. Entomol. 20, 1674-0858. doi: 10. 3969 /j. issn.

Silva GBD, Junior., Vasconcelos AG, Junior., Rocha, A.M.T., Vasconcelos, V.R., Barros J, Neto., Fujishima, J.S., Ferreira, N.B., Barros, E.J.G., Daher, E.F., 2017. Acute kidney injury complicating bee stings - a review. Rev Inst Med Trop Sao Paulo. 59, e25. doi: 10.1590/S1678-9946201759025.

Park, J.H., Kum, Y.S., Lee, T.I., Kim, S.J., Lee, W.R., Kim, B.I., Kim, H.S., Kim, K.H., Park, K.K., 2011.Melittin attenuates liver injury in thioacetamide-treated mice through modulating inflammation and fibrogenesis. Exp. Biol. Med (Maywood). 236, 1306-1313. doi: 10.1258/ebm.2011.011127.

Lee, M.J., Jang, M., Choi, J., Lee, G., Min, H.J., Chung, W.S., Kim, J.I., Jee, Y., Chae, Y., Kim, S.H., Lee, S.J., Cho, I.H., 2016. Bee Venom Acupuncture Alleviates Experimental Autoimmune Encephalomyelitis by Upregulating Regulatory T Cells and Suppressing Th1 and Th17 Responses.Mol. Neurobiol. 53, 1419-1445. doi: 10.1007/s12035-014-9012-2.

Terra, R.M., Guimarães, J.A., Verli, H.J.,2007. Structural and functional behavior of biologically active monomeric melittin. Mol. Graph. Model. 25, 767-772.

Mingomataj, E.C., Bakiri, A.H., Ibranji, A., Sturm, G.J., 2014. Unusual reactions to hymenoptera stings: what should we keep in mind. Clin. Rev. Allergy. Immunol. 47, 91-99. doi: 10.1007/s12016-014-8434-y.

Arruda, V.M., Alves, V.V. JR., Moraes M.M, Chaud Netto, J., Suárez, Y.R., 2007. Morphologic analysis of the venom gland of *Apis mellifera* L. (Hymenoptera: Apidae) in populations of Mato Grosso do Sul, Brazil. Neotrop. Entomol. 36, 203-209.

Wanner, K. and Robertson, H., 2008. The gustatory receptor family in the silkworm moth

Bombyx mori is characterized by a large expansion of a single lineage of putative bitter receptors.

Insect Mol.Biol. 17, 621-629. doi: 10.1111/j.1365-2583.2008.00836.x.

King, A.M. and MacRae, T.H., 2015. Insect heat shock proteins during stress and diapause.

Annu. Rev. Entomol. 60: 59-75. doi: 10.1146/annurev-ento-011613-162107.

Huang, H.J., Xue, J., Zhuo, J.C., Cheng, R.L., Xu, H.J., Zhang, C.X., 2017. Comparative analysis

of the transcriptional responses to low and high temperatures in three rice planthopper species.

Mol. Ecol. 26, 2726-2737. doi: 10.1111/mec.14067.

Abbreviations
DEGs: differentially expressed genes

FPKM: fragments per kilobase of transcript per million mapped reads

STEM: Short Time-series Expression Miner

MRJPs: major royal jelly

sHSP: small heat shock protein

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

SRA: Sequence Read Archive

Highlights

- Newly emerged worker bees had the most distinct antennal DEG profile.

- Expression ofgenes for MRJPs and venom wasenriched in juvenilebees.

- Small HSPswere significantly higherexpressedin bees foraging outside the hive.

- Some metabolism-related genes were more highly expressed atthe foraging stage.