Accepted Manuscript

Uncovering the immune responses of Apis mellifera ligustica larval gut to Ascosphaera apis infection utilizing transcriptome sequencing

Dafu Chen, Rui Guo, Xijian Xu, Cuiling Xiong, Qin Liang, Yanzhen Zheng, Qun Luo, Zhaonan Zhang, Zhijian Huang, Dhiraj Kumar, Weijun Xi, Xuan Zou, Min Liu

PII: S0378-1119(17)30273-1
Reference: GENE 41871
To appear in: Gene
Received date: 7 December 2016
Revised date: 30 March 2017
Accepted date: 14 April 2017

Please cite this article as: Dafu Chen, Rui Guo, Xijian Xu, Cuiling Xiong, Qin Liang, Yanzhen Zheng, Qun Luo, Zhaonan Zhang, Zhijian Huang, Dhiraj Kumar, Weijun Xi, Xuan Zou, Min Liu, Uncovering the immune responses of Apis mellifera ligustica larval gut to Ascosphaera apis infection utilizing transcriptome sequencing. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Gene(2017), doi: 10.1016/j.gene.2017.04.022

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Uncovering the immune responses of *Apis mellifera ligustica* larval gut to *Ascosphaera apis* infection utilizing transcriptome sequencing

Dafu Chen¹#, Rui Guo¹**, Xijian Xu¹#, Cuiling Xiong¹, Qin Liang¹, Yanzhen Zheng¹,
Qun Luo², Zhaonan Zhang¹, Zhijian Huang¹, Dhiraj Kumar³, Weijun Xi⁴, Xuan Zou¹,
Min Liu¹

¹ College of Bee Science, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China
² Apiculture institute of Jiangxi province, Nanchang, Jiangxi, China
³ School of Biology and Basic Medical Science, Soochow University, Suzhou, Jiangsu, China
⁴ Jinhua Academy of Agricultural Sciences, Jinhua, Zhejiang, China

# These authors contributed equally to this work.

* Corresponding author:

E-mail address: fufu_ruigu@126.com
Abstract: Honeybees are susceptible to a variety of diseases, including chalkbrood, which is capable of causing huge losses of both the number of bees and colony productivity. This research is designed to characterize the transcriptome profiles of *Ascosphaera apis*-treated and un-treated larval guts of *Apis mellifera ligustica* in an attempt to unravel the molecular mechanism underlying the immune responses of western honeybee larval guts to mycosis. In this study, 24, 296 and 2157 genes were observed to be differentially expressed in *A. apis*-treated *Apis mellifera* (4-, 5- and 6-day-old) compared with un-treated larval guts. Moreover, the expression patterns of differentially expressed genes (DEGs) were examined via trend analysis, and subsequently, Gene Ontology analysis and KEGG pathway enrichment analysis were conducted for DEGs involved in up- and down-regulated profiles. Immunity-related pathways were selected for further analysis, and our results demonstrated that a total of 13 and 50 DEGs were annotated in the humoral immune-related and cellular immune-related pathways, respectively. Additionally, we observed that many DEGs up-regulated in treated guts were part of cellular immune pathways, such as the lysosome, ubiquitin mediated proteolysis, and insect hormone biosynthesis pathways and were induced by *A. apis* invasion. However, more down-regulated DEGs were restrained. Surprisingly, a majority of DEGs within the Toll-like receptor signaling pathway, and the MAPK signaling pathway were up-regulated in treated guts, while all but two genes involved in the NF-κB signaling pathway were down-regulated, which suggested that most genes involved in humoral immune-related pathways were activated in response to the invasive fungal pathogen. This study’s findings provide
valuable information regarding the investigation of the molecular mechanism of immunity defenses of A. m. ligustica larval guts to infection with A. apis. Furthermore, these studies lay the groundwork for future researches on key genes controlling the susceptibility of A. m. ligustica larvae to chalkbrood.

**Keywords:** Immune response, *Apis mellifera ligustica*, Larval gut, *Ascosphaera apis*, RNA-seq, Transcriptome

1. Introduction

The honeybee *Apis mellifera* is a well-known social insect; it is of great significance not only as a model for studying development, neurobiology, social behavior, epigenomics, and host-parasite relationships (Galizia et al., 2012; Begna et al., 2012; Zayed 2012; Foret et al., 2012; Kurze et al., 2016) but also for its role in agriculture and ecosystem functions (Committee on the status of pollinators in North America, 2007). Bees are vital pollinators for as much as 70% of all crop species throughout the world as well as for wild flora (Klein et al., 2007; Elke, 2010). Hence, they are crucial for an economic, sustainable agriculture and for food security. However, honeybees are attacked by various factors including nutritional stress, parasites and diseases. One representation of fungal disease is chalkbrood (Hornitzky, 2010; Aronstein and Murray, 2010), which is caused by *Ascosphaera apis* (Maassen ex Claussen) (Spiltoir, 1955; Spiltoir and Olive, 1955). Larvae ingest the fungal spores when feeding, the fungal spores germinate in the larval gut after ingestion, and finally mycelia cross the gut lining and proliferate through the body cavity. The stretched larvae are killed and then dry out, leaving behind a chalkbrood mummy containing millions of ascospores (Heath,
The chalk become dark if fruiting bodies of the fungi are formed (sporulated mummies). It is speculated that CO$_2$ produced by the larval tissues accumulates in the closed hindgut of the larvae and stimulates spore germination (Heath and Gaze, 1987). A wide spectrum of fungicides had been tested against _A. apis_, and none of them are approved for controlling chalkbrood disease (Galizia et al., 2012). Thus, improved genetic lines, good management and sanitation practices are the most preferred tactics against this disease (Aronstein and Murray, 2010).

Until recently, despite a large number of genomic and functional studies for such insects as _Drosophila melanogaster_ and _Bombyx mori_ (Kari et al., 2016; Kučerová et al., 2016; Guo, 2015; Xue et al., 2015), much is still unknown about antifungal immune defenses in honeybee. Evans and colleagues identified most components of the honeybee humoral immune defense, and described members of the two principal NF-$\kappa$B/Rel immune signaling pathways, Toll and the Immune Deficiency (IMD) pathways (Evans et al., 2016). This opened the door for further research of honeybee immune responses to specific pathogens. A previous study used amplified fragment length polymorphism (cDNA-AFLP) technology and real-time PCR to examine transcriptional changes in _A. apis_-infected and un-infected _A. mellifera_ larvae, and found that the differentially expressed genes (DEGs) are involved in primary functions related to transcriptional regulation, nutritional regulation, RNA processing and apoptotic degradation of ubiquitinated proteins (Aronstein et al., 2010). In comparison with the fly orthologues, the immune system in the honeybee genome consists only of a reduced set of immunity related molecules (Evans et al., 2016). To enhance the immune
responses within an individual bee, the social immunity against parasites is derived from the behavioral cooperation among individuals (Cremer et al., 2007), including hygienic behavior (Spivak and Gilliam, 1993; Evans and Spivak, 2010).

Publication of the genomic sequence of *A. mellifera* provided an invaluable source for further exploring this significant biological model (Honeybee Genome Sequencing Consortium, 2006). To the best of our knowledge, we for the first time used high-throughput RNA sequencing technology to survey the immune response of *Apis mellifera ligustica* larval gut to *A. apis* infection at the transcriptome level. In the current investigation, we compared DEGs between *A. apis*-treated and un-treated larval guts, examined the expression patterns of DEGs via trend analysis, and further carried out Gene ontology (GO) classification and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DEGs within the up- and down-regulated cluster groups. Briefly, our findings provided valuable and precise transcriptome information for investigation of the molecular mechanism underlying immunity defenses of *A. m. ligustica* larval gut to *A. apis* infection. We further lay the groundwork for future studies on key genes controlling the susceptibility of *A. m. ligustica* larvae to chalkbrood.

2. Materials and Methods

2.1. Honeybee larvae and experiment procedures

*A. m. ligustica* larvae used in this study were sampled from the teaching apiary of the College of Bee Science in Fujian Agriculture and Forestry University. *A. apis* was conserved in the Honeybee Protection Laboratory of the College of Bee Science in
Honeybee larvae were reared according to the method developed by Peng et al. (Peng et al., 1992). Their diet was mixed and frozen in smaller aliquots and was pre-heated to 34°C before being used for feeding. 2-day-old larvae were taken from the combs with a Chinese grafting tool and carefully transferred to a droplet of 10 μL diets. Larval age was estimated by size because the mean weights of different age groups differs significantly (Brødsgaard et al., 1998). The larvae were fed once a day with 20 μL (3-day-old), 30 μL (4-day-old), 40 μL (5-day-old) and 50 μL (6-day-old) adding up to 150 μL diet in total. The risk of injury was avoided because no additional transfer of the larvae was performed using this feeding method. Fresh spores of A. apis were purified as previously described (Jensen et al., 2013), and subsequently used to feed 3-day-old larvae at a final concentration of 10^7 spore/mL to cause effective infection. The larvae consumed all diet were used for further study and the larvae cannot consumed all diet were discarded. Four day old larvae reared with an artificial diet without A. apis spores was used as the control group. Culture plates (NEXT, China) were incubated at 95% RH and 33°C as described by Aronstein et al (Aronstein et al., 2006). As the honeybee larval gut is the main site A. apis parasitizes and larvae usually die before the prepupa stage, 21 guts of 4-, 5- or 6-day-old honeybee larvae from A. apis-treated groups and 21 guts of 4-day-old honeybee larvae from the un-treated group were collected, frozen in liquid nitrogen and stored at −80°C until RNA-seq and real time quantitative reverse transcription PCR (RT-qPCR) experiments were performed.

2.2. Total RNA isolation, cDNA library preparation and Illumina sequencing
Total RNAs were extracted from each sample (pool of seven larval guts) using a TRIzol Kit (Promega, USA) according to the manufacturer’s instructions. RNA quality was examined using a 2100 Bioanalyzer (Agilent Technologies, USA) and RNase-free agarose gel electrophoresis. The total RNA concentration was measured using a 2100 Bioanalyzer. RNA samples were stored at -80°C for later library construction and sequencing. Experiments were performed using three replicates for each sample from both the A. apis-treated groups and the control group.

A total of twelve RNA libraries were constructed, representing samples from A. apis-treated groups and the control group. The libraries were as follows: AmT1-1, AmT1-2 and AmT1-3 are replicate libraries for 4-day-old larvae from the A. apis-treated group, AmT2-1, AmT2-2 and AmT2-3 are replicate libraries for 5-day-old larvae from the A. apis-treated group, AmT3-1, AmT3-2 and AmT3-3 are replicate libraries for 6-day-old larvae from the A. apis-treated group, and AmCK-1, AmCK-2 and AmCK-3 are replicate libraries for 4-day-old larvae from the control group. Oligo (dT)s were used to isolate poly (A) mRNA, which was subsequently fragmented followed by cDNA synthesis using random hexamers. Second-strand cDNAs were synthesized using RNase H and DNA polymerase I. The double-stranded cDNAs were then purified using the QiaQuick PCR extraction kit (QIAGEN, Germany). The required fragments were purified by agarose gel electrophoresis followed by enrichment through PCR amplification. Finally, the amplified fragments were sequenced on the Illumina HiSeq™ 2500 platform (GeneDenovo Co., Guangzhou, China) using 125 bp pair-end sequencing. All reads produced in this research have
have been deposited in the National Centre for Biotechnology Information (NCBI) and could be accessed in the short Read Archive (SRA) Database under accession number SRP456722.

2.3. Mapping reads to the reference genome

The original sequencing image data were transferred into sequence data through base calling, which is defined as raw data or raw reads stored in the fastq format. Raw reads of all twelve samples were pre-processed by removing adaptor sequence and reads with more than 5% unknown nucleotides. Low-quality reads were removed as well, defined as reads where the percentage of low-quality bases of quality value \( \leq 5 \) was more than 50% in a read. Finally, the clean reads were aligned to the *A. mellifera* genome assembly Amel_4.5 (http://www.ncbi.nlm.nih.gov/genome/48?genome_assembly_id=22683) using SOAP aligner/soap2 with the threshold that no more than two mismatches were permitted in the alignment.

2.4. Expression annotation and DEGs analysis

The correlation coefficient between two replicas was calculated using R package (version 2.16.2) to provide an evaluation of the reliability of experimental results as well as operational stability. To determine the separation of expression patterns across samples, principle component analysis (PCA) was performed on all genes level using R package (gmodels, version 2.16.2).

For analysis of gene expression, the number of unique-match reads was calculated and normalized to the FPKM (fragments per kilobase of transcript per million mapped
reads). Expression levels of each gene between the two groups were compared using the edgeR package (Robinson et al., 2010) to give an expression difference value. The \( p \) value corresponded to differential gene expression at statistically significant levels (Lu et al., 2014). FDR (false discovery rate) was used to determine the \( p \) value threshold. DEGs were defined as FDR \( \leq 0.05 \) and absolute value of log\(_2\)Fold change \( \geq 1 \). Further, Venn analysis was carried out to identify the shared DEGs among AmT1 vs AmCK, AmT2 vs AmCK and AmT3 vs AmCK using the OmicShare tools, a free online platform for data analysis (www.omicshare.com/tools).

### 2.5. Trend analysis

Gene expression data \( \psi \) (from 4 d to 6 d of \( A. \) apis treatment) were normalized to log\(_2\) (4 d/4 d), log\(_2\) (5 d/4 d), log\(_2\) (6 d/5 d). Subsequently, DEGs were used for further expression pattern analysis via STEM (Short Time-series Expression Miner, v1.3.8) (Ernst and Bar-Joseph, 2006). The union set of all DEGs were profiled into twenty six expression patterns. Among them, the clustered profiles with \( p \)-values \( \leq 0.05 \) were considered as significantly expressed. Functional classification of the DEGs within up-regulated or down-regulated clusters was carried out using WEGO software (Ye et al., 2006), and KEGG pathway annotation by Blastall software against the KEGG (kyoto encyclopedia of genes and genomes) database (http://www.kegg.jp/).

### 2.6. Real time quantitative PCR validation of RNA seq data

First-strand cDNA was generated from 1 \( \mu \)g total RNA isolated from the gut samples of \( A. \) apis-treated groups and the control group using the superscript first-strand synthesis system (TaKaRa, China). Primers for RT-qPCR were designed with
DNAMAN software (Lynnon Biosoft, USA) and synthesized by Sangon Biotech (Shanghai, China) Co., Ltd. To confirm the RNA-seq data in this study, nine genes were randomly selected that were consistent with those derived from the Illumina sequencing and were subjected to RT-qPCR. The actin gene (gene9102) was chosen as an internal reference gene. RT-qPCR assays were performed on an ABI 7500 Real-time PCR Detection System (Applied Biosystems technologies, USA) with SYBR Green Dye (Vazyme, China). The thermal cycling conditions were as follows: 95°C for 1 min, followed by 40 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 45 s. The Ct values were determined for three biological replicates, with three technical replicates for each value. The relative expression of the cDNAs were calculated by first normalizing to endogenous reference gene and then normalizing the expression level in the untreated samples based on the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The specific primers used for RT-qPCR are listed in Table S1.

3. Results

3.1. An overview of Illumina sequencing datasets

To examine the immune responses of *A. m. ligustica* larval gut to *A. apis* infection, twelve cDNA libraries were prepared and subjected to next-generation sequencing using the Illumina HiSeq2500 platform. After removing adaptor sequences and low-quality reads, approximately 30.85, 29.52, 29.02 and 30.44 million raw reads were obtained for AmCK, AmT1, AmT2 and AmT3 transcriptomes. After filtering the adaptor sequences, regions containing N sequences and low quality sequences, over 28.69 million clean reads were produced in each library. The percentage of clean reads
among raw tags in each library ranged from 97.16% to 99.03% (Table S2), suggesting that the obtained high quality RNA-seq data could be used for further analysis. An overview of the sequencing statistics is displayed in Table 1. An average of 25.76 million reads per sample was mapped to the western honeybee genome (ranging from 22.93 to 27.36 million). Of the total reads, the rate of match reads was more than 86.25% (ranging from 86.25% to 89.45%) (Table S3), while the remaining reads were unmatched.

In this study, the correlation values of three biological replicates for each sample were more than 0.9 ($R^2 > 0.9$) based on the RPKM values. Meanwhile, PCA analysis was performed and the result showed that samples from the A. apis-treated groups and the control group clustered into four separated groups (Fig. 1). These results were indicative of sufficient reproducibility and rationality of sampling.

### 3.2. DEGs in response to A. apis infection

To explore genes that were differentially expressed between the A. apis-treated groups and the un-treated group, DEGs for AmT1 vs AmCK, AmT2 vs AmCK and AmT3 vs AmCK were identified using the edgeR package, with an FDR $\leq 0.05$ and $|\log_2 FC| \geq 1$. In the AmT1 vs AmCK group, 4 genes were up-regulated, while 20 genes were down-regulated. In the AmT2 vs AmCK group, the number of up-regulated and down-regulated genes were 162 and 134, respectively. In the AmT3 vs AmCK group, 980 and 1177 genes were up-regulated and down-regulated (Table 2). The volcano plots of DEGs between every two groups are displayed in Fig. 2.

Moreover, based on Venn analysis, we found 0 and 4 genes were commonly
up-regulated and down-regulated in AmT1 vs AmCK, AmT2 vs AmCK and AmT3 vs AmCK. These genes include 1, 16 and 831 genes that were uniquely up-regulated, and 13, 31 and 1071 genes that were uniquely down-regulated in each group, respectively (Fig. 3). This finding implied that these shared up-/down-regulated genes are likely to play a pivotal role during the whole process of A. m. ligustica larval gut in response to A. apis infection, while those genes that are uniquely up-/down-regulated likely function at different stages of fungal infection.

3.3. Trend analysis of DEGs

To monitor the significant gene expression profiles of all DEGs, the gene expression data (from 4 d to 6 d of A. apis treatment) were normalized to log2 (4 d/4 d), log2 (5 d/4 d), log2 (6 d/4 d). The 2774 DEGs could be clustered into twenty six profiles on the basis of their expression level via STEM software (-u 1 -pro 20 -ratio 1.0) (Fig. 4A). Of those, 2384 DEGs were grouped into just nine profiles (p < 0.05), including five up-regulated patterns (profile 13, 15, 16, 22 and 25) (Fig. 4E-I) and three down-regulated patterns (profile 9, 10 and 12) (Fig. 4B-D). Profiles 13, 15, 16, 22 and 25 contained 344, 109, 383, 66 and 73 DEGs (Table S4-S8), while Profile 9, 10 and 12 contained 844, 145 and 420 DEGs (Table S9-S11), respectively. The aforementioned findings demonstrate that the number of DEGs in the up-regulated pathways (975) was less than that of DEGs in the down-regulated pathways (1409), which suggested that some DEGs were activated by fungal infection, but more DEGs were suppressed by infection.

3.4. GO term analysis of DEGs
The DEGs within up- or down-regulated cluster groups ($p < 0.05$) were subjected to GO term analysis, where they were classified into three main categories including biological process, cellular component and molecular function. On the basis of WEGO, DEGs within up-regulated cluster groups were categorized into 37 functional groups. In these groups, most genes were enriched in “binding” (209 genes, 21.4%), followed by “catalytic activity” (197 genes, 20.26%) and “metabolic process” (178 genes, 18.3%) (Fig. 5A). DEGs within down-regulated cluster groups were classified into 38 functional groups, including “binding” (495 genes, 35.1%), “metabolic process” (461 genes, 32.7%) and “catalytic activity” (457 genes, 32.4%) (Fig. 5B). This finding implied that the honeybee larvae metabolic function may be seriously affected by mycosis.

Further investigation result suggested that 104 DEGs were enriched in “response to stimulus” (Table S12), including 67 genes showing up-regulated trends and 37 genes displaying down-regulated profiles; 46 DEGs with a up-regulated trend and 10 DEGs with a down-regulated trend were enriched in “signaling” (Table S13); 3 DEGs were enriched in “immune system processes” (Table S14), and all of them showed up-regulated profiles. It’s speculated that DEGs involved in the aforementioned three GO terms are likely to be in close relation to the immune responses of A. m. ligustica larval gut to A. apis infection.

3.5. KEGG pathway enrichment analysis of DEGs

KEGG pathway enrichment analysis was conducted for DEGs within the up-/down-regulated cluster groups (Fig. 6). The 10 top KEGG pathways with the
highest representation of DEGs within the up-/down-regulated clusters are shown in **Table S15** and **Table S16**, respectively. For DEGs within the up-regulated cluster groups, the FoxO signaling pathway (ko04068) was mostly enriched, followed by biosynthesis of amino acids (ko01230) and Carbon metabolism (ko01200). For DEGs within the down-regulated cluster groups, the mostly enriched pathway was that of RNA transport (ko03013), followed by protein processing in endoplasmic reticulum (ko04141) as well as ribosome biogenesis in eukaryotes (ko03008).

Insects, such as honeybees, lack an acquired immune system but have a well-developed innate response (Lavine and Strand, 2002). When challenged by chalkbrood, *A. m. ligustica* larvae launched an array of immune defenses via the activation of humoral and cellular immune reactions. To illustrate the immune responses of *A. m. ligustica* larval gut after *A. apis* infection, immune-related pathways were selected for further analysis based on the result of KEGG pathway enrichment analysis.

The number of DEGs involved in cellular immune-related and humoral immune-related pathways during 4 d to 6 d of *A. apis* treatment were calculated and displayed in **Table 3** and **4**. A total of 50 DEGs were annotated in the cellular immune-related pathways, such as lysosome, ubiquitin mediated proteolysis, endocytosis, phagosome, melanogenesis, insect hormone biosynthesis and apoptosis pathways (**Table 3**), while 13 DEGs were annotated in the humoral immune-related pathways, including Mitogen-activated protein kinase (MAPK) signaling pathway, NF-κB signaling pathway as well as Toll-like receptor signaling pathway (**Table 4**).
Furthermore, these DEGs that are related to cellular immunity and humoral immunity had significantly different expression patterns in different samples, as shown in Fig. 7A and B, respectively.

Cellular immune responses start immediately after an invasion is detected in the haemolymph (Govind, 2008). In the lysosome pathway, 5 out of 12 DEGs showed up-regulated profiles in the larval gut, while 7 genes displayed down-regulated trends. Similar results were found in the endocytosis pathway. In the ubiquitin mediated proteolysis pathway, 2 genes encoding ubiquitin-conjugating enzyme E2 R2-like isoform X2 (gene2331, profile 13) and dentin sialophosphoprotein-like (gene7256, profile 13) displayed up-regulated profiles, whereas 10 genes encoding anaphase-promoting complex subunit 1 (gene11191, profile 9), SUMO-activating enzyme subunit 1 (gene2181, profile 9) and ubiquitin conjugation factor E4 A-like isoform X2 (gene717, profile 11) among others, showed down-regulation in the gut. In the insect hormone biosynthesis pathway, it was noticed that only 1 DEG encoding cytochrome P450 314A1 isoform X1 (gene4743, profile 16) was up-regulated, while and 2 DEGs encoding cytochrome P450 302a1, mitochondrial (gene12185, profile 9) and cytochrome P450 306a1 isoform X1 (gene9960, profile 9) were down-regulated. Notably, all DEGs in the phagosome pathway and the apoptosis pathway showed down-regulated profiles, suggesting that both pathways were greatly suppressed by A. apis. In the melanogenesis pathway, 5 DEGs encoding protein Wnt-10b (gene1279, profile 16), protein pangolin, isoforms A/H/I/S-like isoform X1 (gene3343, profile 13), adenylate cyclase type 5-like (gene6716, profile 16), protein kinase shaggy isoform
X4 (gene7560, profile 13) as well as adenylate cyclase type 8-like (gene7836, profile 15) were up-regulated, implying that this pathway was dramatically activated by mycosis. These results jointly demonstrated that DEGs in cellular immune pathways were induced by *A. apis* invasion, however, more DEGs were restrained.

The challenge caused by invading microorganism results in activation of the honeybee humoral immunity, which leads to induction of synthesis of AMPs and bacteriolytic enzymes as well as activation of the prophenoloxidase system. These act in combination to inactivate or kill the pathogens (Hedengren-Olcott et al., 2004; Stanley et al., 2009). Strikingly, two DEGs involved in the Toll-like receptor signaling pathway (gene2489, profile 13, gene7642, profile 13) showed up-regulated profiles, indicating that these two pathways were significantly induced by *A. apis*. In contrast, both gene8189 (profile 12) and gene900 (profile 12) in the NF-κB signaling pathway were down-regulated, suggesting that this pathway was inhibited by mycosis to a large extent. In the MAPK signaling pathway, 10 out of the 11 DEGs were up-regulated, only 1 DEG (gene8031, profile 9) encoding nucleoprotein TPR was down-regulated. This finding implied that this pathway was largely activated by fungal infection (Fig. 7B). These results jointly demonstrated that most genes involved in these signaling pathways were activated in response to the invasive fungal pathogen.

3.6. Confirmation of RNA seq data by RT-qPCR

To validate the accuracy and reproducibility of the transcriptome results, nine DEGs were randomly selected for RT-qPCR verification. Total RNA isolated from *A. apis*-treated group (AmT3) and the un-treated group (AmCK) were used as templates.
The results showed that the expression patterns of the candidate genes were consistent with those from RNA-seq (Fig. 8), which confirmed that the measured changes in gene expression detected by RNA-seq reflected the actual transcriptome differences between the different libraries.

4. Discussion

Chalkbrood is the most common fungal bee brood disease and results in huge losses for apiculture by lowering the number of newly emerged bees, which reduces colony productivity (Aronstein and Murray, 2010; Jensen et al., 2013). In the present research, we utilized RNA-seq technology to monitor DEGs within A. apis-treated and un-treated western honeybee larval guts, and further identified immune-related genes that are differentially expressed during the disease process. In previous studies (Vojvodic et al., 2011; Lee et al., 2013), the commonly used number of A. apis spores for feeding honeybee larvae is $10^6$. To control the spore numbers, we changed the volume of artificial diet used for feeding 3-day-old larvae from 5 μL to 20 μL, and found that it's easier for the larvae to consume all artificial diet (containing about $10^6$ spores of A. apis). In our previous experiments, approximately 70.83% of A. apis-treated larvae died due to chalkbrood during prepupae stage (5 d post inoculation of about $10^6$ spores), so we believe that this spore quantity is able to result in an efficient infection of honeybee larvae. The larval gut is the main place parasitized by A. apis. Meanwhile, A. apis spores consumed by honeybee larvae slightly germinated inside the larval midgut at early stage, and when the midgut and hindgut connected at prepupae stage, the spores rapidly germinated and the mycelial heavily grew inside
the larval hindgut (Li et al., 2012). Thus, 4-, 5- and 6-day-old larval guts were chosen as sequencing materials in this study, and the transcriptome information obtained here more precisely reflects the immune responses of *A. m. ligustica* larvae to mycosis at the transcriptome level.

In a previous study, Aronstein and colleagues used cDNA-AFLP technology to profile transcripts in un-infected and *A. apis*-infected western honeybee larvae, and obtained 98 reproducible polymorphic fragments. However, they did not detect any transcriptional changes in genes of the Toll and IMD pathways, though some previously identified immune-related genes were tested using RT-qPCR (Aronstein et al., 2010). In our research, we found that 24, 196 and 2157 genes were differentially expressed in AmT1 vs AmCK, AmT2 vs AmCK and AmT3 vs AmCK, and trend analysis demonstrated that 975 and 1409 DEGs displayed up- and down-regulated profiles. As a result, the number of DEGs obtained from high-throughput sequencing is much more than the previous findings. Moreover, in response to *A. apis* infection, there were two genes involved in the Toll-like receptor signaling pathway and two genes within the NF-κB signaling pathway identified in current research.

Our primary target is to illustrate the immune responses of honeybee larval gut to infection with *A. apis*. Following trend analysis, KEGG pathway enrichment analysis for the obtained 2774 DEGs was performed, and those genes involved in immune pathways were further analyzed. For insects including honeybees, cuticles and peritrophic membranes are the primary physical barriers (Orihel, 1975). When the first defending line is breached, the pathogenic microorganisms encounter a set of efficient
cellular and humoral defenses including encapsulation, melanization, phagocytosis, enzymatic degradation of pathogens and secretion of antimicrobial peptides (Gliński and Jarosz, 2001; Glinski and Buczek, 2003). For bees, phagocytosis and encapsulation are the most common defense mechanisms against entomopathogenic fungi (Glinski and Buczek, 2003). The ubiquitin proteasome degradation system is the principal mechanism for eliminating unneeded or damaged cells. It serves a crucial means for the regulation of most cellular processes, such as cell cycle and division, differentiation, development, and immunity (William et al., 2003). In insects, apoptosis or programmed cell death (PCD) are strategies employed for antiviral defense (Narayan, 2004; Guo et al., 2015). Strikingly, we found in the present research that a majority of DEGs involved in the lysosome, endocytosis, ubiquitin mediated proteolysis, insect hormone biosynthesis, apoptosis as well as melanogenesis pathways were down-regulated, demonstrating that cellular immunity of A. m. ligustica larvae was extremely inhibited by A. apis, which is believed to lay the foundation for the outbreak of chalkbrood during the late stage.

When A. apis invasion advanced, humoral immunity is activated to induce synthesis of AMPs and bacteriolytic enzymes and activate the prophenoloxidase system. They can work in collaboration to inactivate or kill the invading microorganisms (Stanley et al., 2009). In Drosophila, the humoral response is composed of the Toll and IMD pathways. The Toll/Dif pathway is induced by Gram-positive bacterial and fungal infections, while the IMD/Relish pathway is mostly activated by Gram-negative and some Gram-positive bacterial pathogens (Evans and Spivak, 2010; Govind, 2008).
Evans and colleagues identified most components of the western honeybee NF-κB signaling pathway (Evans et al., 2006), which offered meaningful information for further investigation of the immune response of honeybee at a molecular level. In their study, they found that immune-related molecules were substantially reduced in the bee genome when compared to fly orthologues (Evans et al., 2006). Findings in our research showed that 2, 2 and 11 genes were differentially expressed in the Toll-like receptor signaling pathway, NF-κB signaling pathway as well as MAPK signaling pathway, respectively. The Toll/Dif signaling pathway is confirmed to be activated by fungal and Gram-positive bacterial pathogens (Tanji et al., 2007). Activation of this pathway leads to synthesis of AMPs (Hu et al., 2004; Weber et al., 2003). Aronstein et al. isolated and characterized a novel gene Am18w from the honeybee A. mellifera, which encodes for the Toll-like receptor. Furthermore, they monitored the expression of Am18w after immune challenge, and found a 3.1-fold increase in gene expression occurred 24 h post infection in response to A. apis (Aronstein and Saldivar, 2005). In the current investigation, further analysis suggested that genes within the Toll-like receptor signaling pathway showed up-regulated profiles, which is in accordance with Aronstein’s study. The MAPK cascade, activated by the IMD pathway, is known as the JNK pathway (Davis et al., 2008). In addition to the JNK pathway, the p38 MAPK family (including p38a and p38b) has been implicated in attenuation of the immune response in Drosophila (Han et al., 1998). Data from this research demonstrated that most of the genes involved in MAPK signaling were up-regulated after mycosis invasion, indicating that this pathway was significantly activated by A. apis infection.
Taken together, these results demonstrated that activation of the humoral immunity in the western honeybee larval gut was induced by invasive *A. apis*. To our knowledge, all types of septic injections resulted in rapid induction of AMPs, but none of the responses seemed pathogen-specific. Davis et al. found that fungal septic injections could activate components of both the Toll signaling pathway and the JNK branch of the IMD signaling pathway (Davis et al., 2008), which was in line with findings in this study. Unexpectedly, no component of the Jak-STAT signaling pathway was indentified in this research, which suggested that this pathway in *A. m. ligustica* larval gut was not induced by *A. apis* infection. However, in another investigation, we found that two genes within the Jak-STAT signaling pathway in the eastern honeybee (*Apis cerana cerana*) larval gut invaded by *A. apis* had were up-regulated (data unpublished). In apiculture production, the western honeybee larvae is susceptible to *A. apis*, while the eastern honeybee larvae has strong *A. apis* resistance. The Jak-STAT signaling pathway may partly contribute to the resistance difference between these two honeybee species.

This is the first study on transcriptome analysis of immune responses of the western honeybee larval gut to chalkbrood using next-generation sequencing technology. It provides valuable information not only for further unraveling the molecular mechanism underlying *A. m. ligustica* larvae individual immune responses to *A. apis* challenge, but also for illustrating the interaction between host and pathogen during chalkbrood disease process. A wide range of fungicides has been tested against *A. apis*; unfortunately, none of them are approved for application in beehives (Hornitzky, 2001). A better understanding of the immunity mechanisms of honeybee larvae in response to
infection with *A. apis* will facilitate development of increasingly effective control strategy of chalkbrood disease. Promisingly, based on the findings from our present research, the *A. apis*-resistance difference between *A. m. ligustica* and *A. c. cerana* larvae is expected to be unveiled with the help of comparative transcriptomics, which is a goal for the near future.

**Acknowledgements**

We gratefully acknowledge financial support of the earmarked fund for China Agriculture Research System (CARS-45-KXJ7), the National Natural Science Foundation of China (30800806), the science and technology development fund of Fujian Agriculture and Forestry University (KF2015123), and the undergraduate innovation and entrepreneurship training program of Fujian province (3135602028).

We thank Prof. Bingfeng Zhou for providing assistance for rearing honeybee larvae.

**References**


Orihel, T.C., 1975. The peritrophic membrane: its role as a barrier to infection of the


ubiquitin/proteasome system in cellular responses to radiation. Oncogene 22(37), 5755-5773.


Table 1 Throughput and quality of RNA seq datasets.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Raw reads</th>
<th>Clean reads</th>
<th>GC percentage (%)</th>
<th>Q20 percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AmCK</td>
<td>30846191</td>
<td>30496725</td>
<td>41.37</td>
<td>98.45</td>
</tr>
<tr>
<td>AmT1</td>
<td>29519207</td>
<td>29196197</td>
<td>41.45</td>
<td>96.15</td>
</tr>
<tr>
<td>AmT2</td>
<td>29015564</td>
<td>28690943</td>
<td>41.21</td>
<td>98.47</td>
</tr>
<tr>
<td>AmT3</td>
<td>30442169</td>
<td>29779715</td>
<td>42.56</td>
<td>98.01</td>
</tr>
</tbody>
</table>

Table 2 Summary of DEGs in AmT1 VS AmCK, AmT2 VS AmCK and AmT3 VS AmCK.
<table>
<thead>
<tr>
<th>Pathway</th>
<th>Num (599)</th>
<th>p value</th>
<th>q value</th>
<th>Pathway ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysosome</td>
<td>12</td>
<td>0.855273</td>
<td>0.997777</td>
<td>ko04142</td>
</tr>
<tr>
<td>Ubiquitin mediated proteolysis</td>
<td>12</td>
<td>0.6378</td>
<td>0.997777</td>
<td>ko04120</td>
</tr>
<tr>
<td>Endocytosis</td>
<td>11</td>
<td>0.962283</td>
<td>0.997777</td>
<td>ko04144</td>
</tr>
<tr>
<td>Phagosome</td>
<td>6</td>
<td>0.96621</td>
<td>0.997777</td>
<td>ko04145</td>
</tr>
<tr>
<td>Melanogenesis</td>
<td>5</td>
<td>0.600985</td>
<td>0.997777</td>
<td>ko04916</td>
</tr>
<tr>
<td>Insect hormone biosynthesis</td>
<td>3</td>
<td>0.951536</td>
<td>0.997777</td>
<td>ko00981</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>2</td>
<td>0.974164</td>
<td>0.997777</td>
<td>ko04210</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Num (599)</th>
<th>p value</th>
<th>q value</th>
<th>Pathway ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPK signaling pathway</td>
<td>6</td>
<td>0.859625</td>
<td>0.997777</td>
<td>ko04010</td>
</tr>
<tr>
<td>MAPK signaling pathway - yeast</td>
<td>4</td>
<td>0.665628</td>
<td>0.997777</td>
<td>ko04011</td>
</tr>
<tr>
<td>MAPK signaling pathway - fly</td>
<td>3</td>
<td>0.737105</td>
<td>0.997777</td>
<td>ko04013</td>
</tr>
<tr>
<td>NF-kappa B signaling pathway</td>
<td>2</td>
<td>0.816125</td>
<td>0.997777</td>
<td>ko04064</td>
</tr>
<tr>
<td>Toll-like receptor signaling pathway</td>
<td>2</td>
<td>0.816125</td>
<td>0.997777</td>
<td>ko04620</td>
</tr>
</tbody>
</table>

Figure captions

Fig. 1. Principal component analysis (PCA) of the transcriptome of four different samples. The numbers in parentheses represent the proportion of variance explained by
that principal component. Samples in specific circle were from different conditions. PC1 and PC2 represent the top two dimensions of the genes showing differential expression among these samples, which accounts for 77.5% and 10.3% of the expressed genes, respectively.

**Fig. 2. The volcano plots of DEGs between every two groups.** (A) The volcano plot of DEGs in AmT1 VS AmCK. (B) The volcano plot of DEGs in AmT2 VS AmCK. (C) The volcano plot of DEGs in AmT3 VS AmCK. The horizontal line and vertical line indicate the significance threshold (FDR ≤ 0.05) and two-fold change threshold (|log₂Ratio|≥1), respectively. The green dots indicate down-regulated genes, the black dots indicate genes without differential expression, the red dots indicate up-regulated genes.

**Fig. 3. Venn diagram of DEGs among different groups.** (A) Venn diagram of up-regulated genes among AmT1 VS AmCK, AmT2 VS AmCK and AmT3 VS AmCK. (B) Venn diagram of down-regulated genes among AmT1 VS AmCK, AmT2 VS AmCK and AmT3 VS AmCK.

**Fig. 4. Cluster trajectory profiles across stages of A. apis-treatment.** The profiles are ordered on basis of the p value significance of the number of genes assigned versus expected. The colored square frame denotes significant profile (p value ≤ 0.01). Each graph displays the mean pattern of expression (black lines) of the profile genes. The x-axis represents stages, and the y-axis represents log₂FC in gene expression. (A) STEM clusters of expression profiles. Profiles 9, 10 and 12 indicate a down-regulated trend (B-D). Profiles 13, 15, 16, 22 and 25 indicate a down-regulated trend (E-I).
Fig. 5. Gene Ontology classification of DEGs. (A) GO classification of DEGs with an up-regulation trend. (B) GO classification of DEGs with a down-regulation trend. The up-regulated genes are union of DEGs in profiles 13, 15, 16, 22 and 25, the down-regulated genes are union of DEGs in profiles 9, 10 and 12. The results are summarized in three main categories: biological process, cellular component and molecular function. The x-axis indicates the second term of gene ontology, the y-axis indicates the percentage of genes.

Fig. 6. KEGG pathway enrichment analysis of DEGs. (A) KEGG pathway enrichment analysis of DEGs with up-regulated profiles. (B) KEGG pathway enrichment analysis of DEGs with down-regulated profiles.

Fig. 7. Heatmap of the expressed genes assigned to immune-related pathways. (A) Heatmap of the expressed genes assigned to cellular immune pathways. Gene10351, MD-2-related lipid-recognition protein-like; gene10626, protein Malvolio isoform X1; gene12148, sphingomyelin phosphodiesterase-like isoform X3; gene12403, alpha-L-fucosidase; gene12711, N-sulphoglucosamine sulphohydrolase-like; gene2148, battenin isoform X2; gene3564, cation-independent mannose-6-phosphate receptor isoform X1; gene5785, alpha-N-acetylglucosaminidase-like; gene7075, CD63 antigen; gene7454, beta-hexosaminidase subunit beta-like; gene7629, probable beta-hexosaminidase fdl isoform X3; gene84, lysosomal acid lipase/cholesteryl ester hydrolase-like isoform X2; gene11191, anaphase-promoting complex subunit 1; gene11365, anaphase-promoting complex subunit 13-like; gene11805, S-phase kinase-associated protein 2 isoform X1; gene1284, cell division cycle protein 23
homolog; gene2181, SUMO-activating enzyme subunit 1; gene2331, ubiquitin-conjugating enzyme E2 R2-like isoform X2; gene3932, retinol dehydrogenase 13-like; gene4523, DNA damage-binding protein 1-like, gene7256, dentin sialophosphoprotein-like; gene7397, cullin-4B-like; gene747, ubiquitin conjugation factor E4 A-like isoform X2; gene8039, SUMO-activating enzyme subunit 2 isoform X2; gene11052, fibroblast growth factor receptor homolog 1 isoform X1; gene11317, heat shock protein Hsp70Ab-like; gene12161, vacuolar protein sorting-associated protein VTA1 homolog; gene12409, cytohesin-1-like isoform X1; gene3564, cation-independent mannose-6-phosphate receptor isoform X1; gene4877, epidermal growth factor receptor-like, partial; gene4880, epidermal growth factor receptor-like isoform X1; gene4889, charged multivesicular body protein 7-like; gene7014, WASH complex subunit 7-like; gene7946, WASH complex subunit strumpellin-like isoform X2; gene9885, WASH complex subunit FAM21-like; gene12277, V-type proton ATPase subunit E isoform 3; gene4058, V-type proton ATPase subunit C isoform X1; gene6326, calnexin isoform X1; gene6398, calreticulin; gene6778, protein transport protein Sec61 subunit alpha isoform 2; gene7578, protein transport protein Sec61 subunit gamma-like isoform X1; gene1279, protein Wnt-10b; gene3343, protein pangolin, isoforms A/H/I/S-like isoform X1; gene6716, adenylate cyclase type 5-like; gene7560, protein kinase shaggy isoform X4; gene7836, adenylate cyclase type 8-like; gene12185, cytochrome P450 302a1, mitochondrial; gene4743, cytochrome P450 314A1 isoform X1; gene9960, cytochrome P450 306a1 isoform X1; gene8189, serine-protein kinase ATM isoform X1; gene8420, cytochrome c. (B)
Heatmap of the expressed genes assigned to humoral immune pathways. Gene11052, fibroblast growth factor receptor homolog 1 isoform X1; gene11317, heat shock protein Hsp70Ab-like; gene2489, stress-activated protein kinase JNK isoform X6; gene4877, epidermal growth factor receptor-like, partial; gene4880, epidermal growth factor receptor-like isoform X1; gene7642, transcription factor AP-1; gene13033, catalase-like, partial; gene13245, catalase-like, partial; gene5179, catalase; gene8031, nucleoprotein TPR; gene5197, torso-like isoform X1; gene8189, serine-protein kinase ATM isoform X1; gene900, poly [ADP-ribose] polymerase.

Fig. 8. Real time quantitative PCR verification of RNA seq data. (A) transcriptional regulator for cysteine regulon (gene3591). (B) probable salivary secreted peptide-like (gene1952). (C) chymotrypsin inhibitor-like isoform X1 (gene11350). (D) threo-3-hydroxyaspartate ammonia-lyase-like (gene4563). (E) protein lethal(2)essential for life-like (gene1792). (F) retinoid-inducible serine carboxypeptidase-like isoform X1 (gene6186). (G) FK506-binding protein 2-like (gene4800). (H) aldose reductase-like isoform 1 (gene11302). (I) peptidyl-prolyl cis-trans isomerase-like (gene11961).
Fig. 1
Fig. 2
Fig. 3

A

B

ACCEPTED MANUSCRIPT
Fig. 5

A

B

[Diagram showing bar charts for GO terms at level 2.xls, with categories for biological process, cellular component, and molecular function.]
Fig. 8
Highlights

1. This is the first study on transcriptome analysis of immune responses of the western honeybee larval gut to chalkbrood using next-generation sequencing technology.

2. Cellular immunity of *Apis mellifera ligustica* larvae was extremely inhibited by *Ascosphaera apis*, which is believed to lay the foundation for the outbreak of chalkbrood during the late stage.

3. Activation of the humoral immunity in the western honeybee larval gut was induced by *Ascosphaera apis*.

4. No component of the Jak-STAT signaling pathway was indentified, which suggested that this pathway in *Apis mellifera ligustica* larval gut was not induced by *Ascosphaera apis* infection.
Abbreviations list

A. mellifera: *Apis mellifera*; A. m. ligustica: *Apis mellifera ligustica*; A. apis: *Ascosphaera apis*; A. cerana: *Apis cerana*; A. c. cerana: *Apis cerana cerana*; DEGs: differentially expressed genes; IMD pathway: Immune Deficiency pathway; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; RT-qPCR: real time quantitative reverse transcription PCR; SRA: short Read Archive; PCA: principle component analysis; FDR: false discovery rate; MAPK: Mitogen-activated protein kinase; PCD: programmed cell death;