Journal of Asia-Pacific Entomology 17 (2014) 37–43

Contents lists available at ScienceDirect

Journal of Asia-Pacific Entomology

journal homepage: www.elsevier.com/locate/jape

Mating flight causes genome-wide transcriptional changes in sexually mature honeybee queens

Xiaobo Wu [1], Zilong Wang [1], Fei Zhang, Yuanyuan Shi, Zhijiang Zeng *

Honeybee Research Institute, Jiangxi Agricultural University, Nanchang, Jiangxi 330045, China

ARTICLE INFO

Article history:
Received 29 August 2013
Revised 25 September 2013
Accepted 4 October 2013

Keywords:
Apis mellifera
Mating flight
Digital gene expression
miRNA
Expression difference

ABSTRACT

In this study, we analyzed the gene and miRNA expression differences between the courted virgin queen (CVQ) and non-courted virgin queen (NCVQ) of *Apis mellifera* using a high-throughput sequencing method. Through Digital Gene Expression (DGE) sequencing, 452 genes were differentially expressed, out of which, 90 genes were up-regulated and 362 genes were down-regulated in CVQ compared with NCVQ. Through small RNA sequencing, 27 miRNAs showed significant expression difference between these two samples. Moreover, 9 of the differentially expressed genes are the targets of the 11 differentially expressed miRNAs. Besides, 47 novel miRNA candidates were predicted in these two samples. Our results provided valuable information for understanding the molecular mechanism of the transition to functional queens.

Introduction

Breeding offspring is an important instinct of all animals. In order to breed high-quality offsprings, animals have evolved different sexual selection and mating behaviors to find a high quality mate (Maan and Seehausen, 2011). Studies of transcription changes during the mating process in females have been conducted in *Drosophila melanogaster* (Lawniczak and Begun, 2004; Mack et al., 2006; Dalton et al., 2010), and several differentially expressed genes between unmated and mated females have been identified by microarray analysis.

The honeybee is an important economic insect. Usually, a honeybee colony comprises a queen, thousands of drones and hundreds of thousands of workers. The queen is mainly responsible for reproduction, and virgin queens reach sexual maturity at 5–10 days after emergence. But the sexual maturity virgin queens will not attract drones before they fly to a drone congregation area (DCA), where hundreds or thousands of unrelated drones gather for mating. Then, the drones pursue the queens and usually 7–17 drones could mate with a queen in the mating flight (Taber and Wendel, 1958; Adams et al., 1977). Obviously, the queens maybe released some kind of pheromone to attract the drones during this process. After mating, the ovaries of the mated queens become fully activated to lay eggs. From the virgin queen to functional egg laying queen, the physiological and behavioral changes are substantial and un-reversible (Winston, 1987). Previous research has shown that there are large-scale transcriptional changes in the brains and ovaries between virgin queens, mated queens and egg laying queens (Kocher et al., 2008). Moreover, transcriptional changes in the brains and the ovaries appear to be uncoupled, unlike the physiological and behavioral changes that are coupled during the mating process. Another study detected several differentially expressed genes between naturally mated queens and artificially inseminated queens, and between semen inseminated queens and those treated with a saline control (Kocher et al., 2010).

miRNAs are a special kind of molecules which induce gene silence in organisms. They are involved in the regulation of cell growth, development, gene transcription and translation, and many other life activities (Yang et al., 2005). miRNA has been reported to be present in the reproductive tract (Carletti and Christenson, 2009; Nothnick, 2012), but so far there are no reports of miRNAs being involved in mating behavior. Studies on miRNAs in the honeybee have been performed by several research groups (Weaver et al., 2007; Chen et al., 2010; Greenberg et al., 2012; Liu et al., 2012; Shi et al., 2012). Weaver et al. identified a total of 65 non-redundant candidate miRNAs from the honeybee genome and verified their expression in honeybee (Weaver et al., 2007). Chen et al. (2010) predicted 267 novel honeybee miRNAs by SOLiD sequencing. While Liu et al. (2012) discovered nine significantly differentially expressed known miRNAs between nurses and foragers and 67 novel miRNAs. Another study indicated that miRNAs are important regulators of social behavior (Greenberg et al., 2012). Such results strongly suggest that miRNA may be widely involved in the regulation of various behaviors in the honeybee. Recently, Shi et al. (2012) for the first time reported that miRNA from heterospecific royal jelly can modify gene expression in honey bees.

* Corresponding author.
E-mail address: bees1965@sina.com (Z. Zeng).
[1] These authors contributed equally to this work.

1226-8615/$ – see front matter.
http://dx.doi.org/10.1016/j.aspen.2013.10.001

Since mating flight is crucial for queens to attract drones and causes physiological changes in queens, this process may be accompanied by great changes in gene and miRNA expression. Therefore, in this study, we adopted the DGE and small RNA sequencing approaches to identify differentially expressed genes and miRNAs between CVQ and NCVQ, and survey the relationship between the differentially expressed genes and miRNAs. This work will help us to understand the molecular mechanism of the queen mating flight.

Materials and methods

Insect

Apis mellifera were sampled from the Honey bee Research Institute, Jiangxi Agricultural University, China. In order to ensure a higher genetics similarity among the experimental queens, which should increase the signal to noise ratio when searching for gene expression differences, all the queens used in this experiment were from a single drone inseminated queen and artificially bred at normal conditions. After emergence, these young queens were caged and placed in a normal sized colony without other queens. 12 days after emergence, the mating flight experiment was performed. In order to easily catch the queens after mating flight, a net (5 m × 3 m × 4 m) was installed, then, 6 sexually mature virgin queens and hundreds of sexually mature drones were placed in this net. The drones courted the queens, but in this condition, the queens didn't mate with the drone, because their flying height didn't achieve the height of the mating flight in natural condition. After 20 min of mating flight, each queen was carefully checked to make sure she hasn't mated, then, unmated queens were sampled. The whole bodies of the queens were sampled except the intestines, which were removed to prevent contamination. Finally, four courted virgin queens were sampled. At the same time, four non-courted virgin queens were directly sampled as control from the colonies. All these samples were stored at −80 °C until use.

Digital gene expression library preparation and sequencing

Total RNAs were isolated from each sampled queens, respectively. Then, total RNAs from four courted virgin queens and four non-courted virgin queens were respectively pooled, and each pooled RNA sample was divided into two portions. One portion was used to construct the DGE libraries, and another for small RNA sequencing. The protocols for DGE library construction and sequencing, and the steps and parameters for raw sequence filtration were the same as in our previous study (Wang et al., 2012). All the remained clean tags were deposited in the NCBI sequence read archive (SRR868713 for courted virgin queens and SRR868714 for non-courted virgin queens).

Mapping of DGE tags to genes and genome

Before mapping, two tag libraries containing all the possible CATG + 17 nt tag sequences were created using all the available mRNA sequences and genome sequences of *A. mellifera* downloaded from Genbank database (ftp://ftp.ncbi.nih.gov/genomes/Apis_mellifera). Then all the clean tags were first mapped to the tag database of reference mRNA sequences allowing only one nucleotide mismatch. Clean tags that mapped to multiple genes were filtered. The remained clean tags were designed as unambiguous clean tags. For gene expression analysis, the number of unambiguous clean tags for each gene was calculated and normalized to TPM (number of transcripts per million clean tags). Those tags that cannot be mapped to any gene in the tag database of mRNA sequences were then mapped to the tag database of the reference genome sequence.

Evaluation of DGE libraries

To identify the differentially expressed genes between the CVQ and NCVQ libraries, a rigorous statistical algorithm was developed, using the method described by Audic and Claverie (1997), to statistically analyze the tag frequency in each DGE library. The false discovery rate (FDR) was used to determine the threshold p-value in multiple tests. A FDR ≤ 0.001 and an absolute E-value of the $\log_2$ (CVQ/NCVQ) ≥ 1 were used as the threshold to determine significant differences in gene expression. The identified differentially expressed genes were used for GO (Gene Ontology) and KO (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis.

GO enrichment analysis firstly maps all differentially expressed genes to GO terms in the database (http://www.geneontology.org/) by BLASTX alignments, calculating gene numbers for each term, then using hypergeometric test to find significantly enriched GO terms in differentially expressed genes comparing to the reference gene background. The following formula was used to calculate the p-value of significantly enriched GO terms:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

where N is the number of all genes with GO annotation; n is the number of differentially expressed genes in N; M is the number of all genes that are annotated to the certain GO terms; and m is the number of differentially expressed genes in M. The Bonferroni Correction was used to correct the p-value. GO terms with corrected p-value ≤ 0.05 are defined as significantly enriched in differentially expressed genes.

KEGG pathway enrichment analysis is similar to GO analysis, it firstly maps all differentially expressed genes to terms in the KEGG database (http://www.genome.jp/kegg/) by BLASTX alignments. Then use hypergeometric test to find significantly enriched terms in differentially expressed genes comparing to the whole reference gene background. The formula for calculating p-value of significantly enriched KEGG terms is the same as that of GO analysis. The Q-value (false discovery rate) was used to determine the threshold p-value in multiple tests. Q-value ≤ 0.05 is defined as significantly enriched terms.

Construction of small RNA libraries and high-throughput sequencing

To construct small RNA libraries, total RNAs were from the same queen samples as in the DGE experiment. Then, the total RNAs were subjected to 15% (w/v) denaturing by 15% TBE-Urea gel and the small RNA fragments of 10–36 nt were isolated from the gel and purified. Then, a 3′ adaptor and a 5′ adaptor were sequentially ligated to the purified small RNA molecules. Subsequently, the ligation products were reverse transcribed into cDNA and subjected to PCR amplification. The amplification products were purified using 6% TBE-Urea gel. The purified DNA fragments were used for sequencing by Hiseq™ 2000.

Data cleaning and length distribution

The sequencing-received raw image data was converted into sequence data by the base calling step. Such sequence data called raw data or raw reads and stored them in a ".fq" file. The raw data is processed by getting rid of low quality reads, reads with 5′ primer contaminants or with poly A, reads without 3′ primer or insert tag, and reads shorter than 10 nt. The remained clean tags were deposited in the NCBI sequence read archive (SRR869567 for courted virgin queens and SRR869568 for non-courted virgin queens).

Then, all the small RNA tags were mapped to the *A. mellifera* genome by SOAP software to analyze their expression and distribution on the genome.