Genome-wide characterization of coding and non-coding RNAs in the ovary of honeybee workers and queens

Xiao CHEN¹, Wei SHI^{1,2}

¹Institute of Apicultural Research, Chinese Academy of Agricultural Sciences, Beijing, China

²Xiangshan Beigou No. 1, Haidian District, Beijing 100093, China

Received 10 September 2019 – Revised 24 February 2020 – Accepted 30 March 2020

Abstract – Adult honeybee queens and workers drastically differ in ovary state and ovary size. However, this reproductive bias is only partially understood from the view of a single RNA type. In this study, we predicted 10,271 mRNAs, 7235 lncRNAs, 11,794 circRNAs, and 164 miRNAs in the ovary of honeybee workers through bioinformatics. Combining RNA sequencing data of honeybee virgin queens, 4385 mRNAs, 2390 lncRNAs, 5602 circRNAs, and 75 miRNAs were differentially expressed in workers compared with virgins. Compared with egg-laying queens, 6536 mRNAs, 3130 lncRNAs, 5751 circRNAs, and 81 miRNAs were differentially expressed in workers. Further, functional annotation revealed that neural regulation was closely related to ovary state. Moreover, the potential interactions among circRNAs, miRNAs, lncRNAs, and mRNAs revealed that vitellogenin, ecdysone-induced protein 74, *ame_circ_0001176*, and *ame_circ_0001243* might play critical roles in the competing endogenous RNA network. These findings suggest that the integrative RNA networks have potential effects in ovarian phenotype differences in honeybees.

Apis mellifera ligustica / ovary / caste difference / coding RNA / non-coding RNA

1. INTRODUCTION

Honeybee (*Apis mellifera ligustica*) female ontogenies are governed by identical genomes (Macedo et al. 2016). However, differences in the quality of the food offered to young larvae trigger specific developmental pathways, producing highly specialized female phenotypes: queens and workers (Macedo et al. 2016). Adult honeybee queens and workers drastically differ in ovary size and ovarian state. Honeybee queens are equipped with huge ovaries (150–200 ovarioles). And after mating, the queen's ovaries become activated and even bigger. Then, she initiates egg

laying. By contrast, the worker bees start with queen like ovaries and during development degeneration occurs leading to the worker ovary phenotype, i.e., small ovaries (1–20 ovarioles) and non-activated ovaries (Graham et al. 2011). The queen is capable of laying over 1000 eggs per day, whereas workers typically do not lay eggs in the presence of a queen (Winston 1987). This adult ovarian phenotype difference between queens and workers is due to different developmental pathways (Ashby et al. 2016; Duncan et al. 2016; Macedo et al. 2016). Characterizing caste-specifically expressed genes in ovaries of honeybees has advanced the knowledge on gene regulatory mechanism associated with ovarian phenotype differences in the honeybees (Humann and Hartfelder 2011; Shi et al. 2015; Lago et al. 2016). For instance, microRNAs (miRNAs) were identified and predicted to be involved in caste-independent ovarian activity in queens and

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13592-020-00760-7>) contains supplementary material, which is available to authorized users.

Manuscript editor: Klaus Hartfelder

workers. Ame-miR-184 and ame-miR-315 have been reported to play important roles in ovary development and caste determination in honeybees (Ashby et al. 2016; Macedo et al. 2016). Ame-miR-14 and ame-miR-8 have been suggested to be associated with juvenile hormones (*JH*) and ecdysteroids (*Ec*), which play key roles in ovary development and caste determination in honeybees (Riddiford 1994; Hoover et al. 2003; Flatt et al. 2005). However, the gene regulatory mechanism of ovarian phenotype differences in honeybees is only partially understood from the view of a single RNA type, i.e., mRNA or miRNA (Humann and Hartfelder 2011; Guo et al. 2013; Nunes et al. 2013; Shi et al. 2015; Ashby et al. 2016; Lago et al. 2016; Macedo et al. 2016).

Recent studies reported that RNA–RNA crosstalk plays critical roles in the gene regulation. The RNA–RNA crosstalk includes new layers of gene regulation which involve in interactions between different RNA species (Tay et al. 2014). One of the best studied examples is the posttranscriptional regulation of RNA transcripts by miRNAs (Tay et al. 2014). Because miRNA-binding sites exist on a wide variety of RNA transcripts, all RNA transcripts that contain miRNA-binding sites can communicate with each other and regulate each other by competing specifically for shared miRNAs. These RNA transcripts, including not only coding RNAs but also non-coding RNAs, act as competing endogenous RNAs (ceRNAs) (Salmena et al. 2011; Tay et al. 2014). Thus, the ceRNA crosstalk extends beyond the non-coding RNAs, potentially confers an additional non-protein-coding function to protein coding mRNAs, and further regulates the expression of RNA and physiological processes (Sen et al. 2014; Tay et al. 2014). Studies showed that the ceRNA crosstalk network played important roles in various biological processes, i.e., tumorigenesis, cancer progression, and ovarian function (Hu et al. 2017; Jia et al. 2019).

The ceRNAs include various types of RNAs, such as long non-coding RNAs (lncRNAs), circular RNAs (circRNAs) and messenger RNAs (mRNAs). lncRNA is a class of non-coding RNA transcripts longer than 200 nucleotides (Kapranov et al. 2007). Previous study showed

that lncov1 might be involved in the autophagic cell death of ovarioles during worker embryogenesis, and lncov1 might affect ovary development in honeybees (Humann and Hartfelder 2011). Our previous study also showed that lncRNAs were involved in ovary development in honeybees (Chen et al. 2017). Moreover, lncRNAs can be targeted by miRNAs and further regulate the expression of mRNAs (Fan et al. 2015; Gong et al. 2016).

CircRNA is a type of RNA molecule that recently joined the growing list of known non-coding RNAs types. CircRNAs are circularized through 5–3' linkages as well as through 2–5' linkages formed by branch point nucleophilic attack during splicing (Lasda and Parker 2014). CircRNAs show greater stability, and therefore, circRNAs are better conserved among mammalian cells than miRNAs and lncRNAs (Hansen et al. 2013). In addition, circRNAs have been demonstrated to be associated with the gene expression regulation of various biological processes (Memczak et al. 2013). Our previous study showed that the expression of circRNAs related to ovarian activation in honeybees (Chen et al. 2018).

The difference in caste development is controlled by some remarkable features. For instance, hormones and genes are predicted to pleiotropically influence caste differences (Lago et al. 2016; Macedo et al. 2016; Pandey and Bloch 2015). CeRNA crosstalk includes new layers of gene regulation that involve interactions between diverse RNA species (Tay et al. 2014). Therefore, it is important to construct a ceRNA network to explore the molecular mechanisms underlying the caste difference of the honeybee's ovary. In this study, we examined the miRNA, lncRNA, circRNA, and mRNA expression profiles in ovaries of worker bees using high-throughput sequencing method. Then, we downloaded RNA-sequenced data of honeybee virgin queens and egg-laying queens, and compared the RNA expression patterns of honeybee workers with that of honeybee virgins and queens to identify candidate genes and/or RNAs that contribute to caste differences. Furthermore, a ceRNA network was constructed using bioinformatics approaches to explore the interactions between coding RNAs and

non-coding RNAs. The objective of this study is to identify coding and non-coding RNAs in workers, and provide valuable information on ovarian phenotype differences in honeybees.

2. MATERIALS AND METHODS

2.1. Sampling

All samples were obtained from *Apis mellifera ligustica* colonies. Pools of workers' ovaries were used in this study (number of pools = 3). Ovaries of worker bees sampling were carried out according to the procedure of Macedo, with minor modifications (Macedo et al. 2016). Briefly, 3 colonies with sister queens were reared using standard beekeeping techniques (Harbo 1986). The frames with emerging broods from these colonies were placed in an incubator (34 °C and 80% relative humidity). Newly emerged workers were paint marked and returned to their original colony. After 14 days, the marked workers were collected. Their ovaries were dissected, snap frozen in liquid nitrogen and stored at − 80 °C. Ovaries from 40 workers were in a pool. These 40 workers were from one single colony and their ovaries were all non-activated ovaries. Each biological sample had 40 ovary pairs and this is equivalent to a pool. Totally, three samples were used for RNA sequencing. Sample information of virgins and egg-laying queens are briefly described as follows. For the virgins, 6-day-old virgins were used. Ovaries of virgins were not activated. For egg-laying queens, activated ovaries are collected at the third day after they initiated laying eggs.

2.2. RNA isolation and sequencing

Total RNA was extracted from ovary samples using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. After quality control, a total amount of 10 µg RNA per sample was used for library preparation.

The preparation of whole transcriptome libraries and the sequencing were performed by the Annoroad Gene Technology Corporation (Beijing, China). LncRNA and mRNA library preparation was carried out using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. Paired-end reads of 150 bp were generated using the Illumina Hiseq 4000 platform. Small RNA library preparation was carried out using NEBNext®

Multiplex Small RNA Library Prep Set for Illumina® (NEB, USA) following manufacturer's recommendations. Single-end reads of 50 bp were generated using the Illumina Hiseq 2500 platform. CircRNA library preparation was carried out using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. Paired-end reads of 150bp were generated using the Illumina Hiseq 4000 platform. After quality control, clean reads were aligned to the reference genome (*Amel_4.5*). For RNA identification, mRNAs, lncRNAs, miRNAs, and circRNA were identified by Chen's methods (Chen et al. 2017, 2018). The classification, genome distribution, length distribution, exon number, and expression level of all RNAs were evaluated. Concerning the exon number of RNAs, lncRNA and circRNAs were analyzed. One category of lncRNA, sense lncRNAs, expressed from the same direction of the adjacent protein coding genes (Kour and Rath 2016). The location of sense lncRNAs usually overlap with exons of the coding genes. So sense lncRNAs have exons. CircRNAs are divided into four categories: exonic circRNAs, circular RNAs from introns, exon-intron circRNAs, and intergenic circRNAs (Meng et al. 2017). Exonic circRNAs are predominantly generated from back-spliced exons, where 3' splice donors of the pre-mRNA are covalently linked to 5' splice acceptors in reverse order (Meng et al. 2017). Exon-intron circRNAs are a type of circular RNA that is circularized with both exons and introns at the same time (Meng et al. 2017). So, exonic circRNAs and exon-intron circRNAs have exons. Complete methods are provided in [Supporting Information Materials and Methods](#).

2.3. Differentially expressed mRNA, lncRNA, miRNA, and circRNA identification

To analyze the expression levels of RNAs in different castes, we downloaded the lncRNA, mRNA, miRNA, and circRNA sequencing data of honeybee virgin queens and egg-laying queens from the public database GEO with accession number GSE93028 and GSE103860. The downloaded data are high-throughput ovary sequences of virgin queens and egg laying. Differentially expressed RNAs (Benjamini & Hochber method corrected p value < 0.05 and fold change > 1) were identified using DESeq R package (v2) for each of the following comparisons: (1) virgin queens vs workers; (2) egg-laying queens vs workers.

2.4. Prediction of lncRNA and miRNA target genes

The potential trans role of lncRNAs (acting on non-neighboring genes) can be assessed by correlating expression levels between lncRNAs and mRNAs. The transcriptional role of lncRNAs on coding genes was examined based on the correlation coefficient of expression (Pearson correlation > 0.9). To predict miRNAs targets, we searched for the targets in the 3'UTR of gene models. For genes lacking a predicted 3'UTR, the region 1000 bp downstream of the stop codon were included. The prediction was performed with Miranda with the following parameter: free energy < -20 kcal/mol and score > 160 (Enright et al. 2003).

2.5. Functional enrichment analysis

The potential functions of differentially expressed mRNA, lncRNA, miRNA, and circRNAs were analyzed by Gene ontology (GO) and KEGG pathway functional annotation. Gene ontology (GO) enrichment analysis was implemented by Goseq R package (Young et al. 2012). KEGG pathways analysis was performed using KOBAS to determine the involvement of genes in different biological pathways (Mao et al. 2005).

2.6. Chromosomal localization of the differentially expressed lncRNAs and mRNAs in QTL for ovary size

The localizations of the differentially expressed lncRNAs and mRNAs on *Apis mellifera* chromosomes were accessed from NCBI database (*Amel_4.5*). Each RNA location was estimated in centimorgan and was compared with the location of a significant quantitative trait locus (QTL) previously identified for ovary size. This QTL locates on chromosome 11 (Chr11) between the position 8.9 Mb and 12.2 Mb (Linksvayer et al. 2009; Graham et al. 2011). Genes or RNAs which locate within the QTL confidence intervals were accepted as candidate genes for ovary size.

2.7. Construction of the ceRNA network

In order to make clear the roles of lncRNA, mRNA, circRNA, and miRNA with mediated ceRNA network, we built the co-expression network of lncRNA/circRNA-miRNA-mRNA interactions. The mRNA/miRNA, lncRNA/miRNA, and circRNA/miRNA interactions were predicted based on miRanda (Enright et al. 2003). The

ceRNA network was visualized using Cytoscape v3.5.0 software (Smoot et al. 2011).

2.8. Real-time PCR

In order to confirm sequencing results, the expression of 5 circRNAs, 5 lncRNAs, 5 mRNAs, and 5 miRNAs were validated by real-time PCR using the same ovary samples used for sequencing. Data normalization of lncRNA, mRNA, and circRNA were carried out using the Actin reference gene (gene id: 406122). Data normalization of miRNA was carried out using the U6 reference gene. The correlation between the results of sequencing and PCR was calculated using correlation test.

2.9. Data availability

The authors state that all data necessary for confirming the conclusions presented in the manuscript are represented fully within the manuscript. All the sequence data are available through the GEO database with accession number GSE119256, GSE93028, and GSE103860.

3. RESULTS

3.1. Prediction of coding and non-coding RNAs in ovaries of honeybee workers

Sequencing of all mRNA and lncRNA libraries generated 249,438,672 raw paired-end reads. We obtained 192,919,168 clean reads from workers' ovaries. We detected 10,271 mRNAs, 2805 known lncRNAs, and 4433 novel lncRNAs (Supplemental Table S1).

Sequencing of all miRNA libraries generated 39,997,807 raw single-end reads. We obtained 21,789,917 clean reads from workers' ovaries and detected 152 known miRNAs and 12 novel miRNAs (Supplemental Table S2).

Sequencing of all circRNA libraries generated 286,102,258 raw single-end reads. We obtained 271,334,086 clean reads from the workers ovaries and detected 11,794 circRNAs (Supplemental Table S3).

3.2. Properties of coding and non-coding RNAs in ovaries of honeybee workers

The classification, genome distribution, length distribution, exon number, and expression level of all lncRNAs in workers' ovaries were evaluated. The bio-type of the novel lncRNAs were antisense (41%) and intergenic transcripts (59%). According to the genome distribution of lncRNAs, the largest proportion of lncRNAs was transcribed from Chr1 (Figure 1a). Characterization of the genomic structure revealed that 93% of lncRNAs had 2–5 exons (Figure 1b). The mean length of lncRNAs were 1881 nt (range from 200 to 20,674). The properties of lncRNAs did not show significantly difference in ovaries of workers, virgin queens and egg-laying queens.

In workers, the miRNA sequence length distribution was mainly concentrated at 20 to 23 nt (90% miRNAs, ANOVA, $p < 0.01$) (Figure 2a). However, in virgins and egg-laying queens, the miRNA sequence length distribution was mainly concentrated at 20 to 22 nt (82% miRNAs, ANOVA, $p < 0.01$) (Figure 2a). According to the genome distribution of miRNAs in workers, the largest proportion of miRNAs was transcribed from Chr15 (ANOVA, $p < 0.05$) (Figure 2b). However, in virgins and egg-laying queens, the largest proportion of miRNAs was transcribed from Chr2 (ANOVA, $p < 0.05$) (Figure 2b). But the numbers of miRNAs transcribed from Chr2 and Chr15 are very close in all three groups (workers, virgins and egg-laying queens).

The majority (> 80%) of the circRNAs is consisted of protein-coding exons only, whereas the smaller fractions also contained introns, non-coding RNA gene regions, and unannotated intergenic regions (Figure 3a–c). According to the host gene locations, circRNAs are widely distributed across all chromosomes (Figure 3d). Chr1 produces the most circRNAs (Figure 3e). Interestingly, the number of circRNAs detected on each chromosome in workers is always more than that in virgins and queens. We found that the majority (> 57%) of the exon circRNAs were composed of 2–3 exons within the same parental gene (Figure 3d). In our study, 71% (1657/2321) of the parental genes have the ability to produce more than one circRNA. In comparison, about 40% of

the parental genes have the ability to produce more than one circRNA in cattle (Zhang et al. 2016).

3.3. Expression profiles of mRNAs, lncRNAs, miRNAs, and circRNAs in ovaries of honeybee workers

The RNA expression profiles in workers, virgins, and egg-laying queens were analyzed. Between workers and virgins, 4385 mRNAs show differential expression. Of these, 2165 are up-regulated and 2220 are down-regulated in workers relative to virgins. The differential expression of 6536 mRNAs (3471 up-regulated and 3065 down-regulated) is observed in workers relative to egg-laying queens (Figure 4). Between workers and virgins, 2390 lncRNAs show differential expression (865 up-regulated and 1525 down-regulated). The differential expression of 3130 lncRNAs (1789 up-regulated and 1341 down-regulated) is observed in workers relative to egg-laying queens (Figure 4). Between workers and virgins, 75 miRNAs show differential expression. Of these, 42 are up-regulated and 33 are down-regulated in workers relative to virgins. The differential expression of 81 miRNAs (39 up-regulated and 42 down-regulated) is observed in workers relative to egg-laying queens (Figure 4). Between workers and virgins, 5602 circRNAs show differential expression. Of these, 1614 are up-regulated and 3988 are down-regulated in workers relative to virgins. The differential expression of 5751 circRNAs (1545 up-regulated and 4206 down-regulated) is observed in workers relative to egg-laying queens (Figure 4). A full list of the differentially expressed circRNAs, lncRNAs, miRNAs, and mRNAs is shown in Supplemental Tables S4–S7.

3.4. Functional analysis

For mRNAs, relative to virgins and egg-laying queens, workers show enrichment in genes associated with immune process, reproductive process, multicellular organismal process, and molecule catabolic process (Table 1 and Supplemental Table S8). For lncRNAs, relative to virgins and egg-laying queens, workers show enrichment in genes associated with gene expression, immune and metabolic processes, including

Figure 1. Profiling of lncRNAs in ovary of honeybees. **a** Distribution of all lncRNAs on *Apis mellifera* genome. **b** Distribution of exons found within lncRNAs. The criteria is “exon number ≥ 2 ”.

Figure 2. Profiling of miRNAs in ovary of honeybees. **a** Length distribution of all miRNAs. **b** Distribution of all miRNAs on *Apis mellifera* genome.

Figure 3. Profiling of circRNAs in ovary of honeybees. **a** Categories and distribution of circRNAs in honeybee workers. **b** Categories and distribution of circRNAs in honeybee virgin queens. **c** Categories and distribution of circRNAs in honeybee egg-laying queens. **d** Distribution of exons found within circRNAs. **e** Distribution of all circRNAs on *Apis mellifera* genome.

posttranscriptional regulation of gene expression, positive regulation of immune response and sterol metabolism (Table I and Supplemental Table S9). For miRNAs and circRNAs, relative to virgins and egg-laying queens, workers show enrichment in genes associated with issue development, including neuron differentiation, anatomical structure development, animal organ development (Table I, Supplemental Tables S10 and S11).

In addition, differentially expressed RNAs enrichment ($p < 0.05$) was seen in KEGG pathways (Supplemental Tables S8–S11). Compared with virgins and egg-laying queens, mRNAs in

workers showed enrichment in metabolism pathways, including tryptophan metabolism and ascorbate and aldarate metabolism (Figure 5 and Supplemental Table S8). Compared with virgins and egg-laying queens, miRNAs and circRNAs in workers showed enrichment in cell signaling pathways, including hippo signaling pathway and Wnt signaling pathway (Figure 5 and Supplemental Tables S10 and S11).

However, no pathways were significantly enriched with the differentially expressed lncRNAs in workers compared with virgin queens and egg-laying queens.

Figure 4. The Venn diagram of the differentially expressed RNAs. V, virgin queens; Q, egg-laying queens; W, worker bees.

3.5. Chromosomal localization of the differentially expressed lncRNAs and mRNAs in QTL region for ovary size

If the differentially expressed lncRNAs and mRNAs were found located within the confidence interval of the QTL for ovary size, they could be regarded as candidate genes for ovary size (Linksvayer et al. 2009; Graham et al. 2011). In this way, 189 candidate genes and 88 lncRNAs were identified (Supplemental Table S12).

3.6. Construction of the lncRNA/circRNA-miRNA-mRNA co-expression network

According to ceRNA hypothesis, RNA transcripts effectively communicate with one another. In this study, mRNA-miRNA, lncRNA-miRNA, and circRNA-miRNA pairs were constructed. Then, a ceRNA network was constructed based on the lncRNA/circRNA/mRNA-miRNA pairs which contained at least one differentially expressed RNAs (Supplemental Table S13). In the network, 173 lncRNAs were predicted to share binding sites with 49 miRNAs and 1078 mRNAs bind 55 miRNAs. Considering that graphics cannot display the enormous amount of network information, we selected miRNAs, whose expression correlated with ovary state in honeybees, to make the network diagram (Figure 6 and Supplemental Table S13). Figure 6 shows the RNA-RNA crosstalk of 55 miRNAs with 351 lncRNAs, 38 circRNAs, and 1865 target genes.

This result implies that the differentially expressed non-coding RNAs may influence the expression of target genes and further affect biological processes by sponging miRNAs. Through functional annotation analysis, mRNAs in ceRNA networks were predicted to be involved in the hippo signaling pathway, Wnt signaling pathway, notch signaling pathway, ubiquitin-mediated proteolysis, inositol phosphate metabolism, insulin resistance and Fanconi anemia pathway (Supplemental Table S14).

3.7. Quantitative RT-PCR validation

In order to validate the sequencing results, the expression of 5 mRNAs, 5 lncRNAs, 5 miRNAs, and 5 circRNAs were tested by using real-time PCR with the same RNA samples used for sequencing (Supplemental Table S15). The expression profiles of these RNAs detected by real time PCR were consistent with those obtained by sequencing, which confirmed the reliability of our sequencing results (Figure 7).

4. DISCUSSION

In honeybees, queens have large, activated ovaries and workers have small, non-activated ovaries. The genetic architecture of ovary phenotype has been studied in detail in *Drosophila* (Bergland et al. 2008; Orgogozo et al. 2006). Previous studies exploring the gene regulatory mechanism of ovarian phenotype differences in honeybees found that mRNAs and miRNAs played critical roles (Lago et al. 2016; Macedo et al. 2016). More recently, emerging evidence has indicated that coding and non-coding RNA interactions participate in gene expression by modulating transcription, protein, miRNA functions, and further affect biological processes (Sen et al. 2014; Tay et al. 2014). Moreover, our preliminary study found that coding and non-coding RNAs interactions play roles in ovary activation and egg-laying regulation in honeybees (Chen et al. 2017, 2018). In this study, we construct a hypothetical ceRNA network by analyzing differential expression of lncRNAs, circRNAs, miRNAs, and mRNAs. We try to extend our analysis towards a comprehensive understanding of the genetic

Table 1. Functional annotation results of the differentially expressed RNAs in the comparison between egg-laying queens and worker bees and the comparison between virgin queens and worker bees.

GO_biological process	GO id	P value	No. of genes	GO_biological process	GO id	P value	No. of genes
GO_mRNA_workers vs egg-laying queens				GO_mRNA_workers vs virgins			
Reproductive process	GO:0022414	0.000046	343	Response to xenobiotic stimulus	GO:0009410	8.8E-07	26
Regulation of multicellular organismal process	GO:0051239	0.000051	414	Cellular response to xenobiotic stimulus	GO:0071466	9.3E-07	25
DNA conformation change	GO:0071103	0.000079	60	Xenobiotic metabolic process	GO:0006805	9.3E-07	25
Single organism reproductive process	GO:0044702	0.000094	297	Small molecule catabolic process	GO:0044282	4.2E-06	68
Multicellular organismal process	GO:0032501	0.00014	1366	Response to DDT	GO:0046680	0.000013	15
Biological regulation	GO:0065007	0.00016	1897	Xenobiotic catabolic process	GO:0042178	0.000015	17
DNA packaging	GO:0006323	0.00016	46	Alpha-amino acid metabolic process	GO:1901605	0.000019	57
Single-organism developmental process	GO:0044767	0.00017	1234	Secondary metabolite catabolic process	GO:0090487	0.000034	14
Developmental process	GO:0032502	0.00017	1245	Insecticide catabolic process	GO:0046701	0.000034	14
Multicellular organism development	GO:0007275	0.00021	1071	Insecticide metabolic process	GO:0017143	0.000034	14
GO_lncRNA_workers vs egg-laying queens				GO_lncRNA_workers vs virgins			
Posttranscriptional regulation of gene expression	GO:0010608	0.00052	257	Cellular process	GO:0009987	0.00026	5587
Microtubule cytoskeleton organization	GO:0000226	0.00448	234	Epithelial cell migration	GO:0010631	0.00058	122
Positive regulation of immune response	GO:0050778	0.00497	59	Epithelium migration	GO:0090132	0.00058	122
Sterol metabolic process	GO:0016125	0.00544	58	Autophagy	GO:0006914	0.00087	147
Mitotic nuclear division	GO:0007067	0.00581	229	Organelle organization	GO:0006996	0.0009	1466
Synaptic vesicle localization	GO:0097479	0.00639	79	Single-organism organelle organization	GO:1902589	0.00127	1020
Positive regulation of cell death	GO:0010942	0.0072	115	Tissue migration	GO:0090130	0.00139	127
Synaptic vesicle transport	GO:0048489	0.00749	77	Positive regulation of immune response	GO:0050778	0.00143	59
Establishment of synaptic vesicle localization	GO:0097480	0.00749	77	Positive regulation of immune system process	GO:0002684	0.0021	92

Table 1 (continued)

GO biological process	GO id	P value	No. of genes	GO biological process	GO id	P value	No. of genes
Autophagy	GO:0006914	0.00787	147	Response to nutrient levels	GO:0031667	0.00211	188
GO miRNA_workers vs egg-laying queens				GO_miRNA_workers vs virgins			
Single-multicellular organism process	GO:0044707	1.00E-30	1051	Single-multicellular organism process	GO:0044707	3.40E-28	1017
Multicellular organismal process	GO:0032501	1.00E-30	1077	developmental process	GO:0032502	2.60E-27	955
Developmental process	GO:0032502	1.00E-30	999	Multicellular organismal process	GO:0032501	4.60E-27	1035
Multicellular organism development	GO:0007275	1.00E-30	874	single-organism developmental process	GO:0044767	2.30E-26	944
Single-organism developmental process	GO:0044767	1.00E-30	990	Multicellular organism development	GO:0007275	1.80E-25	831
Anatomical structure development	GO:0048856	4.30E-29	868	Anatomical structure development	GO:0048856	9.50E-22	824
System development	GO:0048731	1.70E-28	720	System development	GO:0048731	1.40E-21	682
Anatomical structure morphogenesis	GO:0009653	4.20E-28	550	Nervous system development	GO:0007399	1.90E-20	448
Biological regulation	GO:0065007	9.80E-27	1403	Single organism signaling	GO:0044700	2.60E-20	700
Neuron differentiation	GO:0030182	2.00E-26	310	Anatomical structure morphogenesis	GO:0009653	3.10E-20	514
GO_circRNA_workers vs egg-laying queens				GO_circRNA_workers vs virgins			
Biological regulation	GO:0065007	1.00E-30	1232	Neuron differentiation	GO:0030182	1.00E-30	216
Signaling	GO:0023052	1.00E-30	502	Single-multicellular organism process	GO:0044707	1.00E-30	662
Anatomical structure development	GO:0048856	1.00E-30	567	Multicellular organism development	GO:0007275	1.00E-30	544
Cellular developmental process	GO:0048869	1.00E-30	460	Biological regulation	GO:0065007	1.00E-30	1207
Cell development	GO:0048468	1.00E-30	319	Anatomical structure development	GO:0048856	1.00E-30	551
Single-organism developmental process	GO:0044767	1.00E-30	619	Developmental process	GO:0032502	1.00E-30	606
Cell communication	GO:0007154	1.00E-30	512	Animal organ development	GO:0048513	1.00E-30	349
Multicellular organism development	GO:0007275	1.00E-30	558	Single-organism developmental process	GO:0044767	1.00E-30	600
Animal organ development	GO:0048513	1.00E-30	354	Multicellular organismal process	GO:0032501	1.00E-30	673
Anatomical structure morphogenesis	GO:0009653	1.00E-30	372	System development	GO:0048731	1.00E-30	484

Figure 5. Heat map of enriched pathways of differentially expressed mRNAs, miRNAs, and circRNAs in two comparison groups. V, virgin queens; Q, egg-laying queens; W, worker bees.

Figure 6. The ceRNA network constructed with miRNAs which have known roles in reproduction in honeybees. Red triangle nodes represent circRNAs. Red ellipse nodes represent miRNAs. Light blue rectangle nodes represent mRNAs. Dark blue rectangle nodes represent mRNAs.

Figure 7. The correlation analysis between the result of sequencing and RT-PCR. $R > 0.8$ means significantly positive correlation.

architecture of the ovary trait in honeybees, and analyzed the different aspects of ovary phenotype by exploiting coding and non-coding RNAs interactions.

Our RNA sequencing library revealed that more than 98% of the detected mRNAs are expressed in all three groups (workers, virgins and egg-laying queens) (Figure 8). However, for non-coding RNAs, it is different. Our results showed that 95% of lncRNAs were detected in all three groups (Figure 8). And around 60% of the currently known honeybee miRNAs were

Figure 8. The Venn diagram of all identified RNAs. **a** The Venn diagram of all detected mRNAs. **b** The Venn diagram of all detected lncRNAs. **c** The Venn diagram of all detected miRNAs. **d** The Venn diagram of all detected circRNAs. V, virgin queens; Q, egg-laying queens; W, worker bees.

expressed in all three groups (Figure 8). For circRNAs, only around 27% of the identified honeybee circRNAs were expressed in all three groups (Figure 8). In specific, 5722 circRNAs showed specific expression in ovary of workers, while 1078 circRNAs showed specific expression in ovary of virgins and 1537 circRNAs showed specific expression in ovary of egg-laying queens (Figure 8). The results revealed that the difference in non-coding RNAs is more obvious in ovaries of workers than in virgins and egg-laying queens.

In our study, according to GO and KEGG analysis, no biological process related to ovary development or caste differences was significantly enriched with the differentially expressed mRNAs and lncRNAs, whereas many biological processes and pathways were significantly enriched with the differentially expressed miRNAs and circRNAs. Especially, the functional category neural regulation was highly represented for the differentially expressed miRNAs and circRNAs in workers compared with virgins and egg-laying queens. This is consistent with the reported enrichment of miRNAs in neural tissues of many species, where miRNAs appear to play a critical role in synaptic plasticity, axon guidance, as well as learning and memory (Ashby et al. 2016). For instance, miR-92, which was differentially expressed in the comparison between egg-laying queens and workers, were found associated with neural regulation in our results and also in previous studies (Macedo et al. 2016). Neural regulation was previously reported associated with caste differences in honeybee larvae (Ashby et al. 2016). Additionally, circRNAs were found closely relating to neural development and neural function in neural tissues in *Drosophila* (Westholm et al. 2014). Furthermore, recent findings have revealed that circRNAs can act as miRNA sponges to regulate gene expression, and furthermore protein synthesis [19, 36]. CircRNAs can also chelate miRNAs more efficiently (Hansen et al. 2013). Taken together, the circRNA-miRNA interactions in honeybees' ovary may play roles in neural development and neural function. Therefore, we presume that the differentially expressed circRNAs and miRNAs in our results may interact with each other by ceRNA-crosstalk and further affect ovary state through neural regulation in honeybees.

Further analyses of the interaction networks indicated that the differentially expressed coding and non-coding RNAs play key roles in honeybees’ ovary by regulating RNA expression. In our study, 1078 mRNAs were predicted to interact with 54 common miRNAs, along with 173 lncRNAs and 161 circRNAs. For example, in the network, vitellogenin (*Vg*) was targeted by miR-3747b, miR-375 and miR-9891. *Vg* is among the most important genes in regulating egg development in honeybees (Nunes et al. 2013). Meanwhile, miR-375 was reported to relate with caste differences (Guo et al. 2013). Furthermore, through the network analysis, we found that miR-3747b, miR-375, and miR-9891 targeted the gene ecdysone-induced protein 74 (*E74*). *E74* is *Ec* responsive genes and participates in *Ec* signaling regulating reproductive process (Pandey and Bloch 2015). *Ec* is one of the most important hormones in honeybees. Previous studies suggested that large amounts of *Ec* was produced in the ovary of queens and strongly influenced the ovary state (Flatt et al. 2005; Pandey and Bloch 2015). Additionally, in the network, *E74* was targeted by other miRNAs, including miR-34 and miR-92. MiR-34 has been reported to involve in caste differentiation in honeybees, whereas miR-92 was previously found negatively correlating with *Vg*. The results indicated that *Vg* and *E74* may communicate with and regulate each other by competing specifically for shared miRNAs—miR-3747b, miR-375, miR-9891, and miR-92. This ceRNA crosstalk would further extend the role of *Vg* and *E74* in ovary development and caste differences by the complex network.

In the ceRNA network, two circRNAs, *ame_circ_0001176*, and *ame_circ_0001243* are of particular interest. Both *ame_circ_0001176* and *ame_circ_0001243* were found significantly differently expressed in workers compared with virgins and egg-laying queens. In the ceRNA network, *ame_circ_0001176* was found to potentially interact with miR-124 and miR-8, whereas *ame_circ_0001243* was found to potentially interact with miR-1. MiR-124, miR-8, and miR-1 were all *Ec* responsive miRNAs. MiR-1 was also *Vg* responsive miRNAs and associated with reproductive status in honeybees (Nunes et al. 2013; Mello et al. 2014; Macedo et al. 2016). Therefore,

we deduced that *ame_circ_0001176* and *ame_circ_0001243* might be *Ec* or *Vg* responsive circRNAs by competing specifically for shared miRNAs - miR-124, miR-8, and miR-1. This circRNA-miRNA-mRNA crosstalk may further affect ovarian phenotype differences.

Through KEGG pathway analysis, mRNAs in the ceRNA network were predicted to involve in hippo signaling pathway, Wnt signaling pathway, notch signaling pathway, ubiquitin-mediated proteolysis, inositol phosphate metabolism, insulin resistance, and Fanconi anemia pathway. These pathways were previously recognized to involve in the regulation of workers’ ovary state (Ronai et al. 2016). Previous studies also demonstrated that hippo, Wnt and notch pathways involved in caste determination in honeybee larvae (Ashby et al. 2016). Moreover, the hippo signaling pathway can coordinate with Wnt and notch signaling pathways and further affect ovary development (Barry and Camargo 2013; Ye et al. 2017). Concerning ubiquitin-mediated proteolysis, although there is no report showing that it directly related with ovary development or caste differences, previous studies found that ubiquitin-mediated proteolysis promotes the activation of Wnt pathway (Chitalia et al. 2008). These results indicate that there is a crosstalk between hippo, Wnt, notch and ubiquitin-mediated proteolysis pathways. And this crosstalk is important in regulating of ovary development and caste differences. We hypothesize that this crosstalk may be built based on the ceRNA network. For the insulin resistance pathway, studies have shown that it may be involved in the growth and development of oocytes and also be essential for reproductive hormones in mammals (Navratil et al. 2009). So insulin resistance pathway may affect ovary development or caste differences by interacting with reproductive hormones. However, for inositol phosphate metabolism and Fanconi anemia pathways, there is no report showing them related with ovary development or caste differences.

5. CONCLUSION

In the present study, numerous of differentially expressed coding and non-coding RNAs were identified in worker bees compared with virgins

and egg-laying queens. These observations suggest that both coding and non-coding RNAs were involved in ovary development and caste differences related processes in honeybees. Further, the interactions between coding and non-coding RNAs were investigated through bioinformatics analysis, suggesting that such interactions might have potential effects in the gene regulatory mechanism of ovarian phenotype differences in honeybees.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: Xiao Chen and Wei Shi. Wrote the manuscript: Xiao Chen. Collection of the sample, RNA extraction and data analyze: Xiao Chen. ceRNA network constructing: Xiao Chen. Critically revised the manuscript: Xiao Chen and Wei Shi. All authors read and approved the final manuscript.

FUNDING INFORMATION

This work was supported by the National Natural Science Foundation of China (No. 31802145), the Agricultural Science and Technology Innovation Program (CAAS-ASTIP-2018-IAR), and the earmarked fund for Modern Agro-Industry Technology Research System (CARDS-45-KXJ1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors thank Feng Wei for his help in field beekeeping and sampling.

COMPLIANCE WITH ETHICAL STANDARDS

The apiaries for honeybees sample collection were maintained by Institute of Apicultural Research, Chinese Academy of Agricultural Sciences (IAR, CAAS), Beijing, China. The sample collection was approved by the ethics committee of IAR, CAAS. Bees were humanely sacrificed as needed to prevent suffering. No specific permits were required for the described studies.

Conflict of interest The authors declare that they have no conflict of interest.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Caractérisation à l'échelle du génome des ARN codants et non codants dans l'ovaire des ouvrières et des reines d'abeilles.

Apis mellifera ligustica / ovaire / différence de caste / ARN codant / ARN non codant.

Genomweite Charakterisierung von kodierenden und nichtkodierenden RNAs im Ovar von Arbeiterinnen und Königinnen der Honigbiene.

Apis mellifera ligustica / Ovar / Kastenunterschied / kodierende RNA / nichtkodierende RNA.

REFERENCES

- Ashby R., S. Forêt, I. Searle, R. Maleszka. (2016) MicroRNAs in Honey Bee Caste Determination. *Scientific reports* **6**: 18794.
- Barry E.R., F.D. Camargo. (2013) The Hippo superhighway: signaling crossroads converging on the Hippo/Yap pathway in stem cells and development. *Current opinion in cell biology* **25** (2): 247-253.
- Bergland A.O., A. Genissel, S.V. Nuzhdin, M. Tatar. (2008) Quantitative Trait Loci Affecting Phenotypic Plasticity and the Allometric Relationship of Ovariole Number and Thorax Length in *Drosophila melanogaster*. *Genetics* **180** (1): 567-582.

- Chen X., C. Ma, C. Chen, Q. Lu, W. Shi, et al. (2017) Integration of lncRNA–miRNA–mRNA reveals novel insights into oviposition regulation in honey bees. *PeerJ* **5** (12.9): e3881.
- Chen X., W. Shi, C. Chen. (2018) Differential circular RNAs expression in ovary during oviposition in honey bees. *Genomics* **online**. <https://doi.org/10.1016/j.ygeno.2018.03.015>.
- Chitalia V.C., R.L. Foy, M.M. Bachschmid, L. Zeng, M.V. Panchenko, et al. (2008) Jade-1 inhibits Wnt signalling by ubiquitinating beta-catenin and mediates Wnt pathway inhibition by pVHL. *Nature Cell Biology* **10** (10): 1208.
- Duncan E.J., O. Hyink, P.K. Dearden. (2016) Notch signalling mediates reproductive constraint in the adult worker honeybee. *Nature Communications* **7**.
- Enright A.J., B. John, U. Gaul, T. Tuschl, C. Sander, et al. (2003) MicroRNA targets in *Drosophila*. *Genome biology* **5** (1): R1.
- Fan C., Z. Hao, J. Yan, G. Li. (2015) Genome-wide identification and functional analysis of lincRNAs acting as miRNA targets or decoys in maize. *BMC genomics* **16** (1): 1–19.
- Flatt T., M.P. Tu, M. Tatar. (2005) Hormonal pleiotropy and the juvenile hormone regulation of *Drosophila* development and life history. *Bioessays* **27** (10): 999–1010.
- Gong Z., Y. Qian, Z. Zeng, W. Zhang, X. Li, et al. (2016) An integrative transcriptomic analysis reveals p53 regulated miRNA, mRNA, and lncRNA networks in nasopharyngeal carcinoma. *Tumor Biology* **37** (3): 1–13.
- Graham A.M., M.D. Munday, O. Kaftanoglu, R.E. Page, G.V. Amdam, et al. (2011) Support for the reproductive ground plan hypothesis of social evolution and major QTL for ovary traits of Africanized worker honey bees (*Apis mellifera* L.). *BMC evolutionary biology* **11** (1): 95.
- Guo X., S. Su, G. Skogerboe, S. Dai, W. Li, et al. (2013) Recipe for a busy bee: microRNAs in Honey Bee caste determination. *PLoS one* **8** (12): e81661.
- Hansen T.B., T.I. Jensen, B.H. Clausen, J.B. Bramsen, B. Finsen, et al. (2013) Natural RNA circles function as efficient microRNA sponges. *Nature* **495** (7441): 384–388.
- Harbo J.R. (1986) Propagation and Instrumental Insemination - Bee Genetics and Breeding - CHAPTER 15. *Bee Genetics & Breeding* **01** (02): 361–389.
- Hoover S.E.R., C.I. Keeling, M.L. Winston, K.N. Slessor. (2003) The effect of queen pheromones on worker honey bee ovary development. *Naturwissenschaften* **90** (10): 477–480.
- Hu T., X. Qi, F. Zhang, N. Zhang, L. Yang, et al. (2017) Circular RNA profiling reveals chi_circ_0008219 function as microRNA sponges in pre-ovulatory ovarian follicles of goats (*Capra hircus*). *Genomics* **110** (4): 257–266.
- Humann F.C., K. Hartfelder. (2011) Representational Difference Analysis (RDA) reveals differential expression of conserved as well as novel genes during caste-specific development of the honey bee (*Apis mellifera* L.) ovary. *Insect Biochemistry & Molecular Biology* **41** (8): 602–612.
- Jia Z., J. Peng, Z. Yang, J. Chen, L. Liu, et al. (2019) Long non-coding RNA TP73-AS1 promotes colorectal cancer proliferation by acting as a ceRNA for miR-103 to regulate PTEN expression. *Gene* **685**: 222–229.
- Kapranov P., J. Cheng, S. Dike, D. Nix, R. Duttagupta, et al. (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316** (5830): 1484–1488.
- Kour S., P.C. Rath. (2016) Long noncoding RNAs in aging and age-related diseases. *Ageing Research Reviews* **26**: 1–21.
- Lago D.C., F.C. Humann, A.R. Barchuk, K.J. Abraham, K. Hartfelder. (2016) Differential gene expression underlying ovarian phenotype determination in honey bee, *Apis mellifera* L, caste development. *Insect Biochemistry & Molecular Biology* **79**: 1–12.
- Lasda E., R. Parker. (2014) Circular RNAs: diversity of form and function. *Rna-a Publication of the Rna Society* **20** (12): 1829.
- Linksvayer T.A., O. Rueppell, A. Siegel, O. Kaftanoglu, Page R.E. Jr, et al. (2009) The genetic basis of transgressive ovary size in honeybee workers. *Genetics* **183** (2): 693.
- Macedo L.M.F., F.M.F. Nunes, F.C.P. Freitas, C.V. Pires, E.D. Tanaka, et al. (2016) MicroRNA signatures characterizing caste-independent ovarian activity in queen and worker honeybees (*Apis mellifera* L.). *Insect molecular biology* **25** (3): 216–226.
- Mao X., T. Cai, J.G. Olyarchuk, L. Wei. (2005) Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **21** (19): 3787–3793.
- Mello T.R.P., A.C. Aleixo, D.G. Pinheiro, F.M.F. Nunes, M.M.G. Bitondi, et al. (2014) Developmental regulation of ecdysone receptor (EcR) and EcR-controlled gene expression during pharate-adult development of honeybees (*Apis mellifera*). *Frontiers in genetics* **5**: 445.
- Memczak S., M. Jens, A. Elefsinioti, F. Torti, J. Krueger, et al. (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495** (7441): 333–338.
- Meng S., H. Zhou, Z. Feng, Z. Xu, Y. Tang, et al. (2017) CircRNA: functions and properties of a novel potential biomarker for cancer. *Molecular Cancer* **16** (1): 94.
- Navratil A.M., H. Song, J.B. Hernandez, B.D. Cherrington, S.J. Santos, et al. (2009) Insulin augments gonadotropin-releasing hormone induction of translation in LbetaT2 cells. *Molecular & Cellular Endocrinology* **311** (1–2): 47.
- Nunes F.M., K.E. Ihle, N.S. Mutti, Z.L. Simões, G.V. Amdam. (2013) The gene vitellogenin affects microRNA regulation in honey bee (*Apis mellifera*) fat body and brain. *Journal of Experimental Biology* **216** (Pt 19): 3724.
- Orgogozo V., K. Broman, D. Stern. (2006) High-Resolution Quantitative Trait Locus Mapping Reveals

- Sign Epistasis Controlling Ovariole Number Between Two *Drosophila* Species. *Genetics* **173** (1): 197-205.
- Pandey A., G. Bloch. (2015) Juvenile hormone and ecdysteroids as major regulators of brain and behavior in bees. *Current Opinion in Insect Science* **12**: 26-37.
- Riddiford L.M. (1994) Cellular and Molecular Actions of Juvenile Hormone I. General Considerations and Premetamorphic Actions. *Advances in Insect Physiology* **24**: 213-274.
- Ronai I., V. Vergoz, B. P. (2016) The mechanistic, genetic, and evolutionary basis of worker sterility in the social Hymenoptera. *Advances in the Study of Behavior* **48**: 251-317.
- Salmena L., L. Poliseno, Y. Tay, L. Kats, P.P. Pandolfi. (2011) A ceRNA hypothesis: the Rosetta stone of a hidden RNA language? *Cell* **146** (3): 353-358.
- Sen R., S. Ghosal, S. Das, S. Baiti, J. Chakrabarti. (2014) Competing Endogenous RNA: The Key to Posttranscriptional Regulation. *The Scientific World Journal*, 2014, (2014-2-2) **112** (1): 89-98.
- Shi Y.Y., H.J. Zheng, Q.Z. Pan, Z.L. Wang, Z.J. Zeng. (2015) Differentially expressed microRNAs between queen and worker larvae of the honey bee (*Apis mellifera*). *Apidologie* **46** (1): 35-45.
- Smoot M.E., K. Ono, J. Ruscheinski, P.L. Wang, T. Ideker. (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27** (3): 431-432.
- Tay Y., J. Rinn, P.P. Pandolfi. (2014) The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505** (7483): 344-352.
- Westholm J.O., P. Miura, S. Olson, S. Shenker, B. Joseph, et al. (2014) Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell reports* **9** (5): 1966-1980.
- Winston M.L. (1987) *The Biology of the Honey Bee*. Harvard University Press, Cambridge.
- Ye H., X. Li, T. Zheng, C. Hu, Z. Pan, et al. (2017) The Hippo Signaling Pathway Regulates Ovarian Function via the Proliferation of Ovarian Germline Stem Cells. *Cellular Physiology & Biochemistry International Journal of Experimental Cellular Physiology Biochemistry & Pharmacology* **41** (3): 1051.
- Young M.D., M.J. Wakefield, G.K. Smyth, A. Oshlack. (2012) *goseq*: Gene Ontology testing for RNA-seq datasets.
- Zhang C.L., W. Hui, Y.H. Wang, S.Q. Zhu, J.Q. Liu, et al. (2016) Circular RNA of cattle casein genes are highly expressed in bovine mammary gland. *Journal of Dairy Science* **99** (6): 4750-4760.

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