OXFORD

On the Molecular Basis of Division of Labor in *Solenopsis invicta* (Hymenoptera: Formicidae) Workers: RNA-seq Analysis

Hua-Long Qiu,[1],* Cheng-Yin Zhao,[3],* and Yu-Rong He[1,2]

[1]Department of Entomology College of Agriculture, South China Agricultural University, Guangdong, Guangzhou 510642, China (qiuhualong2008@163.com), [2]Corresponding author, e-mail:yrhe@scau.edu.cn.com, and [3]Department of Life Science Luoyang Normal University, Henan, Luoyang 471000, China

*These authors contributed equally to this work.

Abstract

The fire ant *Solenopsis invicta* Buren is an important invasive pest. Among *S. invicta* workers behavioral changes depend on age where younger ants are nurses and older ants foragers. To identify potential genes associated with this division of labor, we compared gene expression between foragers and nurses by high-throughput sequencing. In total, we identified 1,618 genes significantly differently expressed between nurses and foragers, of which 542 were upregulated in foragers and 1,076 were upregulated in nurses. Several pathways related to metabolism were significantly enriched, such as lipid storage and fatty acid biosynthesis, which might contribute to the division of labor in *S. invicta*. Several genes involved in DNA methylation, transcription, and olfactory responses as well as resistance to stress were differentially expressed between nurses and foragers workers. Finally, a comparison between previously published microarray data and our RNA-seq data in *S. invicta* shows 116 genes overlap, and the GO term myofibril assembly (GO: 0030239) were simultaneously significantly enriched. These results advance knowledge of potentially important genes and molecular pathways associated with worker division of labor in *S. invicta*. We hope our dataset will provide . candidate target genes to disrupt organization in *S. invicta* as a control strategy against this invasive pest

Key words: biological invasions, eusociality, myrmecology, transcriptomics

Social insects are characterized by a marked division of labor (Smith et al. 2008, Duarte et al. 2011). Reproductive division of labor in social insects is a hallmark of their evolution, in which queens lay eggs and generated workers will care for the brood and undertake all the other tasks (Hölldobler 1990, Lach et al. 2010). In addition, labor division in eusocial insects also typically involves age-dependent polyethism among workers, who will switch from various indoor duties to foraging outside the colony as they grow older (Wilson 1971, Mirenda and Vinson 1981). Specifically, workers start their adult lives among the brood pile acting as nurses. As they age they will gradually move their activities to the nest periphery, finally starting foraging outside the nest (Tschinkel 2006). Colony members will benefit from labor division specialization, which is probably a key aspect behind the ecological success of eusocial insects (Hölldobler 1990). The regulatory mechanisms in division of labor in social insects is thus a main topic of research (Smith et al. 2008). Many factors have been demonstrated to play a role in workers switching from nurses to foragers, such as developmental and nutritional factors, morphology, gained experience, and genetic variation (Libbrecht et al. 2013 and authors therein).

The red imported fire ants, *Solenopsis invicta* Buren (Formicidae: Myrmicinae) are eusocial insects noted for their invasive capacity, painful stings, as well as conspicuous nest mounds. Fire ant workers are characterized by age-dependent division of labor (Mirenda and Vinson 1981). The transition from nurse to forager typically occurs at about 30–50 days of age, and is associated with many changes in physiology and gene expression (Tschinkel 2006). Social behaviors of *S. invicta* are complex but experimentally accessible (Qiu et al. 2014, 2015). Due to being one of the most extensively studied ants, the fire ants are considered a model for the study of behavior, physiology, genomics, and genetic regulation of social organization.

Microarray and RNA-seq studies have been used to identify potential genes and molecular pathways associated with division of labor in social insect colonies (Kucharski and Maleszka 2002, Feldmeyer et al. 2014, Manfredini et al. 2014, Berens et al. 2015, Harrison et al. 2015, Morandin et al. 2015). Recent studies based

on microarray and RNA-seq technologies suggest that some genes and especially molecular pathways and gene networks (e.g., metabolism and nutrition) involved in the division of labor in social insects are convergent among the different lineages, honeybees, wasps, ants (Markiewicz and O'Donnell 2001, Toth and Robinson 2005, Lattorff and Moritz 2013, Berens et al. 2015). In *S. invicta*, the molecular basis for worker division of labor has been characterized using microarray (Manfredini et al. 2014). Numerous genes were differentially expressed between nursing and foraging workers, mainly enriched among some pathways and GO terms. However, all data currently available on *Solenopsis* castes come from microarray platforms. Here, we use RNA-seq by Illumina Genome Analyzer to reassess gene expression differences between foragers and nurses of *S. invicta*. RNA-seq not only provides a wide and unbiased coverage of entire genes, it is capable also of identifying differentially expressed genes (DEGs) without annotation information (Morozova and Marra 2008, Wang et al. 2009). We hypothesized that gene pathways involved in *S. invicta* worker division of labor would be similar to those described from other hymenopteran social insects. As far as we know, this is the first study comparing the transcriptome of fire ants by high-throughput RNA sequencing. Our study provides important information on how critical molecules and metabolic pathways affect worker behavior and task transition at the transcriptional level.

Materials and Methods

Ant Rearing and Sample Preparation

Three polygyne *S. invicta* colonies were collected from South China Agricultural University (SCAU, Guangzhou, China) campus in May 2014. Colonies were reared inside plastic boxes (50 cm × 40 cm × 15 cm) under controlled laboratory conditions of 25 ± 2°C, 75 ± 5 relative humidity and 12:12 (L:D) h light. Colonies were fed with *Tenebrio molitor* larvae and 25% sucrose water every other day. Two groups of workers were sampled: foragers from outside the nest (foraging in the food area) and nurses from inside the nest (caring for the brood). We collected 10 foragers and nurses from each of the three colonies at around 10:00 am for most of the foragers re observed outside the nest around this time. Workers were directly frozen in liquid nitrogen. The total of thirty (foragers or nurses) were pooled and were kept at −80°C until further processing. RNA of whole ant bodies was extracted using RNAiso Plus (TaKaRa, Code number 9108). DNase I (RNase-free, TaKaRa) was used to digest genomic DNA from the RNA solution. NanoDrop apparatus (NanoDrop Technologies, Inc.) and Agilent 2100 bioanalyzer (Agilent Technologies) were used to quantify and check for integrity of extracted RNA.

RNA-Seq Library Construction

Obtained mRNA of the two worker samples was purified out of a total of 6 μg of total RNA using oligo (dT) magnetic beads. One cDNA library for foragers and one for nurses were prepared following the manufacturer's instructions of the Illumina kit as in Yang et al. (2014). In brief, first-strand cDNA was synthesized with random hexamer-primers from the purified mRNA; second-strand cDNA was synthesized by using dNTPs, buffer, DNA polymerase I and RNaseH. Obtained ds cDNAs were then purified with QiaQuick PCR extraction kit, fragments washed off with elution buffer and the purified cDNA ligated to sequencing adapters. Only suitable fragments judged by agarose gel electrophoresis were selected for use as templates for PCR amplification. The final cDNA

library was sequenced with Illumina HiSeq 2000 system at Beijing Genomics Institute BGI, Shenzhen, China).

Mapping of RNA-Sequencing Reads

Before mapping reads to the reference genome, ESTs were filtered of adaptor sequences, and by removing low-quality sequences (based on proportion of 'N'); 3′ adapter sequences and empty reads only with 3′ adapter sequences. Processed ESTs were then mapped to scaffold contigs of *S. invicta* reference genome (file *S. invicta* Si_gnF.454scaffolds.fasta.zip) downloaded from NCBI. Mapped reads were assembled as potential genes.

Identification and Functional Analysis of DEGs

DEGs between nurses and foragers were identified using EdgeR package (Robinson et al. 2010), where False Discovery Rate (FDR) is used to determine the threshold of *P* value by multiple tests (Benjamini and Yekutieli 2001). Specifically, gene expression levels were analyzed with a normalization of Reads Per Kb per Million reads (RPKM). The fold change (FC) of each gene was calculated with the following equation:

$$FC = RPKM_{Forager}/RPKM_{Nurse}.$$

Here, we considered genes differentially expressed when FDR values ≤0.01 and Log_2 FC ≥1.

Functional annotation of DEGs was made by BLASTx against the NCBI nr database using an *E*-value <10^{-5}, allied to alignment with Cluster of Orthologous Groups (COG) and Swiss-Prot protein classes in NCBI at *E*-value ≤10^{-5}). BLASTx results were imported to the Blast2Go software to map to Gene Ontology (GO) categories, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Finally, enriched GO terms and KEGG pathways were calculated using Fisher's Exact Test with Multiple Testing Correction (FDR <0.01).

Validation of Differential Expression by qRT-PCR

The expression levels of some selected DEGs were examined. For this test, ten foragers and 10 nurses from the same three polygyne *S. invicta* colonies were collected around 10:00 and frozen in liquid nitrogen, as described above but pooled per colony. Thus, we three samples for foragers and three for nurses were processed for RNA as mentioned above. Candidate genes we selected for qRT-PCR were found as differentially expressed in RNA-seq experiment and are known to play a role in key biological functions of model organisms (followed by published reference): odorant binding protein '*OBP 11*' (olfactory response, Forêt and Maleszka 2006), chemosensory protein '*CSP11*' (expressed in insect sensory appendages and involved in chemical signaling, González et al. 2009), *lysozyme c-1* (antimicrobial response, Kajla et al. 2011), hymenoptaecin '*Hym*' (antimicrobial peptide, Casteels et al. 1993), foraging '*for*' (feeding behavior and polyethism, Ingram et al. 2011), synaptotagmin-1 '*Syt-1*' (neurotransmitter secretion, Brose et al. 1992) and down syndrome cell adhesion molecule-like protein '*Dscam*' (regulating axon configuration in mushroom bodies, Hummel et al. 2003). Specific primers were designed with NCBI primer-blast for these target genes and the reference gene elongation factor 1-beta ('*ef1-beta*', refer to Supp Table 1 [online only]). Three biological replicates and three technical replicates were performed. Reactions were tracked in a qRT-PCR Mx3005P qPCR System (Agilent Technologies Inc., Santa Clara, CA, USA) using SYBR Premix Ex Taq (Takara, Dalian, China). The following PCR settings were used: start 50°C for 2 min, pre-denaturation for 10 min at 95°C; 45 cycles of denaturation for 15 s at 95°C and elongation for 1 min at 56°C. Relative expression

Table 1. Alignment statistics of clean reads of nurses and foragers mapped to the genome of *S.invicta*

Sample	Nurses	Foragers
Total reads	12,617,114 (100.00%)	11,854,765 (100.00%)
Total base pairs	618,238,586 (100.00%)	580,883,485 (100.00%)
Total mapped reads	10,601,620 (84.03%)	9,352,435 (78.89%)
Perfect match	9,256,033 (73.36%)	8,107,742 (68.39%)
≤3 bp mismatch	1,345,587 (10.66%)	1,244,693 (10.50%)
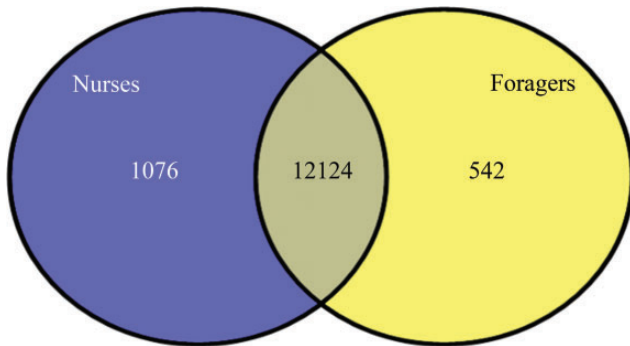
Fig. 1. Numbers of genes expressed in nurses and foragers of *S. invicta*. 1,618 genes were significantly differentially expressed between foragers and nurses, of which 542 genes were more highly expressed in foragers and 1,076 genes were more highly expressed in nurses. 12,124 genes were not significantly differentially expressed between the two age groups.

levels of genes were analyzed using a $2^{-\triangle\triangle CT}$ method. The means and standard error of expression levels in nurses and foragers for each gene were calculated, and *P* values tested with *t*-test by SPSS 13.0 (Schmittgen and Livak 2008).

Comparative Analyses With Microarray Data
The results of microarray (Manfredini et al. 2014) and RNA-seq of gene expression difference between forager and nurse in *S. invicta* were compared. For this purpose, we used our references gene dataset (*S. invicta* SI 2.2.3.cdna.zip) as a database. Blast software (version 2.2.29+) was used to map the microarray sequence to this reference gene database (*E*-value = 1*e*−5). Only sequences with over 80% coverage to query were retained.

Results

Summary of Digital Gene Expression Sequencing
For each sequenced sample 12,617,114 and 11,854,765 clean reads were obtained, nurses, and foragers, respectively. About 84 and 79% of processed reads from nurses and foragers were mapped to the *S. invicta* genome. Among mapped reads, 73 and 68% perfectly mapped onto the genome in nurse and foragers (Table 1).

Based on these mapped reads, we identified 14,055 and 13,643 expressed annotated genes among nurses and foragers, respectively, of which while 13,631 genes were shared, 892 genes were expressed specifically in nurses and 480 genes were exclusive to foragers. Furthermore, 1,618 genes were significantly differentially expressed between nurses and foragers; 542 more highly expressed in foragers, and 1,076 more highly in nurses (Fig. 1).

Sequencing saturation (Suppl Fig. 1 [online only]) is critical for sequencing coverage. From start the number of genes detected
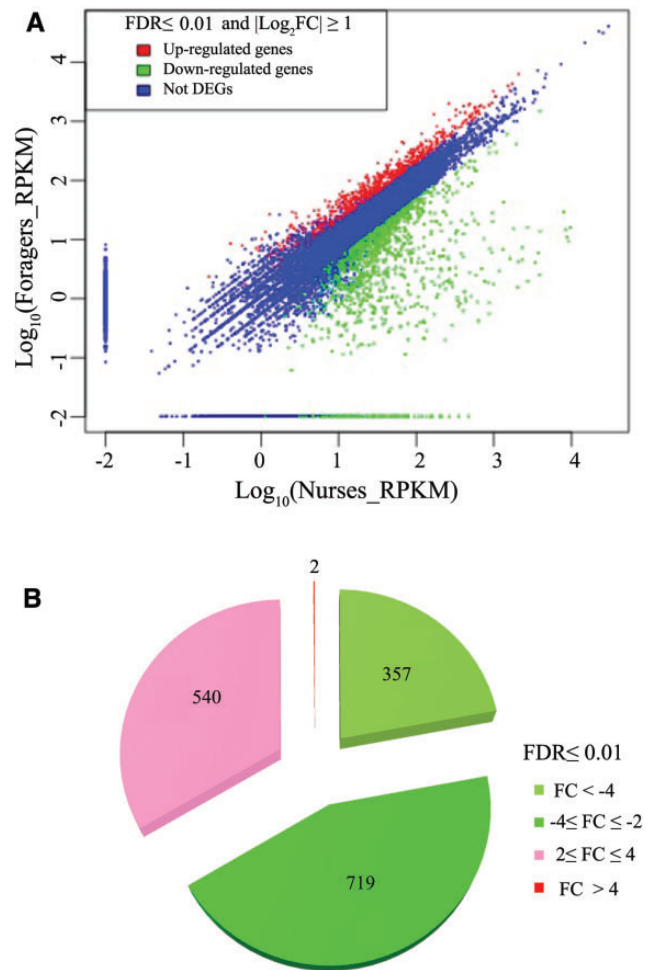
Fig. 2. Genes differentially expressed between *S. invicta* nurses and foragers. (A) represents the genes expressed in nurses and foragers, red points indicate highly expressed in foragers with |log2Ratio| ≥1 and FDR <0.01, in which ratio represent the ratio of foragers to nurses; green points indicate highly expressed genes in nurses with |log2Ratio| ≥1 and FDR <0.01; Blue points indicate not DEGs between foragers and nurses. (B) Represent number of genes differentially expressed with different FCs between nurses and foragers. FC represents the ratio of gene expression level (RPKM) in foragers and nurses. Negative value represents highly expressed genes in nurses. Positive number represents highly expressed genes in foragers.

increased as the amounts of clean reads grew in both nurses and foragers. When reads count reached 2.5 million the number of detected genes tended to saturate culminating in a plateau. This curve suggests sequencing could efficiently detect transcript diversity in *S. invicta*.

Expression levels were tracked by number of reads mapped (Supp Fig. 2 [online only]) to each gene used as reference. In the nurses, the coverage rate of 5,520 (39%) genes ranged within between 90–100%, 2,167 (15%) genes within 70–80%, and 512 (4%) genes were covered at <10%. Similarly, in the foragers, coverage rate of 4754 (35%) genes were between 90 and 100%, 1983 (15%) genes were between 70 and 80%, only 645 (5%) genes' coverage rate were <10%.

DEGs in Nurses and Foragers
A total of 1,618 genes were differentially expressed between nurses and foragers (Supp Table 2 [online only]) (Fig. 2). Among them, 542 were more highly expressed in foragers (FC ≥ 2, |log₂ FC|≥1; FDR

value ≤0.01) and 1,076 significant higher in nurses (FC ≥2, |log2 FC|≥1; FDR value ≤0.01), in which 357 genes had a FC higher than 4 (|$\log_2$ FC|≥2; FDR value ≤0.01). The top 10 DEGs and their FC values are listed on Supp Table 3 [online only].

Annotation of DEGs

The 1,618 DEGs between nurses and foragers were functionally annotated using BLASTX by aligning sequences within diverse protein databases with a cutoff e-value of 10^{-5}, including the non-redundant protein (nr) database, the Swiss-Prot, the KEGG, Cluster of Orthologous Groups of proteins (COG) and Gene Ontology (GO) databases (Table 2). We successfully annotated 1,032 genes (63.78%) in these six databases; 1,003 (61.99%), 597(36.90%), 678 (41.90%), 636 (39.31%), and 299 (18.48%) genes were successfully annotated in nr, Swiss-Prot, KEGG, COG, and GO databases, respectively. Distinct gene sequences were functionally annotated against nr database with a cut-off *E*-value 1×10^{-5}. According to the nr annotation, 63% sequences were matched to *Acromyrmex echinatior*, followed by *Camponotus floridanus* Buckley (18%), *Harpegnathos saltator* Jerdon (11%), and *Megachile rotundata* Fab. (2%) (Supp Fig. 3 [online only]).

Table 2. Functional annotations of DEGs between nurses and foragers in *S.invicta* measured by RNA-seq

| Annotated database	Number of sequences	Percentage
Total genes	1,618	100
nr (*E*-value <10^{-5})	1,003	61.99
Swiss-Prot	597	36.90
KEGG	1,034	63.90
COG	636	39.31
GO	299	18.48

nr: non-redundant protein, KEGG: Kyoto Encyclopedia of Genes and Genomes, COG: Cluster of Orthologous Groups of proteins, GO: Gene Ontology.

Nevertheless, 615 genes (38%) failed to map to nr database, which might have been short sequence reads (Garber et al. 2011).

We successfully annotated 299 DEGs through GO database divided by three categories: biological process, cellular component, and molecular function (Fig. 3 and Supp Table 4 [online only]). Under biological processes category, genes associated with cellular and metabolic processes, single-organism processed, and response to stimulus were highly represented. Within the cellular component category, the subcategories cell, cell part, and organelle were the most highly. The molecular function category was mainly composed of proteins associated with binding, catalytic activity, electron carrier activity, nutrient reservoir activity, structural molecule activity, and transporter activity. These findings provide important information to identify pivotal genes participating in secondary metabolite pathways.

In addition, to further evaluate our annotation for completeness and the effectiveness, DEGs were searched against COG database for functional prediction and classification (Fig. 4 and Supp Table 5 [online only]). COG is a database built on coding sequences from complete genome and system–evolution relationships of bacteria, algae as well as eukaryotes, in which proteins are annotated based on orthologous gene annotations of an ancestral protein. Therefore, genes aligned to COG database can be classified according to their possible functions. In total, 636 genes (39.3% of all DEGs; Table 2) were successfully assigned to COG database and divided into 25 functional groups (Fig. 5). Specifically, the functions 'carbohydrate transport and metabolism' accounted for 24.06%, 'secondary metabolites biosynthesis, transport and catabolism' accounted for 18.71%, 'transcription' accounted for 18.71%, 'cell cycle control, cell division, chromosome partitioning' accounted for 16.35%, and 'cell wall/membrane/envelope biogenesis' accounted for 16.35%. In these groups, 'secondary metabolites biosynthesis, transport and catabolism' can be hypothesized as most relevant for foragers' resistance against insecticides and environmental stress.

The KEGG annotation database was used to obtain metabolic pathway information (Fig. 5 and Supp Table 6 [online only]). Based on sequence homologies, 1,034 genes out of 1,618 DEGs (63.9%)
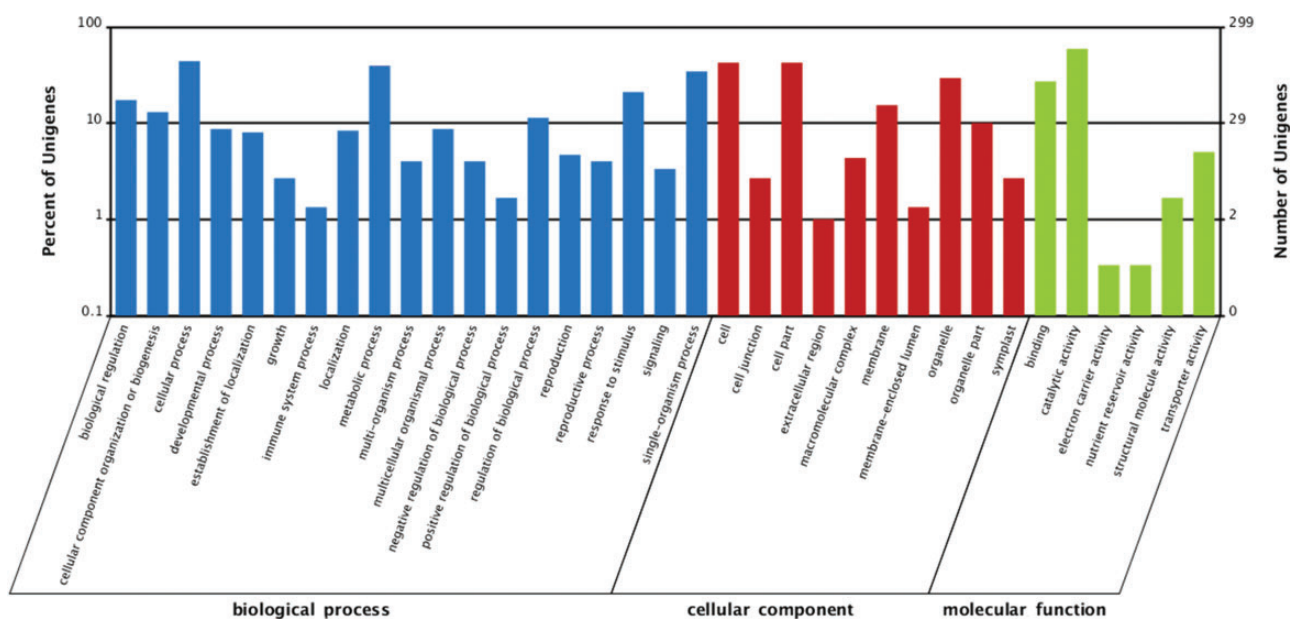
Fig. 3. GO classifications of DEGs between nurses and foragers in *S. invicta*. A total of 1,034 genes (63.90% of the DEGs) were categorized into three Gene Ontology classes, biological processes, cellular components, and molecular functions, respectively.
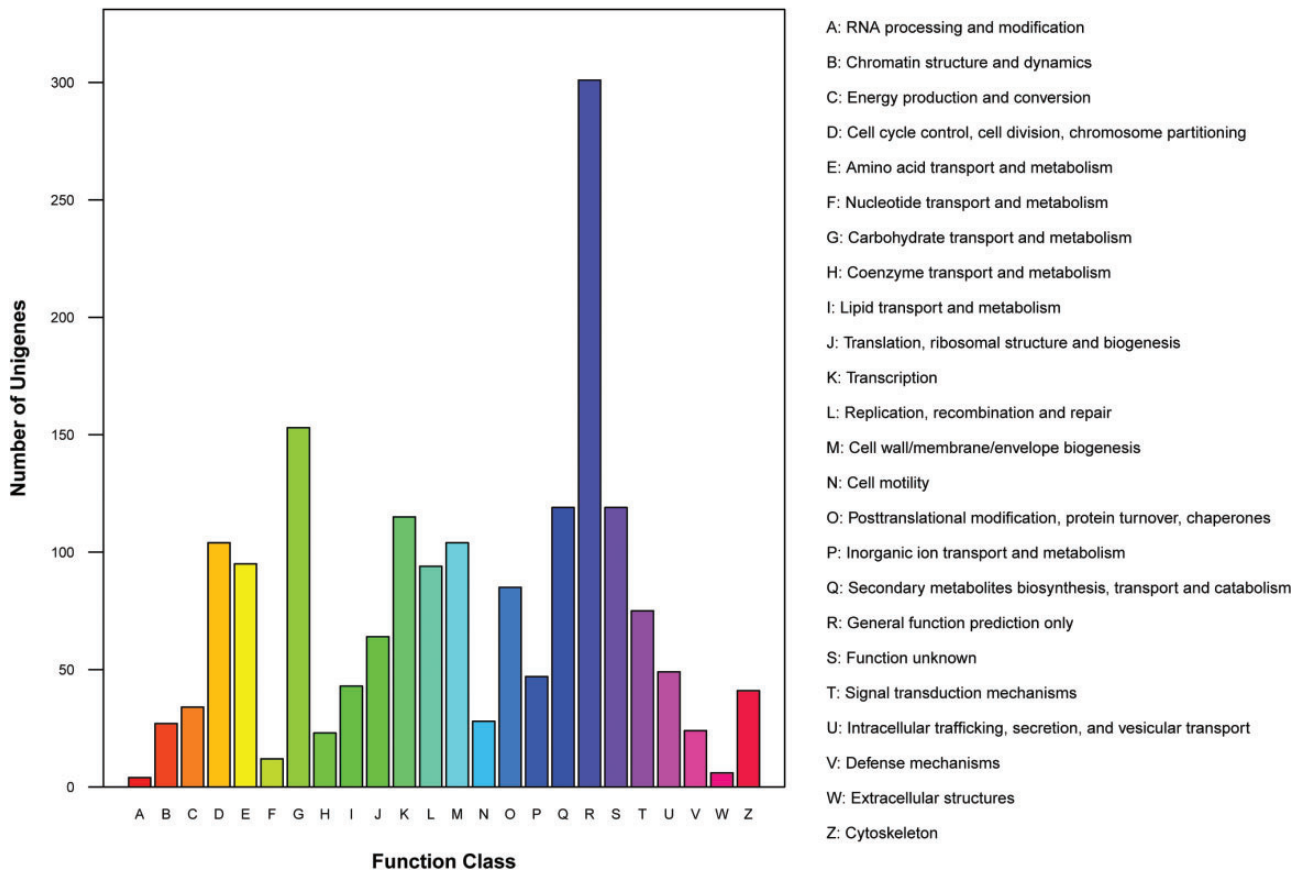
Fig. 4. Cluster of orthologous groups (COG) classification of DEGs between nurses and foragers in *S. invicta*. In total, 636 of the 1,618 DEGs sequences with COG database hits were grouped into 25 COG classifications.

were annotated with KEGG and mapped to 225 pathways, among which 41 pathways were significantly enriched. Pathway enrichment analysis shows that the top four most enriched pathways in DGEs were Insulin signaling pathway, Fatty acid biosynthesis, Pancreatic secretion, and Protein digestion and absorption (Fig. 5). We found that DEGs were prevalent in lipid metabolism, drug metabolism-cytochrome P450, insect hormone biosynthesis, and some secretion pathways. These annotations provide valuable resources to investigate specific processes, functions, and pathways regarding social insect research.

Validation of Differential Expression of Select Genes With qRT-PCR

We carried out qRT-PCR analysis to validate some of the DEGs identified by RNA-seq. All selected genes, excepting *Dscam*, also showed significantly higher expression by qRT-PCR (ANOVA, $P < 0.05$, Table 3). All such genes showed similar expression patterns as indicated by the RNA-seq analysis, suggesting that RNA-seq method could be a powerful and robust tool to discover candidate genes involved in biological functions.

Comparative Analyses With Microarray Data

We carried out comparative analyses between the microarray of Manfredini et al. (2014) and our RNA-seq datasets. Results showed 116 significantly DEGs overlapped between the microarray dataset and the present RNA-seq results, of which 92 genes showed similar trend in expression (Supp Table 7 [online only]). The GO term

myofibril assembly (GO: 0030239) was particularly enriched by both microarray and RNA-seq.

Discussion

Age-related division of labor in social insect colonies involves the performance of various tasks by different worker castes. Nurses feed larvae and care for the queen inside the nest and then shift to forage for food as they age. The molecular mechanism of division of labor in *S. invicta* workers was investigated and our findings confirm division of labor in workers is correlated with transcriptome profile changes. We detected a total of 1,618 DEGs between nurses and foragers, in which part of the genes are likely to be causal effect of gene regulation on the worker division of labor, but some genes may have been altered as a result of different environmental exposure of the workers during their tasks.

We found several pathways involved in nutrition were significantly enriched, including the insulin signaling pathway, fatty acid biosynthesis, amino sugar and nucleotide sugar metabolism, protein digestion and absorption, and lipid metabolism pathways. Previous studies have proposed that worker division of labor was regulated by nutrition and associated with lipid reserves involved in worker feeding behavior (Ament et al. 2010, Silberman et al. 2016). The insulin signaling pathway was also correlated with social regulation of worker–worker as well as worker–queen division of labor in *Apis mellifera* Linnaeus (Wheeler et al. 2006, Ament et al. 2008). Lipid reserves were negatively correlated with the propensity of workers to forage: nurses usually have greater lipid reserves than foragers
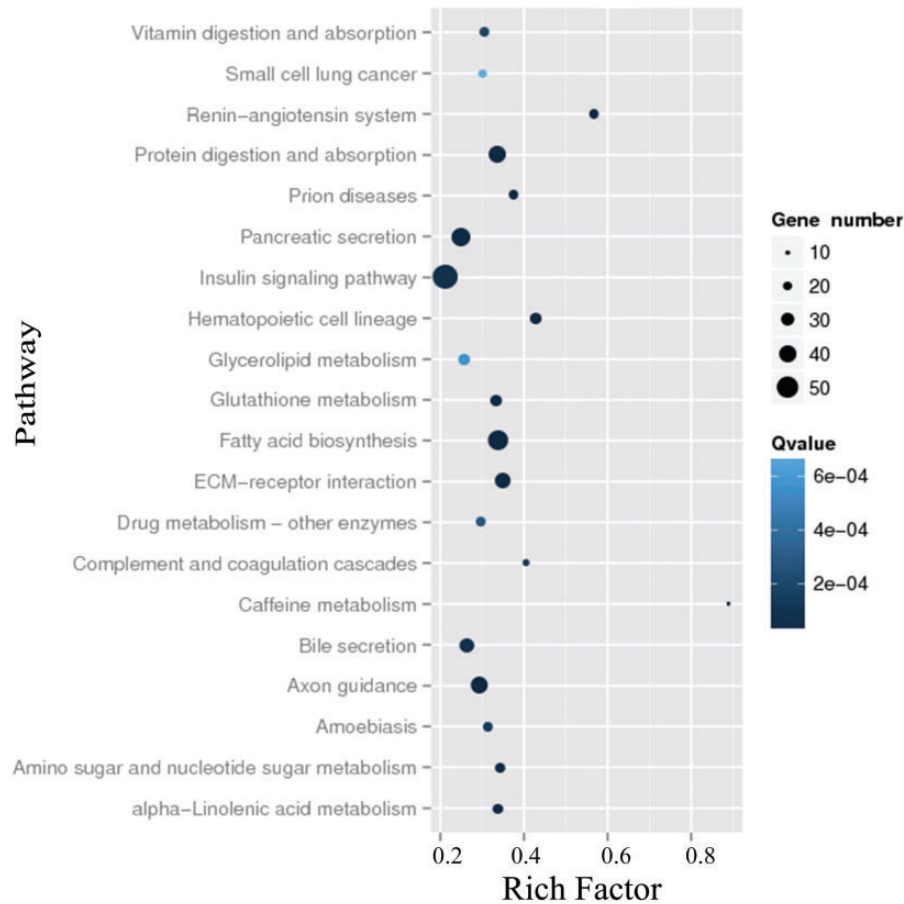
Fig. 5. Scatter plot of top 20 significantly enriched KEGG pathways among DEGs. *Q*-value is calculated based on Hypergeometric Test. Pathways with *Q* value ≤0.05 are significantly enriched in DEGs.

Table 3. Selected DEGs by RNA-Seq re-evaluated by qRT-PCR expression analysis

Gene ID	Description	FC		*P* value
		DGE	qRT-PCR	
SI2.2.0_06972	Ordorant binding protein, *OBP 11*	3.83	4.64	0.012
SI2.2.0_01617	Lysozyme c-1	2.13	2.94	0.001
SI2.2.0_10737	Uncharacterized LOC 105202151	2.10	2.85	0.002
SI2.2.0_16397	Synaptotagmin-1	2.09	2.93	0.001
SI2.2.0_06399	Hymenoptaecin	2.80	3.08	0.042
SI2.2.0_04861	Chemosensory protein, *CSP11*	−5.64	−7.65	0.012
SI2.2.0_08925	Foraging, *for*	−1.09	−1.21	0.337
SI2.2.0_05898	Down syndrom cell adhesion molecule-like protein, Dscam	2.22	1.99	0.263

Positive and negative values represent higher and lower expression in foragers relative to nurses of *S. invicta* workers, respectively. *P* values were calculated with *t*-test in SPSS 13.0.

(Blanchard et al. 2000, Toth and Robinson 2005, Ament et al. 2011, Silberman et al. 2016). Similarly, greater reserves of proteins are found in nurses and treatments reducing nurse protein reserves will induce a shift to foraging (Nelson et al. 2007). Interestingly, in the evolutionary convergent termites *Reticulitermes flavipes* Kollar, nutrition-related genes, such related to fatty acid biosynthesis, have been associated with division of labor (Scharf et al. 2005). Genes enriched in such pathways might be related to structural differentiation in the brain structure and chemistry, playing key roles during the maturation of nurses to foragers (Seid et al. 2005).

DNA methylation is considered a crucial factor in caste determination in ants (Bonasio et al. 2012). By this biochemical process a methyl group is added to a cytosine or adenine DNA nucleotide, with important consequences in epigenetic gene regulation in development and disease (Jaenisch and Bird 2003, Galbraith et al. 2015). Previous studies on mammals suggest that environmental stimuli could alter the epigenetic condition of a genome modulating gene expression via DNA methylation (Jaenisch and Bird 2003). In social insects, caste determination is also determined by epigenetic regulation in the ants *C. floridanus* and *H. saltator*, as well as in the

honeybee *A. mellifera*, but not in clonal ants such as *Cerapachys biroi* Forel (Elango et al. 2009, Bonasio et al. 2010, Libbrecht et al. 2016). Our results shows that the genes for methyl-CpG-binding domain protein 3 and arginine *N*-methyltransferase 10 were significantly downregulated in foragers of *S. invicta*. A previous study also showed that DNA methylation would influence worker development: larvae are likely to develop to queens when DNA methyl transferase 3 is knocked down with RNAi (Kucharski et al. 2008). Thus, it seems likely that these two DEGs related to DNA methylation play key roles in regulating downstream gene expression involved in worker temporal subcaste development.

The genes *foraging* (*for*) and *malvolio* are reported to play key roles in behavioral maturation and division of labor (Ben-Shahar et al. 2004, 2005). We found that the expression level of *for* is not significantly different between nurses and foragers. This contrasts with previous studies with *A. mellifera* and the harvester ant (Ben-Shahar 2005; Ingram et al. 2011, 2016) in which foragers had significantly higher *for* expression. Thus although *for* may be involved in age-dependent transition from nurses to foragers it might not be a main gene involved. Similar result was also observed with *malvolio* gene. *Malvolio* plays an important role in division of labor in *A. mellifera*, but as far as we know, there are no reports of any function in ant division of labor. There are three independent explanations for the detected low fold differences of *for* and *Malvolio* between nurses and foragers. A first one is that these two genes might modify at the protein translational level but not the RNA transcriptional level (Patalano et al. 2012, Yan et al. 2015). A second explanation is that whole-body analysis might obscure tissue-specific profiles for these genes (Whitfield et al. 2003). A last explanation is that gene expression may vary over the day depending on tasks affecting amounts of mRNA level of genes (Ingram et al. 2011). Future studies focusing on the epigenetic processes and specific-tissue expression as taken at multiple time scales between different castes would be more fruitful.

Previous studies showed that the expression of transcription factors are involved in the regulation of development in *Drosophila* and also related to worker division of labor in the honey bee. But the difference in promoter regions between behaviorally related bee genes and their *Drosophila* orthologs indicates different gene regulation mechanisms for transcription factors (Sinha et al. 2006, Whitfield et al. 2006). Transcription factors—also known as sequence-specific DNA-binding factors—control the transcription rate through binding to specific DNA template sequences (Schrader et al. 2015). Transcription factors can promote or block the recruitment of RNA polymerase to specific genes thus acting as activators or repressors (Latchman 1997). In our results, several transcription factor genes were significantly differentially expressed between nurses and foragers, in which the transcription factors 'MafA' and 'kayak' were upregulated in foragers, but 'glial cells missing', 'Sp9-like', AP-2-epsilon-like, and SOX-6 were upregulated in nurses.

There were also some DEGs that are likely to be effects of the different environments the workers of different tasks inhabit. For example, workers foraging outside colony are exposed to different types of pathogens. Thus, the immune strength might be different across worker developmental stages. Indeed, in *A. mellifera*, the immune strength of older foragers is higher than among nurse bees (Wilson-Rich et al. 2008). In our results, AMPs such as defensin-2 and hymenoptaecin were significantly more highly expressed in foragers than nurses, which is consistent with the observation that foragers suffer from contacting more microbes when they are foraging outside the nest. However, lipopolysaccharide binding proteins (LBPs) were upregulated in nurses. LBPs are conserved proteins that bind to bacterial lipopolysaccharides (LPS) in eliciting immune responses (Kang et al. 1998). We believe that this gene might play a role in the innate immunity of nurses. Similarly, enzyme superfamily glutathione s-transferases (GST) exhibit various catalytic functions across a range of organisms. There are many studies on GSTs given a critical role of GST in insecticide resistance in insects. Our finds indicate GST1-1 was more highly expressed in foragers, suggesting foragers might be more resistant to insecticides and other toxic chemicals than nurses.

Differential response thresholds of individuals in a colony might result from the different odor the workers engaged in different tasks may encounter. For example, the response threshold of nurses to brood and queens might be low, and nurses would be more responsive to brood and queens, and foragers might be more responsive to food and environments (Detrain and Pasteels 1991). Accordingly, some genes involved in olfaction were upregulated in foragers as compared with nurses, such as the odorant-binding proteins OBP10, OBP11, and the Odorant receptors OR2, Gustatory and odorant receptor 7, which might play essential roles in the response of foragers to environmental stimuli, such as encounters with natural enemies and recruitment in collecting food. In the same direction several 'chemosensory proteins' (CSP) were upregulated in foragers such as CSP8 and CSP9, which might play key roles in the ability of foragers being able to detect the taste of food. Controversially, 'Down syndrome cell adhesion molecule' gene (*Dscam*) which is reported to play a role in olfaction was downregulated in foragers (Hummel et al. 2003).

Through comparison with microarray data on the same topic, we found DEGs overlapping between the two methods. The GO term 'myofibril assembly' was significantly enriched in both experiments, indicating an importance of the composition of muscles in the differential labors of foragers and nurses. During muscle development, troponin C (TpnC), the calcium-binding subunit of the troponin regulatory complex in the muscle thin filament plays an important function (Agianian et al. 2004). This protein is encoded by multiple genes among insects (Agianian et al. 2004). Here, we found two differentially expressed TpnC genes in nurses and foragers. These two TpnC genes were significantly more expressed in foragers, suggesting a potential function in the formation of muscles involved in foraging activity (Supp Table 2 [online only]).

With the advanced technology of RNA-seq, many DEGs and significantly enriched GO terms and pathways between nurses and foragers were detected which might play a role in the task switch from nurses to foragers in *S. invicta*. These critical genes, their GO terms and KEGG pathways, represent a potential genetic 'tool kit' for future studies to unveil the molecular mechanisms underpinning the relationship between genome and division of labor in ants, especially regarding overlapping genes and annotation terms as detected by microarray and the present RNA-seq experiment.

Supplementary Data

Acknowledgments

References Cited

Agianian, B., U. Kržič, F. Qiu, W. A. Linke, K. Leonard, and B. Bullard. 2004. A troponin switch that regulates muscle contraction by stretch instead of calcium. EMBO J. 23: 772–779.

Ament, S. A., M. Corona, H. S. Pollock, and G. E. Robinson. 2008. Insulin signaling is involved in the regulation of worker division of labor in honey bee colonies. PNAS. 105: 4226–4231.

Ament, S. A., Y. Wang, and G. E. Robinson. 2010. Nutritional regulation of division of labor in honey bees: toward a systems biology perspective. Wiley Interdisciplinary Reviews: Syst. Biol. Med. 2: 566–576.

Ament, S. A., Q. W. Chan, M. M. Wheeler, S. E. Nixon, S. P. Johnson, S. L. Rodriguez-Zas, L. J. Foster, and G. E. Robinson. 2011. Mechanisms of stable lipid loss in a social insect. J. Exp. Biol. 214: 3808–3821.

Ben-Shahar, Y. 2005. The foraging gene, behavioral plasticity, and honeybee division of labor. J. Comp. Physiol. A. 191: 987–994.

Benjamini, Y., and D. Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 34: 1165–1188.

Ben-Shahar, Y., N. L. Dudek, and G. E. Robinson. 2004. Phenotypic deconstruction reveals involvement of manganese transporter malvolio in honey bee division of labor. J. Exp. Biol. 207: 3281–3288.

Berens, A. J., J. H. Hunt, and A. L. Toth. 2015. Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. Mol. Biol. E. 32: 690–703.

Blanchard, G., G. Orledge, S. Reynolds, and N. Franks. 2000. Division of labour and seasonality in the ant *Leptothorax albipennis*: worker corpulence and its influence on behaviour. Anim. Behav. 59: 723–738.

Bonasio, R., G. Zhang, C. Ye, N. S. Mutti, X. Fang, N. Qin, G. Donahue, P. Yang, Q. Li, C. Li, P. Zhang, Z. Huang, S. L. Berger, D. Reinberg, J. Wang, and J. Liebig. 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science. 329: 1068–1071.

Bonasio, R., Q. Li, J. Lian, N. S. Mutti, L. Jin, H. Zhao, P. Zhang, P. Wen, H. Xiang, and Y. Ding. 2012. Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Curr. Biol. 22: 1755–1764.

Brose, N., A. G. Petrenko, T. C. Südhof, and R. Jahn. 1992. Synaptotagmin: a calcium sensor on the synaptic vesicle surface. Science. 256: 1021–1025.

Casteels, P., C. Ampe, F. Jacobs, and P. Tempst. 1993. Functional and chemical characterization of hymenoptaecin, an antibacterial polypeptide that is infection-inducible in the honeybee (*Apis mellifera*). J. Biol. Chem. 268: 7044–7054.

Detrain, C., and J. Pasteels. 1991. Caste differences in behavioral thresholds as a basis for polyethism during food recruitment in the ant, *Pheidole pallidula* (Nyl.) (Hymenoptera: Myrmicinae). J. Insect Behav. 4: 157–176.

Duarte, A., F. J. Weissing, I. Pen, and L. Keller. 2011. An evolutionary perspective on self-organized division of labor in social insects. Annu. Rev. Ecol. Evol. Syst. 42: 91–110.

Elango, N., B. G. Hunt, M. A. Goodisman, and S. V. Yi. 2009. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. PNAS. 106: 11206–11211.

Forêt, S., and R. Maleszka. 2006. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). Genome Res. 16: 1404–1413.

Feldmeyer, B., D. Elsner, and S. Foitzik. 2014. Gene expression patterns associated with caste and reproductive status in ants: worker-specific genes are more derived than queen-specific ones. Mol. Ecol. 23: 151–161.

Galbraith, D. A., X. Yang, E. L. Niño, S. Yi, and C. Grozinger. 2015. Parallel epigenomic and transcriptomic responses to viral infection in honey bees (*Apis mellifera*). PLoS Path. 11: e1004713–e1004713.

Garber, M., M. G. Grabherr, M. Guttman, and C. Trapnell. 2011. Computational methods for transcriptome annotation and quantification using RNA-seq. Nat. Methods. 8: 469–477.

González, D., Q. Zhao, C. McMahan, D. Velasquez, W. E. Haskins, V. Sponsel, A. Cassill, and R. Renthal. 2009. The major antennal chemosensory protein of red imported fire ant workers. Insect Mol. Biol. 18: 395–404.

Harrison, M. C., R. L. Hammond, and E. B. Mallon. 2015. Reproductive workers show queenlike gene expression in an intermediately eusocial insect, the buff-tailed bumble bee *Bombus terrestris*. Mol. Ecol. 24: 3043–3063.

Hölldobler, B.1990. The ants. Harvard University Press, Cambridge.

Hummel, T., M. L. Vasconcelos, J. C. Clemens, Y. Fishilevich, L. B. Vosshall, and S. L. Zipursky. 2003. Axonal targeting of olfactory receptor neurons in *Drosophila* is controlled by Dscam. Neuron. 37: 221–231.

Ingram, K. K., L. Kleeman, and S. Peteru. 2011. Differential regulation of the foraging gene associated with task behaviors in harvester ants. BMC Ecol. 11: 19.

Ingram, K. K., D. M. Gordon, D. A. Friedman, M. Greene, J. Kahler, and S. Peteru. 2016. Context-dependent expression of the foraging gene in field colonies of ants: the interacting roles of age, environment and task. Proc. R. Soc. B. 283: 20160841.

Jaenisch, R., and A. Bird. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 33: 245–254.

Kajla, M. K., L. Shi, B. Li, S. Luckhart, J. Li, and S. M. Paskewitz. 2011. A new role for an old antimicrobial: lysozyme c-1 can function to protect malaria parasites in *anopheles mosquitoes*. PLoS One. 6: e19649.

Kang, D., G. Liu, A. Lundström, E. Gelius, and H. Steiner. 1998. A peptidoglycan recognition protein in innate immunity conserved from insects to humans. PNAS. 95: 10078–10082.

Kucharski, R., and R. Maleszka. 2002. Evaluation of differential gene expression during behavioral development in the honeybee using microarrays and northern blots. Genome Biol. 3: 1-0007.

Kucharski, R., J. Maleszka, S. Foret, and R. Maleszka. 2008. Nutritional control of reproductive status in honeybees via DNA methylation. Science. 319: 1827–1830.

Lach, L., C. L. Parr, and K. L. Abott. 2010. Ant ecology. Oxford University Press.

Latchman, D. S. 1997. Transcription factors: an overview. Intern. J. Biochem. Cell Biol. 29: 1305–1312.

Lattorff, H. M. G., and R. F. Moritz. 2013. Genetic underpinnings of division of labor in the honeybee (*Apis mellifera*). Trends Genet. 29: 641–648.

Libbrecht, R., P. R. Oxley, D. Kronauer, and L. Keller. 2013. Ant genomics sheds light on the molecular regulation of social organization. Genome Biol. 14: 212.

Libbrecht, R., P. R. Oxley, L. Keller, and D. Kronauer. 2016. Robust DNA methylation in the clonal raider ant brain. Curr. Biol. 26: 391–395.

Manfredini, F., C. Lucas, M. Nicolas, L. Keller, D. Shoemaker, and C. M. Grozinger. 2014. Molecular and social regulation of worker division of labour in fire ants. Mol. Ecol. 23: 660–672.

Markiewicz, D. A., and S. O'Donnell. 2001. Social dominance, task performance and nutrition: implications for reproduction in eusocial wasps. J. Compar. Physiol. A. 187: 327–333.

Mirenda, J. T., and S. B. Vinson. 1981. Division of labour and specification of castes in the red imported fire ant *Solenopsis invicta* Buren. Anim. Behav. 29: 410–420.

Morandin, C., K. Dhaygude, J. K. Paviala, C. Trontti, H. Wheat, and H. Helantera. 2015. Caste biases in gene expression are specific to developmental stage in the ant *Formica exsecta*. J. Evol. Biol. 28: 1705–1718.

Morozova, O., and M. A. Marra. 2008. Applications of next-generation sequencing technologies in functional genomics. Genomics. 92: 255–264.

Nelson, C. M., K. E. Ihle, M. K. Fondrk, R. E. Page, Jr, and G. V. Amdam. 2007. The gene vitellogenin has multiple coordinating effects on social organization. PLoS Biol. 5: e62.

Patalano, S., T. A. Hore, W. Reik, and S. Sumner. 2012. Shifting behaviour: epigenetic reprogramming in eusocial insects. Curr. Opin. Cell Biol. 24: 367–373.

Qiu, H. L., L. H. Lu, Q. X. Shi, and Y. R. He. 2014. Fungus exposed *Solenopsis invicta* ants benefit from grooming. J. Insect Behav. 27: 678–691.

Qiu, H. L., L. H. Lu, Q. X. Shi, C. C. Tu, T. Lin, and Y. R. He. 2015. Differential necrophoric behaviour of the ant *Solenopsis invicta* towards fungal-infected corpses of workers and pupae. Bull. Entomol. Res. 105: 607–614.

Robinson, M. D., D. J. Mccarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26: 139–140.

Scharf, M., D. Wu-Scharf, X. Zhou, B. Pittendrigh, and G. Bennett. 2005. Gene expression profiles among immature and adult reproductive castes of the termite *Reticulitermes flavipes*. Insect Mol. Biol. 14: 31–44.

Schmittgen, T. D., and K. J. Livak. 2008. Analyzing real-time PCR data by the comparative CT method. Nat. Protoc. 3: 1101–1108.

Schrader, L., D. F. Simola, J. Heinze, and J. Oettler. 2015. Sphingolipids, transcription factors, and conserved toolkit genes: developmental plasticity in the ant *Cardiocondyla obscurior*. Mol. Biol. Evol. 32: 1474–1486.

Seid, M. A., K. M. Harris, and J. F. Traniello. 2005. Age-related changes in the number and structure of synapses in the lip region of the mushroom bodies in the ant *Pheidole dentata*. J. Comp. Neurol. 488: 269–277.

Silberman, R. E., D. Gordon, and K. K. Ingram. 2016. Nutrient stores predict task behaviors in diverse ant species. Insectes Soc. 63: 299–307.

Sinha, S., X. Ling, C. W. Whitfield, C. Zhai, and G. E. Robinson. 2006. Genome scan for cis-regulatory DNA motifs associated with social behavior in honey bees. PNAS. 103: 16352–16357.

Smith, C. R., A. L. Toth, A. V. Suarez, and G. E. Robinson. 2008. Genetic and genomic analyses of the division of labour in insect societies. Nat. Rev. Genet. 9: 735–748.

Toth, A. L., and G. E. Robinson. 2005. Worker nutrition and division of labour in honeybees. Anim. Behav. 69: 427–435.

Tschinkel, W. R. 2006. The fire ants. Harvard University Press, Cambridge.

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10: 57–63.

Wheeler, D., N. Buck, and J. D. Evans. 2006. Expression of insulin pathway genes during the period of caste determination in the honey bee, *Apis mellifera*. Insect Mol. Biol. 15: 597–602.

Whitfield, C. W., Y. Ben-Shahar, C. Brillet, I. Leoncini, D. Crauser, Y. LeConte, S. Rodriguez-Zas, and G. E. Robinson. 2006. Genomic dissection of behavioral maturation in the honey bee. PNAS. 103: 16068–16075.

Whitfield, C. W., A. M. Cziko, and G. E. Robinson. 2003. Gene expression profiles in the brain predict behavior in individual honey bees. Science. 302: 296–299.

Wilson, E. O. 1971. The insect societies. Harvard University Press, Cambridge.

Wilson-Rich, N., S. T. Dres, and P. T. Starks. 2008. The ontogeny of immunity: development of innate immune strength in the honey bee (*Apis mellifera*). J. Insect Physiol. 54: 1392–1399.

Yan, H., R. Bonasio, D. F. Simola, J. Liebig, S. L. Berger, and D. Reinberg. 2015. DNA methylation in social insects: how epigenetics can control behavior and longevity. Annu. Rev. Entomol. 60: 435–452.

Yang, X., X. Liu, X. Xu, Z. Li, Y. Li, D. Song, T. Yu, F. Zhu, Q. Zhang, and X. Zhou. 2014. Gene expression profiling in winged and wingless cotton aphids, *Aphis gossypii* (Hemiptera: Aphididae). Intern. J. Biol. Sci. 10: 257.