Published as: *Science*. 2010 August 27; 329(5995): 1068–1071.

Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*

Roberto Bonasio[1,*], **Guojie Zhang**[2,3,*], **Chaoyang Ye**[4,*], **Navdeep S. Mutti**[5,*], **Xiaodong Fang**[3,*], **Nan Qin**[3,*], **Greg Donahue**[4], **Pengcheng Yang**[3], **Qiye Li**[3], **Cai Li**[3], **Pei Zhang**[3], **Zhiyong Huang**[3], **Shelley L. Berger**[4,†], **Danny Reinberg**[1,6,†], **Jun Wang**[3,7,†], and **Jürgen Liebig**[5,†]

[1]Department of Biochemistry, New York University School of Medicine, 522 First Avenue, New York, NY 10016, USA

[2]CAS-Max Planck Junior Research Group, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

[3]BGI-Shenzhen, Shenzhen 518083, China

[4]Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

[5]School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

[6]Howard Hughes Medical Institute, New York University Medical School, New York, NY 10016, USA

[7]Department of Biology, University of Copenhagen, Copenhagen DK-2200, Denmark

Abstract

The organized societies of ants include short-lived worker castes displaying specialized behavior and morphology, and long-lived queens dedicated to reproduction. We sequenced and compared the genomes of two socially divergent ant species: *Camponotus floridanus* and *Harpegnathos saltator*. Both genomes contained high amounts of CpG, despite the presence of DNA methylation, which in non-Hymenoptera correlates with CpG depletion. Comparison of gene expression in different castes identified upregulation of telomerase and sirtuin deacetylases in longer-lived *H. saltator* reproductives, caste-specific expression of miRNAs and SMYD histone methyl-transferases, and differential regulation of genes implicated in neuronal function and chemical communication. Our findings provide clues on the molecular differences between castes in these two ants, and establish a new experimental model to study epigenetics in aging and behavior.

As eusocial insects, ants live in populous colonies, in which up to millions of individuals delegate the reproductive role to one or few queens, while non-reproductive workers carry out all tasks required for colony maintenance (1). These mutually exclusive morphologies and behaviors arise from a single genome, and are typically not dictated by genetic differences, but by environmental factors (2). The first fertilized (diploid) eggs laid by a founder queen develop into workers, but as the colony enlarges, some diploid embryos take a different developmental path to become virgin queens, leave the nest, mate, and establish new colonies. As colonies mature, queens transition from a broad behavioral repertoire that

†To whom correspondence should be addressed: bergers@mail.med.upenn.edu (S.L.B.), Danny.Reinberg@nyumc.org (D.R.), wangj@genomics.cn (J.W.), Juergen.Liebig@asu.edu (J.L.).
*These authors contributed equally

allows them to forage, excavate nests and rear offspring, to one restricted to egg-laying and total dependence on workers. They also live on average 10 times longer than workers offspring and up to 500 times longer than males (3).

We compared the genomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*, because of contrasts in their behavioral flexibility, caste specialization, and social organization. *C. floridanus* lives in large organized colonies, in which only the queen lays fertilized eggs; when the queen dies, so does the colony (1). Non-reproductive individuals belong to two separate castes, major and minor workers, which exhibit differences in morphology and behavior established during development purely on environmental grounds. In contrast, the *H. saltator* social system and division of labor are more basal: dimorphism between queens and workers is limited, and when the queen dies she is replaced by workers that become functional queens, called gamergates (4).

These two ant species differ in other respects as well. *C. floridanus* are scavengers, forage diurnally and nocturnally, and lay pheromone trails that mark paths to food sources. *H. saltator* workers prey on small arthropods in a solitary and diurnal fashion. *C. floridanus* exhibits high territoriality, strong nestmate recognition and elaborate task specialization. In contrast, *H. saltator* displays low territoriality, loses nestmate recognition in the laboratory, and has only basic task specialization.

The Illumina Genome Analyzer platform was used to sequence genomic libraries for *C. floridanus* and *H. saltator*, obtaining more than 100-fold coverage. Draft genomic assemblies reached scaffold N50 size of ~600 kb (Table S1), although for *C. floridanus* most genome-wide analyses reported here were conducted on an earlier version (v3), with scaffold N50 size of 444 kb (Table S1). Assembly resulted in only small gaps and large N50 size, which assured us that most genomic features, particularly gene models, were predicted with reasonable accuracy. We verified the assemblies by sequencing 9 (*C. floridanus*) and 10 (*H. saltator*) randomly selected fosmid inserts (average size 37 kb) (Table S2). Additionally, we sequenced ~5,000 ESTs from each ant and mapped them to the assembled scaffolds; more than 95% matched the assemblies (Table S3).

The *C. floridanus* and *H. saltator* assemblies cover more than 90% of the genomes, which we estimate at 240 and 330 Mb in size, respectively (Fig. S1, Table S4). The *H. saltator* assembly contains 45% G+C, similar to *D. melanogaster* (42%), and *N. vitripennis* (42%), whereas the *C. floridanus* genome is A+T rich, with a G+C frequency of 34%, similar to *A. mellifera* (33%) (5) (Table S4). Organisms that use DNA methylation for gene regulation typically display a depletion of CpG dinucleotides in their genome (Table S5); however, CpG dinucleotides are over-represented in both ant genomes (Table S4, Fig. S2), despite the presence of cytosine methylation. CpG dinucleotides are also over-represented in *A. mellifera* (5) and in *N. vitripennis*; thus, this genomic feature may be specific to the Hymenoptera.

Repetitive elements comprise 15% of the assemblies in *C. floridanus* and 26% in *H. saltator* (Table S4, Fig. S3), which is probably an underestimate, because genome regions that cannot be assembled are enriched in repeats. Both genomes contain copies of the piggyBac transposon (234 total and 6 with intact ORFs in *C. floridanus*, 976 total and 23 with intact ORFs in *H. saltator*), which has been used for insect transgenesis (6, 7), and may also prove useful in ants.

Segmental duplications (SDs) account for 9.6% and 14.8% of the *C. floridanus* and *H. saltator* genomes, respectively (Table S6). The 1,810 (*C. floridanus*) and 2,858 (*H. saltator*) gene models found in SDs are enriched in Gene Ontology (GO) terms associated with chemo-detection, olfactory function, and protease/peptidase activity (Table S7). *C.*

floridanus SDs are also enriched for genes belonging to the cytochrome P450 family — perhaps due to detoxifying needs related to their generalist feeder life style.

The predicted proteomes of *C. floridanus* and *H. saltator* share most protein families with *A. mellifera* and *N. vitripennis*, but also contain a significant number of ant-specific (690) and speciesspecific families (3,230 in *C. floridanus* and 2,617 in *H. saltator*) without homologues in the other two Hymenoptera (Fig. S4). When more insect genomes are included in the comparison, one third of the ant protein coding genes are conserved with vertebrates and two thirds are conserved with insects, leaving 506 ant-specific protein families and 2–3,000 species-specific families (Fig. 1). Ant-specific genes were enriched in GO terms “olfactory receptor activity”, “sensory perception of smell”, “G protein-coupled receptor activity”, and “odorant binding”, but also terms related with detoxification, including “monooxygenase activity” and “heme binding” (Table S8).

We annotated 96 miRNA genes in *C. floridanus* and 159 in *H. saltator*. These accounted for the majority of small RNA reads in adult specimens, with the exception of *H. saltator* gamergates, which showed a larger proportion of small RNAs mapping to unannotated regions of the genome (Fig. 2A, Table S9). Gamergates also expressed the most diverse miRNA repertoire (Table S10). The two *C. floridanus* worker castes displayed differential expression of miRNAs, with cflo-mir-64 upregulated in minor workers (Fig. 2B) and cflo-mir-7 in major workers (Fig. S5). This suggests that, in addition to being developmentally regulated (Fig. 2C–D), some miRNAs might contribute to the differences among ant castes.

We examined the ant genomes and transcriptomes for signatures related to aging. Telomere shortening is a hallmark of cellular senescence in multi-cellular eukaryotes, and the enzyme telomerase (TERT), which counteracts telomere shortening, prolongs lifespan upon overexpression (8). *TERT* RNA levels were highest in eggs and lower in adults in both *C. floridanus* and *H. saltator*, but, interestingly, they were upregulated in *H. saltator* gamergates (Fig. 3A). This may be explained by the fact that gamergates acquire many physiological characteristics of queens, including longer lifespan (9). Aging has also been linked to the sirtuin lysine deacetylases enzymes SIRT1 and SIRT6, homologous to the *S. cerevisiae* Sir2p implicated in replicative senescence (10). In *H. saltator* gamergates both of these genes are expressed at higher levels compared to workers (Fig. 3B). These results suggest that the regulation of lifespan in gamergates may share common mechanisms with other organisms.

The caste system in ant societies in general and the social flexibility of *H. saltator* in particular, allow us to study the role of epigenetics on behavior, aging, and development. Here, we use the term epigenetics as the ensemble of molecular pathways that select genomic regions (chromatin domains) for activation or repression, transmit these states through cell division, and stabilize them in differentiated cell types. These mechanisms include DNA methylation, histone post-translational modifications, and *trans*-acting mechanisms involving non-coding RNAs (11).

Unlike in *Drosophila*, DNA methylation pathways in *A. mellifera* and *N. vitripennis* are similar to mammalian systems (5, 12). In the two ant species, we found single copies of cytosine methyl-transferases *DNMT1* and *DNMT3b* genes and the non-catalytic *DNMT3L*, as well as four genes containing methyl-CpG binding domains (Table S11). Consistent with these and previous observations (13), we detected the presence of 5-methylcytosine in the ants genomic DNA. Interestingly, *H. saltator*, which shows more ancestral characteristics, had less DNA methylation than its more derived relative, *C. floridanus* (Fig. 4A).

Histone acetyltransferases (HAT) and deacetylases (HDAC) add and remove acetyl groups on histone tails, regulate genes through chromatin structure (14), and are linked to the aging

process (10). We identified 26 putative HATs in *C. floridanus* and 27 in *H. saltator*, with an expansion of GCNL2 homologues (Table S12). Humans have four classes of HDACs, comprising HDAC1-11 and the NAD^+-dependent sirtuin family proteins (SIRT1-7). In contrast, only 4 HDACs and 6 sirtuins are found in *A. mellifera* (5). We identified 5 HDACs in *C. floridanus* and 4 in *H. saltator*, as well as 6 sirtuins in both ant species (Table S11). Compared to mammals, both ants and honeybee lack SIRT3 (10).

Histone methylation is suspected to participate in epigenetic processes (11). We identified 27 proteins containing the conserved SET histone methyl-transferase domain in *C. floridanus* and 22 in *H. saltator* (Table S13). All major SET families are conserved in ants, including Polycomb/trithorax group, NSDs, SMYDs, and SETDs, and the arginine methyl-transferase family. The SMYD family appears to have undergone multiple duplication events during the evolution of insects (Fig. S6). We found 5 homologs in *C. floridanus* and 6 in *H. saltator* of *SMYD4*, which is involved in muscle development and breast cancer in humans (15, 16). SMYD family members are differentially regulated among different ant castes and developmental stages (Fig. S7); among them, a *SMYD3* homologue (Hsal_14941) is upregulated in gamergates (Fig. 4B), while a *SMYD4* homologue (Hsal_08142) is upregulated in non-reproductive workers (Fig. 4C).

The elaborate social organization in ant colonies is attributed to a sophisticated communication system, made up of chemical signals that elicit behavioral responses via the nervous system (1). Many enriched GO terms associated with differentially expressed genes in *C. floridanus* (Table S14) and *H. saltator* (Table S15) castes are related to neuronal function and chemical communication. In *C. floridanus* major and minor workers we detected 4-fold differences in the expression levels of genes associated with GO terms including “post-synaptic membrane”, “ligand-gated channel activity”, “sensory perception of smell” (Table S14) and “neurotransmitter binding” (Fig. 5A). This suggests that the different behaviors exhibited by distinct workers castes may in part be encoded in the brain at the transcriptional level.

Olfactory receptors (ORs) are G protein-coupled receptors (GPCRs) functioning in behavioral responses and chemical communication in animals (17). We found 139 genes containing the *Drosophila* olfactory receptor domain (IPR004117) in *C. floridanus* and 105 in *H. saltator* (Tables S16–17). Similar to the *A. mellifera* genome, the ant genomes contain fewer gustatory (taste) receptors and odorant binding proteins than *Drosophila* (Table S16–17). The number of identified ORs is much smaller than the number of glomeruli observed in the antennal lobe for both ant species (18, 19), apparently contradicting findings in other insects, which showed that most ORs are individually represented in the brain by a dedicated glomerulus (20). However, our homology-based approach might be insufficient to detect ant-specific ORs.

In ants, cuticular hydrocarbons play a key role in nestmate recognition and regulation of reproduction (21). GO terms associated with the metabolism of hydrocarbons were enriched in antspecific protein families (Table S8), and protein families expanded in ants (Table S18, Fig. S8). *C. floridanus* and *H. saltator* have 19 and 14 genes containing a beta-ketoacyl synthase domain (IPR014030), whereas *A. mellifera* has 3 and *D. melanogaster* has 4. Genes with putative roles in hydrocarbon metabolism show altered expression levels across castes and developmental stages (Fig. S9). Compared to non-reproductive workers, *H. saltator* gamergates upregulated several hydrocarbon metabolism genes, including homologues of a fatty acid synthase (FAS), an acetyl-CoA desaturase (ACOD) and an elongase of very long fatty acid (ELOV) (Fig. 5B). *H. saltator* gamergates advertise their reproductive status to nestmates via long cuticular hydrocarbons (22), and we found that expression of most genes with homology to human ELOVs was upregulated in gamergates (Fig. S10). Because of the

complexity of the metabolic pathways involved, we cannot conclude that transcriptional regulation of these genes is directly linked to communication, however these observations suggest a molecular basis for the interplay between pheromone production, olfaction, and behavior in ants.

Our initial analysis of the genomes of two socially divergent ant species has captured major molecular features that make ants an attractive model for genomic and epigenetic studies. Ant species vary widely in the extent to which eusociality is implemented, ranging from small and less organized colonies to massive and complex societies (1). The diversity and flexibility of ants may provide experimental avenues to address longstanding hypotheses on the relationships among epigenetics, neurobiology, and behavior (23), as well as lifespan regulation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank T. Bloss for ant colony maintenance, B. Kopenhaver for technical help, A. Alekseyenko and J. Wang for bioinformatic support, W. Wang for advice on evolutionary analysis, and P. Voigt and D. Simola for careful reading of the manuscript. R.B. is a fellow of the Helen Hay Whitney Foundation. This work was funded by HHMI Collaborative Innovation Award #2009005 to S.L.B., D.R. and J.L. Raw sequencing data have been deposited in NCBI as SRA020747 (*C. floridanus* genome), SRA020748 (*H. saltator* genome), and GSE22680 (RNA-seq). Assemblies and annotations have been deposited at DDBJ/EMBL/GenBank under the accession AEAB00000000 and AEAC00000000. The versions described in this paper are the first deposited versions: AEAB01000000 and AEAC01000000.

REFERENCES AND NOTES

1. Hölldobler, B.; Wilson, EO. The Ants. Cambridge, MA: Harvard University Press; 1990.
2. Wheeler DE. The American Naturalist. 1986; 128:13.
3. Jemielity S, Chapuisat M, Parker J, Keller L. Age. 2005 Dec 31.27:241. [PubMed: 23598656]
4. Peeters C, Liebig J, Hölldobler B. Insectes Sociaux. 2000 Nov 27.47:325.
5. Consortium HGS. Nature. 2006 Oct 26.443:931. [PubMed: 17073008]
6. Lorenzen MD, et al. Insect Mol Biol. 2007 Jun.16:265. [PubMed: 17316329]
7. Sumitani M, Yamamoto DS, Oishi K, Lee JM, Hatakeyama M. Insect Biochem Mol Biol. 2003 Apr. 33:449. [PubMed: 12650693]
8. Sahin E, Depinho RA. Nature. 2010 Mar 25.464:520. [PubMed: 20336134]
9. Hartmann A, Heinze J. Evolution. 2003 Oct.57:2424. [PubMed: 14628930]
10. Vaquero A, Reinberg D. Genes Dev. 2009 Aug 15.23:1849. [PubMed: 19608767]
11. Allis, CD.; Jenuwein, T.; Reinberg, D. Epigenetics. ed. 1st. Caparros, M., editor. Cold Spring Harbor Laboratory Press; 2007. p. 502
12. Werren JH, et al. Science. 2010 Jan 15.327:343. [PubMed: 20075255]
13. Kronforst MR, Gilley DC, Strassmann JE, Queller DC. Curr. Biol. 2008 Apr 8.18:R287. [PubMed: 18397732]
14. Wang Z, et al. Cell. 2009 Sep 4.138:1019. [PubMed: 19698979]
15. Thompson EC, Travers AA. PLoS One. 2008; 3:e3008. [PubMed: 18714374]
16. Hu L, Zhu YT, Qi C, Zhu YJ. Cancer Res. 2009 May 1.69:4067. [PubMed: 19383909]
17. Hildebrand JG, Shepherd GM. Annu Rev Neurosci. 1997; 20:595. [PubMed: 9056726]
18. Zube C, Kleineidam CJ, Kirschner S, Neef J, Rössler W. J Comp Neurol. 2008 Jan 20.506:425. [PubMed: 18041786]
19. Hoyer SC, Liebig J, Rössler W. Arthropod Structure & Development. 2005 Nov 23.:1.
20. Fishilevich E, Vosshall LB. Curr Biol. 2005 Sep 6.15:1548. [PubMed: 16139209]

21. Blomquist, GJ.; Bagnères, AG. Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology. Cambridge: Cambridge University Press; 2010.
22. Liebig J, Peeters C, Oldham NJ, Markstädter C, Hölldobler B. Proc Natl Acad Sci U S A. 2000 Apr 11.97:4124. [PubMed: 10760282]
23. Dulac C. Nature. 2010 Jun 10.465:728. [PubMed: 20535202]

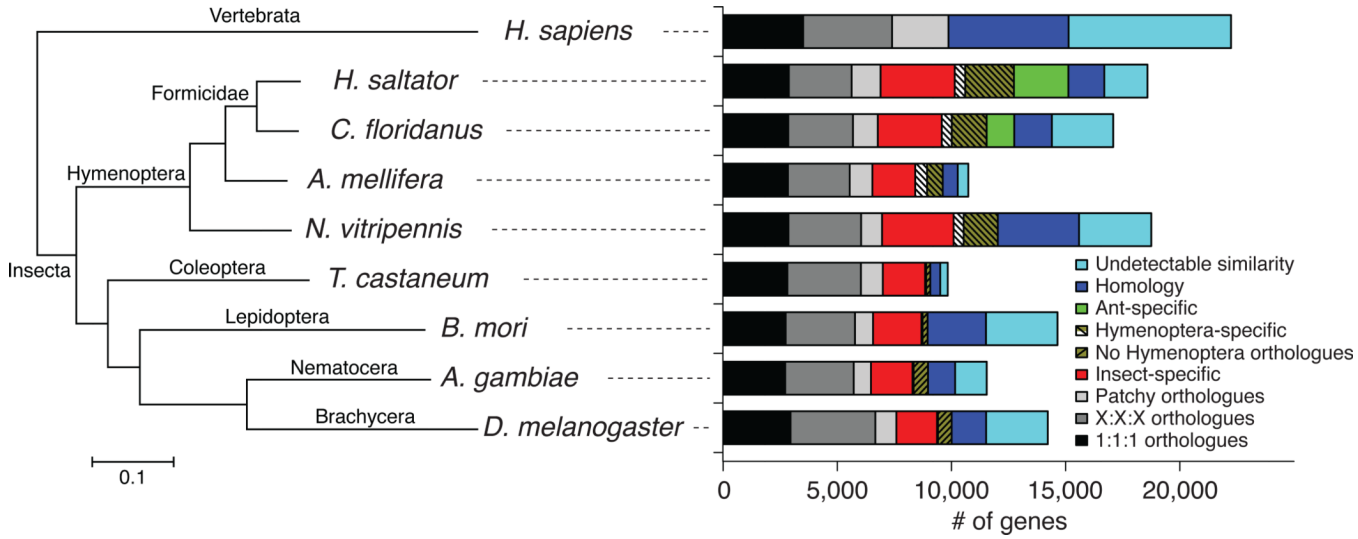

Figure 1.

Ant proteome. Phylogenetic tree based on maximum likelihood analyses of a concatenated alignment of single-copy proteins (left), and orthology relationships in multiple insects (right), using *H. sapiens* as outgroup. The scale bar indicates 0.1 substitution/site.

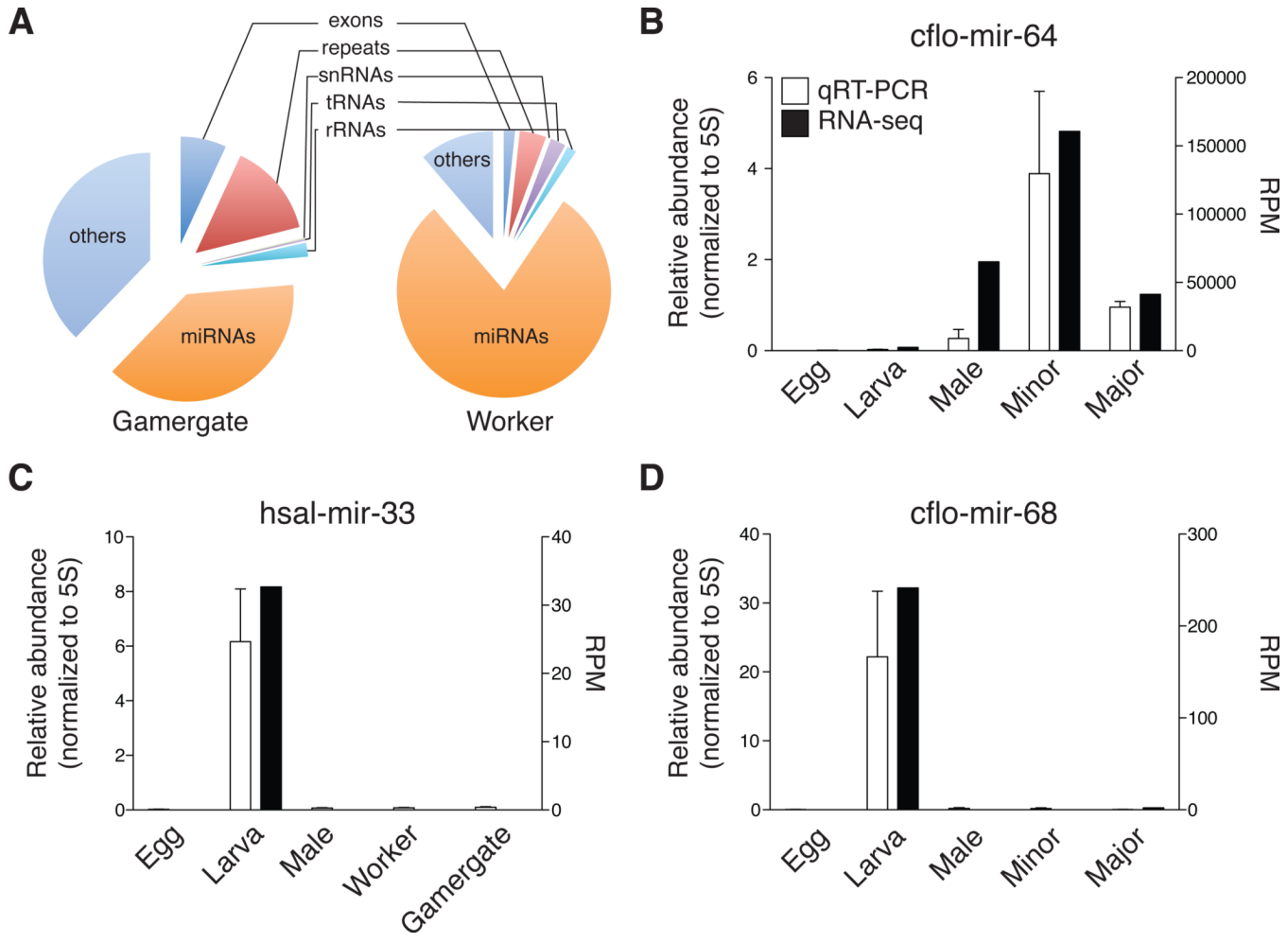

Figure 2.

Small RNA-seq and miRNA. (A) Classification of small RNA-seq reads from *H. saltator* gamergates (left) or workers (right). (B–D) Caste- and stage-specific expression of miRNAs in *C. floridanus* (B and D) and *H. saltator* (C). White bars represent qRT-PCR and indicate the mean + SEM, n=7. Black bars are scaled on the right *y* axis and indicate reads per million (RPM).

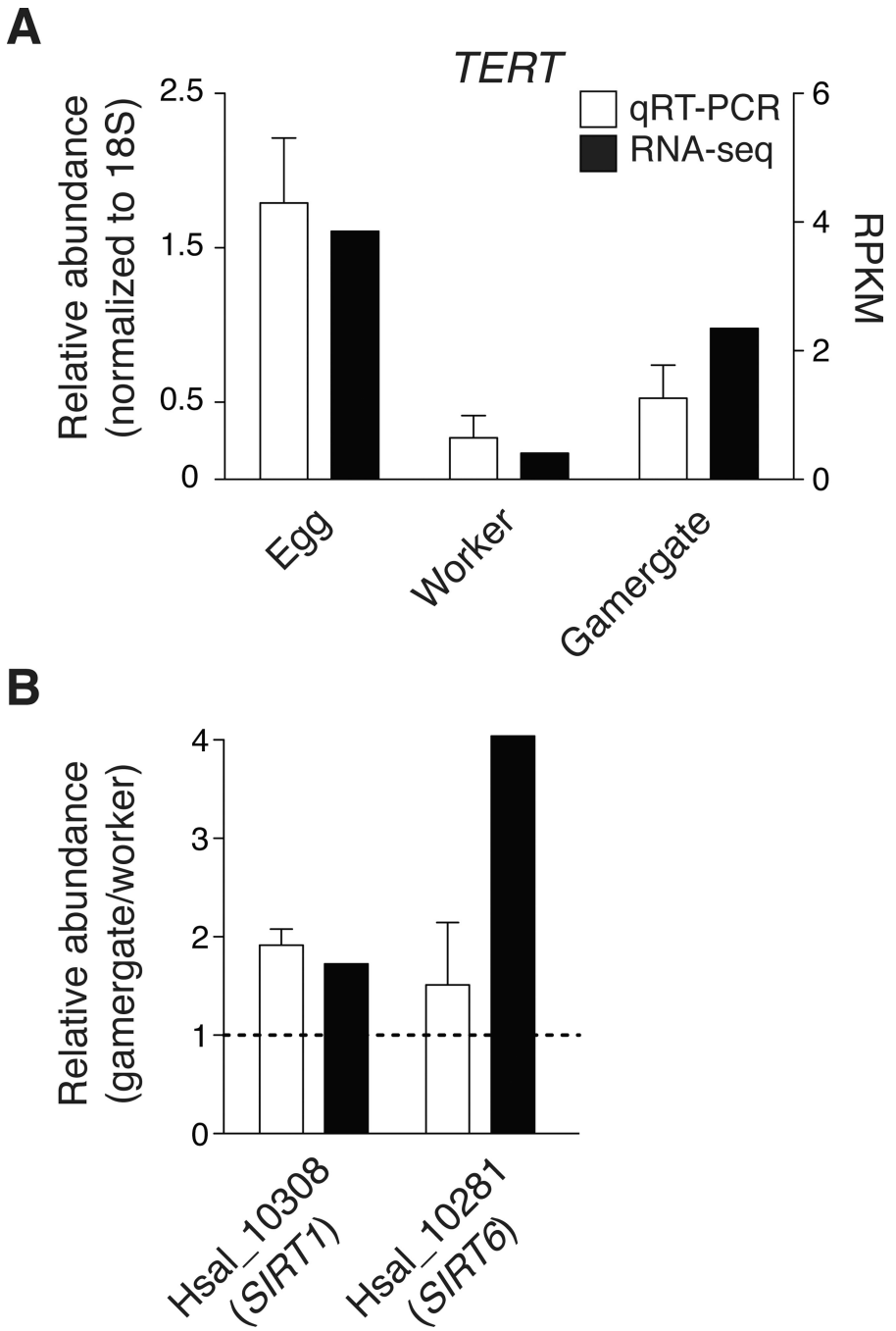

Figure 3.

Aging pathways in *H. saltator*. (A) *TERT* transcript levels in egg, workers, and gamergates. RPKM, read per kilobase per million reads. (B) Transcript levels in gamergates vs. workers for *SIRT1* and *SIRT6* orthologues. White bars, qRT-PCR mean + SEM, n=7. Black bars, RNA-seq.

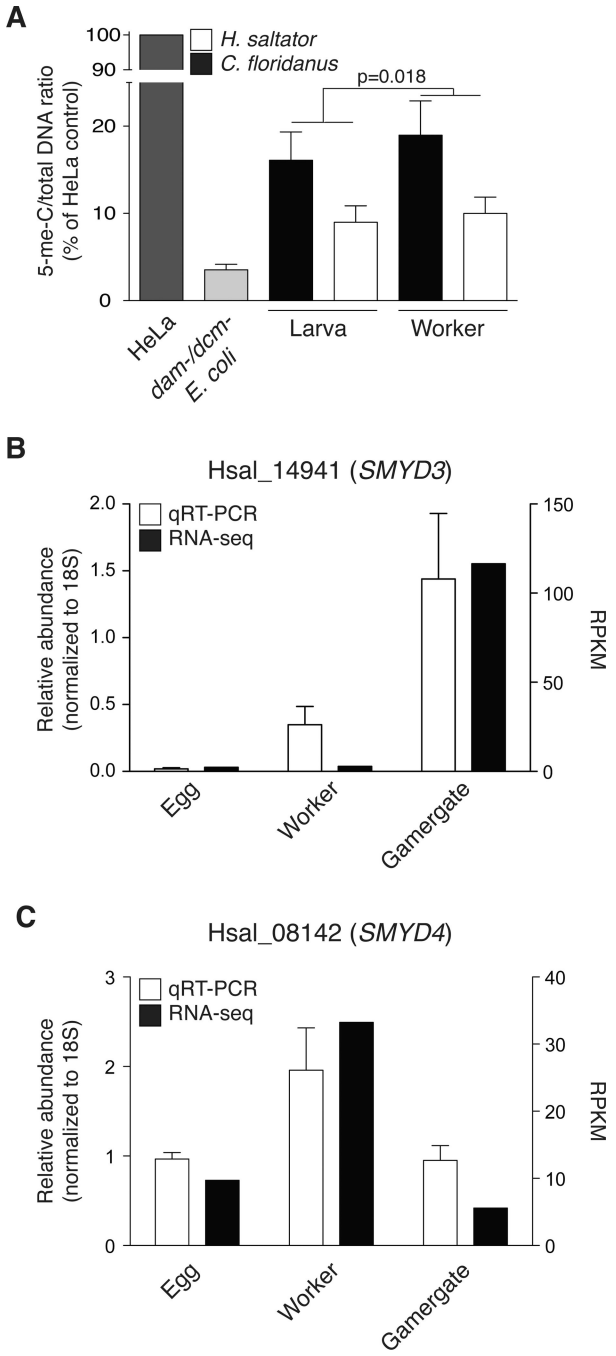

Figure 4.

Potential epigenetic pathways in ants. (A) Ratio of 5-methyl-cytosine/total DNA in *C. floridanus* and *H. saltator*, as determined by dot blot densitometry. Bars show mean + SEM; p value is from repeated measurement ANOVA (Wilks' lambda = 0.52), n=10. (B–C) Quantification of Hsal_14941 (B) and Hsal_08142 (C) expression in different developmental stages and castes. White bars, qRT-PCR mean + SEM, n=7. Black bars, RNA-seq.

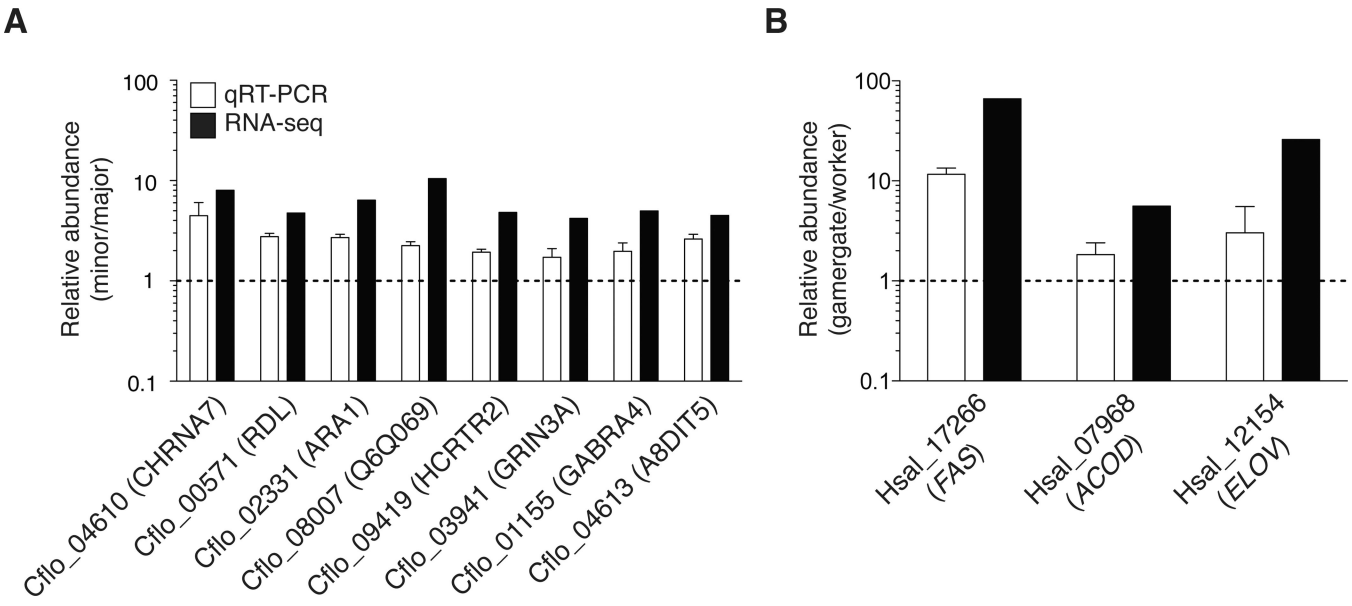

Figure 5.

Neurobiology and communication. (A) Quantification of 8 “GO:neurotransmitter binding” genes in head and thorax from major and minor workers in *C. floridanus*. White bars, qRT-PCR mean + SEM, n=9. Black bars, RNA-seq. (B) Quantification of fatty acid biosynthesis genes in *H. saltator* gamergates compared to non-reproductive workers. White bars, qRT-PCR mean + SEM, n=7. Black bars, RNA-seq.