



Published as: *Curr Biol.* 2012 October 9; 22(19): 1755–1764.

Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*

Roberto Bonasio^{1,*}, Qiye Li^{2,3,*}, Jinmin Lian², Navdeep S. Mutti^{4,†}, Lijun Jin², Hongmei Zhao², Pei Zhang², Ping Wen², Hui Xiang⁵, Yun Ding⁵, Zonghui Jin², Steven S. Shen^{1,6}, Zongji Wang^{2,3}, Wen Wang⁵, Jun Wang^{2,7,8}, Shelley L. Berger^{9,‡}, Jürgen Liebig^{4,‡}, Guojie Zhang^{2,‡}, and Danny Reinberg^{1,‡}

¹Howard Hughes Medical Institute and Department of Biochemistry, New York University School of Medicine, 522 First Avenue, New York, NY 10016, USA

²BGI-Shenzhen, Shenzhen 518083, China

³School of Bioscience & Bioengineering, South China University of Technology, Guangzhou 510006, China

⁴School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

⁵State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China

⁶Center for Health Informatics and Bioinformatics, New York University School of Medicine, 227 E 30th street, Room 7-33, New York, NY 10016

⁷Department of Biology, University of Copenhagen, Universitetsparken 15, København, 2100, Denmark

⁸The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Universitetsparken 15, København, 2100, Denmark

⁹University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

SUMMARY

Background—Ant societies comprise individuals belonging to different castes characterized by specialized morphologies and behaviors. Because ant embryos can follow different developmental trajectories, epigenetic mechanisms must play a role in caste determination. Ants have a full set of DNA methyltransferase and their genomes contain methylcytosine. To determine the relationship between DNA methylation and phenotypic plasticity in ants, we obtained and compared the genome-wide methylomes of different castes and developmental stages of *Camponotus floridanus* and *Harpegnathos saltator*.

Results—In the ant genomes, methylcytosines are found both in CpG and non-CpG contexts and are strongly enriched at exons of active genes. Changes in exonic DNA methylation correlate with

© 2012 Elsevier Inc. All rights reserved.

[‡]To whom correspondence should be addressed: bergers@mail.med.upenn.edu (S.L.B.); Juergen.Liebig@asu.edu (J.L.); zhanggj@genomics.org.cn (G.Z.); Danny.Reinberg@nyumc.org (D.R.).

^{*}These authors contributed equally

[†]Current address: DuPont Agricultural Biotechnology, DuPont Experimental Station, Wilmington, DE, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

alternative splicing events such as exon skipping and alternative splice site selection. Several genes exhibit caste-specific and developmental changes in DNA methylation that are conserved between the two species, including genes involved in reproduction, telomere maintenance, and noncoding RNA metabolism. Several loci are methylated and expressed monoallelically, and in some cases the choice of methylated allele depends on the caste.

Conclusions—These first ant methylomes and their intra- and inter-species comparison reveal an exonic methylation pattern that points to a connection between DNA methylation and splicing. The presence of monoallelic DNA methylation and the methylation of non-CpG sites in all samples suggest roles in genome regulation in these social insects, including the intriguing possibility of parental or caste-specific genomic imprinting.

Keywords

DNA methylation; social insects; ants; epigenetics; splicing

INTRODUCTION

Eusocial insects show extreme phenotypic plasticity, which is particularly pronounced in ant societies, where colony members vary in size, behavior, and physiology [1]. In most ant species, colonies are divided into sexual castes (reproductively active queens, virgin queens, males), and non-reproductive female workers, sometimes divided in distinct sub-castes. With the exception of males, which are haploid and contribute little to the organized life of an ant colony, all ant castes develop from diploid female embryos. Typically, caste determination is not a consequence of genetic differences in these embryos, but occurs in response to environmental stimuli. Therefore, the ant genome must simultaneously encode multiple behavioral, morphological, and physiological phenotypes, among which a specific caste identity is selected for each individual during development and maintained for its lifetime, most likely by epigenetic mechanisms [2, 3].

We recently sequenced the genomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*, which provide intriguing contrasts in behavioral flexibility and social organization [4]. *Camponotus* colonies are organized in a rigid social structure, entirely dependent on the presence of the queen, and with workers that differ in morphology and behavior (majors and minors). When a *Camponotus* queen dies it is not replaced and the colony dies with it. *Harpegnathos* colonies are more flexible, and upon removal of the founding queen, a few dominant workers, called “gamergates”, rise to the social status of acting queens [5]. Although all *Harpegnathos* females are capable of mating and laying fertilized eggs, only queens and gamergates are allowed to produce progeny [6].

These two ant species provide compelling experimental paradigms to investigate epigenetic processes that affect organisms as a whole. In *Camponotus*, alternative developmental trajectories yield fixed phenotypic outcomes: minor, major, and queen adults that look and behave differently. This process is analogous to the process of cell-type differentiation, during which the epigenetic state of pluripotent cells is molded to generate a variety of cell identities all arising from a single genome. In contrast, the worker–gamergate transition in *Harpegnathos* reflects more plastic epigenetic processes, perhaps analogous to the processes of somatic cell reprogramming and transdifferentiation. Thus, we hypothesized that DNA methylation, a well-characterized epigenetic signal involved in differentiation and cell fate decisions in many multicellular organisms [7, 8], may contribute to the phenotypic diversity of ant castes.

DNA methylation is largely absent in four widely studied invertebrates, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Caenorhabditis elegans*, and *Drosophila*

melanogaster. The lack of DNA methylation in Diptera (such as *Drosophila*) appears to be the exception rather than the rule among insects, given that it is present in Lepidoptera, Hemiptera, and Hymenoptera [4, 9–11]. DNA methylation in Hymenoptera may be required for the long-term maintenance of polyphenism in adults, a precondition to caste distinction and social organization. In fact, DNA methylation has been implicated in caste determination and learning in *Apis mellifera* [12, 13].

Here, we report the genome-wide, nucleotide-resolution DNA methylomes for 7 different developmental stages and castes of *Camponotus* and *Harpegnathos*, and we analyze the relationship between DNA methylation, gene expression, and splicing in these social insects.

RESULTS

DNA methylation maps for different developmental stages and adult ant castes

We measured the levels of DNA methylation in embryos, larvae, and 5 adult castes for *Camponotus* and *Harpegnathos* by performing bisulfite conversion and sequencing (BS-seq) of genomic DNA from two libraries (biological replicates) per sample [10]. Anatomical differences between embryos, larvae, and adults and the large amounts of DNA required for BS-seq prohibited the analysis of isolated tissues; therefore, we pooled genomic DNA from whole individuals. Although this strategy yields a complex picture of DNA methylation patterns from various cell types [14], we reasoned that a global DNA methylation profile would still unveil general features, and that inter-caste differences would emerge from the global comparison.

We sequenced ~86 (*Camponotus*) and ~132 (*Harpegnathos*) Gb of bisulfite-converted DNA, which yielded an average depth of ~20× per strand for each sample. More than 92.5% of all cytosines (Cs) were covered by at least two reads per sample. We detected cytosine methylation at ~200,000 sites in *Camponotus* and at ~250,000 sites in *Harpegnathos* (Figure 1A), accounting for 0.3% and 0.21% of all cytosines. After correcting for partially methylated sites, we determined the abundance of mCs at 0.14–0.16% in *Camponotus* and 0.11–0.12% in *Harpegnathos*. The higher ratio of mC/C in *Camponotus* compared to *Harpegnathos* confirms our previous estimates obtained by dot blot analysis [4]. Although this mC/C ratio is lower than in vertebrates, DNA methylation is more prevalent in ants than in the two most established invertebrate model organism, *D. melanogaster*, where it is confined to early embryonic stages [15], and *C. elegans*, which has no DNA methylation at all [16].

Context and degree of cytosine methylation

Methyl-cytosines in eukaryotes are typically found in symmetric CG dinucleotides, although non-CpG sequences (henceforth CH, where H stands for non-G nucleotides) can also be methylated. CH methylation (mCH) is further classified in symmetric mCHG and asymmetric mCHH [7]. In addition to mCGs, we found mCHs in CHG and CHH context in all caste and developmental stages from both species (Figure 1A). Because previous studies on other insects reported that mCHs were attributable to sequencing errors [10, 14], we confirmed their presence in ants by conventional sequencing of 15 loci (Figure S1). Manual verification confirmed that these regions contained mCHs, not only in embryos, where extensive *de novo* DNA methylation is expected, but also in adults (Figure 1B–C). Methylated CHs were mostly in the context of CpA dinucleotides (Figure 1D–E), as in vertebrates [17].

Methylated CGs exhibited a typical bimodal distribution [18], in which the majority of sites were either methylated in > 80% or < 30% of the reads (Figure S2A–B), whereas mCHs were methylated to a much smaller degree (Figure S2C–F). Considering that our methylome

profiles were generated from whole bodies, this difference suggests that mCH in ants is restricted to a small cell population, or that its genomic distribution is highly variable from cell to cell. Methylated CGs were symmetrically methylated on both strands at 75% of sites, whereas less than 1% of mCHGs were symmetrical; therefore, symmetry-based deposition of mCHG is not common in ants.

Methyl-cytosines accumulate at transcribed genes and transposable elements

To quantify the extent of methylation of genomic features, we first calculated the degree of methylation for each mC by dividing the number of methylated reads by the total number of reads covering that cytosine, then summed all these values in the feature of interest and divided by the total number of sites available for methylation. We refer to this value as the “methylation level” of a given region.

Cytosine methylation in both *Camponotus* and *Harpegnathos* exhibited a mosaic distribution, typical of invertebrates [8, 19, 20]: small regions with high methylation levels were interspersed among larger regions devoid of DNA methylation. In all samples, most mCGs were in protein-coding genes, particularly in their coding sequences, with small amounts of methylation also observed at snRNA loci (Figure 2A). We observed similar patterns for mCHG and mCHH (data not shown).

Similar to the case in mammals [21], CpG islands—regions rich in G–C base pairs and CpG dinucleotides—were depleted of mCGs (Figure 2B). Mammalian genomes contain relatively few CpG islands because most CpG dinucleotides are methylated, accumulate unrepaired C–T mutations, and are purged from the genome over evolutionary time. Thus, mammalian CpG islands are preferentially unmethylated and enriched for regulatory functions. In contrast, the ant genome contains a surprisingly high number of CpG dinucleotides [4] and the conventionally accepted parameters that define CpG islands in vertebrates [22] assigned ~5% (*Camponotus*) and ~16% (*Harpegnathos*) of the total genomic sequence to this class. Therefore, the functional significance of these sequences in ants remains unknown, but a mechanism that maintains them in a hypomethylated state seems to exist.

Although transposable elements (TEs) were methylated at genomic background level or lower when taken as a whole (Figure 2A), certain TE classes were more methylated than others (Figure S3A). Several individual TEs were hypermethylated compared to the genome average, whereas others were hypomethylated (Figure S3B). Many TEs showed comparable methylation levels in the two species, but in some cases hypermethylation was species-specific, for example that of L1-Tx1 LINE and Mariner elements in *Camponotus*, and Harbinger and hAT-Blackjack in *Harpegnathos* (Figure S3B). Methylation of TEs correlated positively with their expression level (Figure S3C), suggesting that DNA methylation preferentially marks active TEs. Interestingly, in both species the targets of repeat-associated small RNAs (rasRNAs) showed the highest methylation levels in males (Figure S3D). Given that degenerating testis tissue and stored sperm account for a large fraction of body weight in male ants, this observation suggests that rasRNAs target preferentially active transposons in the male germline.

DNA methylation peaks at the start of the second exon

In the body of the average ant gene, methylation increased sharply at the ATG and decreased in the 3' direction, returning to background levels at the stop codon (Figure 2C, S4A). This pattern was virtually identical across castes and developmental stages and between the two species. DNA methylation peaked ~750 bp downstream of the ATG (Figure S4B), which corresponds to the start of the second exon of methylated genes (Figure S4C). In contrast, the average start of the second exon when all genes were considered—

regardless of their methylation status—was at +1,475 bp (*Camponotus*) and +1,525 bp (*Harpegnathos*). In fact, the average intron size between methylated and unmethylated genes varied greatly, and this difference was also observed in *A. mellifera*, but not *B. mori* (Figure S4C).

Within the gene body, mCs accumulated on exons (Figure 2D, S4D), and were largely absent from introns, as observed in other organisms [23]. We reanalyzed genome-wide DNA methylation data for *A. mellifera* [14] (Figure S4E) and *B. mori* [10] and found similar patterns, although in the latter introns contained detectable amounts of DNA methylation (Figure S4F). The high mC density in exons compared to introns in ants and other organisms suggests a link between DNA methylation and the transcription/splicing machinery (see below).

Relationship between caste-specific methylomes and transcriptomes

To compare DNA methylation and gene expression, we utilized our published RNA-seq datasets [4] and integrated them with newly generated datasets for mature queens and virgin queens. Genome-wide, methylation levels exhibited for the most part a positive correlation with RNA levels, except for the most highly expressed genes (Figure 3A–B), as originally found in plants and other organisms [19, 24]. Methylated genes exhibited lower sample specificity [25] than unmethylated genes (Figure 3C–D), which suggests that they are enriched for constitutively expressed housekeeping genes, which contribute RNA and methylated genomic DNA from all tissues and are therefore more represented in whole-body samples. This is consistent with predictions based on genome-wide patterns of CpG depletion in *A. mellifera* [26].

At the single gene level, the correlation between methylation and RNA levels was not as striking as it was genome-wide (Figure 3E–F), which may explain why the unsupervised hierarchical clustering of embryos, larvae, and adult castes according to their expression or methylome profiles yielded trees with different topologies (Figure S5). In *Camponotus*, the methylome profile drew the clearest distinction between sexual castes and the remaining samples (Figure S5A), whereas at the transcriptome level the highly reproductive queen was drastically different from the other adult castes (Figure S5B). In *Harpegnathos*, the DNA methylation profiles of gamergates and workers were more similar than their transcriptomes (Figure S5C–D), suggesting that not all the transcriptional changes that accompany the worker–gamergate transition translate into stable epigenetic modifications.

Differentially methylated genes and their conservation between ant species

Next, we sought to identify genes with differences in methylation levels. Given that differences in methylation concentrated in small patches within genes, rather than on the entire gene length (Figure 4A–B), we used a 200 bp sliding window to identify genes containing at least one differentially methylated region with > 2-fold changes in methylation levels in all 21 pairwise sample comparisons. In some cases differences in methylation were associated with changes in splicing patterns (Figure 4A), and in others with changes in gene expression (Figure 4B). The number of differentially methylated genes (Figure 4C–D) in each comparison reflected our anticipations based on the biology of these organisms. For example, major and minor workers in *Camponotus* differed by few methylated genes (9), and the number of differentially methylated genes between *Harpegnathos* worker and gamergates (38) was the lowest for this species, consistent with the similarities of their methylomes genome-wide (Fig. S5C). Nonetheless, gamergates had more methylated genes in common with the reproductive queen, suggesting that methylation of at least some of these genes may correlate with the achievement of a reproductive dominant status.

Using our homology-based functional annotation [4], we analyzed the enrichment of gene ontology (GO) terms in genes differentially methylated between adult castes (Table S1–S2). The 9 genes differentially methylated in *Camponotus* minor workers compared to major workers were enriched for GO terms related to fatty acid metabolism (Table S1), which are also associated with differentially expressed, ant-specific genes [4]. Reflecting a more extensive methylome distinction, genes differentially methylated in reproductive compared to non-reproductive individuals were enriched for more diverse GO terms related to metabolic processes, GTPase signaling pathways, and chromatin processes (“chromatin remodeling complex”, “H3K36 demethylation”), among others.

We identified genes differentially methylated in *Camponotus* queens compared to workers and asked whether their orthologs in *Harpegnathos* were also differentially methylated. In all comparisons the number of genes displaying conserved patterns of methylation changes across species was higher than what was expected by chance (Table S3). Half of these genes belonged to three functional categories: reproductive biology, telomere maintenance, and noncoding RNA metabolism (Table S4), whereas genes with conserved differences in methylation between embryos and larvae included several that function in larval development in other organisms (Table S5). These observations suggest that DNA methylation of certain genes has been associated with caste identity and possibly with the regulation of reproduction ever since the ancestors of *Camponotus* and *Harpegnathos* diverged more than 100 million years ago.

DNA methylation levels are altered near alternative splice sites

Previous analyses in *A. mellifera* [14, 27] and humans [28], our observation that mCs accumulate in exons (Figure 2), and the fact that patches of mCs were observed near alternative splicing sites (Figure 4A), all pointed to a link between splicing and DNA methylation. Indeed, skipped exons exhibited significantly ($P < 0.01$) lower methylation level than randomly selected exons (Figure 5A), and the presence of alternative splice site in 5′ or 3′ affected the methylation levels of upstream and downstream exons (Figure 5B–C). The fact that these changes were relatively minor suggests that the link between DNA methylation and alternative splice site selection is either restricted to a subset of genes or that additional RNA-seq data is required to improve the detection of alternative isoforms. Nonetheless, these results and observations in other species [27, 28] support a connection between DNA methylation and the selection of alternative exons and splice sites.

To further investigate the relationship between DNA methylation and alternative splicing, we manually analyzed the *Camponotus* homolog of *lipophorin receptor 2* (*Cflo_09743*), a gene involved in oogenesis [29], the *Harpegnathos* homolog of *ciboulot* (*Hsal_08119*), a gene involved in caste determination in termites [30], and the *Harpegnathos* endonuclease G gene (*Hsal_05204*). In the first two cases, inclusion of an alternative exon correlated with hypomethylation (Figure S6A–B), as observed for the *alk* gene in honeybees [27] and *CD45B* in human cells [28], whereas exon inclusion in *Hsal_05204* correlated with hypermethylation (Figure S6C). This suggests that, if DNA methylation changes affect the inclusion rates of exons, they likely do so through recruitment of (or interference with) different factors in different genes.

Monoallelic DNA methylation correlates with monoallelic gene expression

In vertebrates, allele-specific DNA methylation (ASM) underpins important epigenetic phenomena such as X chromosome inactivation [31] and parental imprinting [32], and ASM at gene promoters correlates inversely with allele-specific expression [33]. To our knowledge, this aspect of DNA methylation has not been investigated in invertebrates. Using single nucleotide polymorphisms (SNPs), we assigned each BS-seq read to one of two

alleles in each sample and we detected patches of ASM in all samples analyzed (Table S6), although only regions with informative A or G SNPs could be interrogated.

Some cases of ASM were caste-specific; for example, in *Camponotus*, an ASM region was methylated on allele #1 in non-reproductives and allele #2 in reproductive individuals (Figure 6A). This region mapped to *Cflo_11155*, a conserved gene involved in reproduction and gamete generation in *C. elegans* and preferentially expressed in *Drosophila* ovaries. Among the regions displaying ASM in *Harpegnathos*, we found one that was devoid of methylation in the embryos but acquired ASM in the adults (Figure 6B). In all samples, ASM associated with allele-specific expression (Figure 6C–D), supporting a relationship between DNA methylation and gene expression in these regions.

DISCUSSION

We generated the first single-nucleotide resolution DNA methylomes in ants with the goal of understanding the relationship between this epigenetic mark and the extensive polyphenism observed in these social insects. Although genetic effects in ant caste determination have been observed [34], they are considered maladaptive in monogynous and monandrous species (such as *Camponotus* and *Harpegnathos*) and thus unlikely to be relevant for this study. Furthermore, our extensive re-sequencing efforts have failed to uncover any allelic bias between *Camponotus* castes.

By analyzing the DNA methylation maps of two developing stages and 5 adult castes in *Camponotus* and *Harpegnathos*, we identified conserved features of the ant methylomes: 1) the presence of non-CpG methylation in developing and adult individuals; 2) the accumulation of mCs on exons; 3) the existence of ASM and its correlation with allele-specific expression. Some differentially methylated genes were conserved in the two species, but many were species-specific, reflecting evolutionary divergence in the targets of the DNA methylation pathway.

Most mCs in eukaryotes are in the context of symmetric CpG dinucleotides [2]; however, low levels of mCHH and mCHG are also found in mouse and human embryonic stem (ES) cells [17, 21]. It has been argued that mCH in animals is a byproduct of DNMT3A, which has low sequence specificity, because they have been detected only in cells with strong *de novo* methyltransferase activity, such as ES cells [17]. The presence of non-CpG methylation in adult ants was unexpected and suggests the possibility of biologically functional non-CpG methylation in ants.

DNA methylation in ants was most prominent inside gene bodies, and in particular on transcribed exons. Our data are consistent with previous findings showing that the bulk of DNA methylation in invertebrates and plants accumulates on intermediate-to-high expression genes [14, 19, 23, 24, 35]. However, the correlation between methylation levels and gene expression was not absolute, indicating that the presence of DNA methylation on the bodies of these genes may reflect a separate layer of regulation, such as that of alternative splicing.

The precise origin and ultimate function of gene body DNA methylation, which is found in all organisms that express DNMTs [21, 35], remains unknown. It was suggested that gene body methylation protects transcribed genes from aberrant transcriptional initiation in the wake of RNA polymerase [36], but recent studies argued against this model [7, 24]. We cannot explain how the DNA methylation machinery distinguishes exons from introns in the DNA sequence, but we speculate that it must result from a molecular cross-talk with the transcription and splicing machinery. Observations in *A. mellifera* [14, 27] and our ant methylomes point to a connection between DNA methylation and the regulation of splicing,

which may be conserved with other organisms, including humans [28]. Indeed, in human cells DNA methylation regulates splicing by inhibiting CTCF binding, which alters the rate of transcription by RNA polymerase and affects exon inclusion [28]. This and other mechanisms may be at work in ants, and although most of the changes that we observe are subtle, they may reflect more pronounced tissue-specific changes that are masked by the heterogeneity of whole-body samples.

Our observations raise intriguing questions regarding the mechanistic details of DNA methylation pathways in ants. In particular, how do methylation patterns change in response to developmental and environmental cues? In honeybees caste determination has a strong nutritional component, such that larvae fed “royal jelly” develop into queens [37]. No equivalent dietary input is known in ants, and, at least for *Harpegnathos*, it seems unlikely, given that larvae feed on live prey captured by workers [38]. Thus, other developmental cues must initiate a cascade of events that culminates in caste polyphenism, but whether and how such signals translate into differential methylation remains to be explained. DNMTs interact with a variety of nuclear components, including transcription factors, epigenetic regulators, histone modification, and noncoding RNAs [2, 39, 40], and some of these may guide caste-specific DNA methylation in ants. The refinement of molecular techniques such as RNAi and transgenesis will be required before we can determine if the observed caste-specific DNA methylation patterns play a direct role in establishing caste identity in ants.

Finally, the discovery of ASM in ants opens a new avenue of investigation on the role of DNA methylation in shaping caste identity and social behavior. Based on the peculiar genetics of haplodiploid sex determination and eusocial living, it was proposed that parental imprinting may be prominent in social insects [41–44]. We identified several loci that exhibit ASM as well as allele-specific expression, which, in mammals, is a sign of parental imprinting. Because of technical limitations we could not trace methylated and unmethylated alleles to the parent of origin and therefore we cannot formally conclude that parental imprinting caused the ASM that we observed. However, at least one case of parent-of-origin effects on social behavior was described in honeybees [45] and it would be interesting to determine whether it originates from imprinted DNA methylation.

MATERIALS AND METHODS

Camponotus and *Harpegnathos* colonies were housed and reared as described [4]. 20 µg of genomic DNA or 10 µg of RNA were isolated for each sample. Gut and poison gland were removed from all adults to minimize microbial contamination and degradation of nucleic acids. Bisulfite conversion was carried out with a modified NH_4HSO_3 -based protocol [46], and polyA+ RNA was selected for RNA-seq. Libraries were sequenced on Illumina Genome Analyzers. Short reads were aligned *Camponotus* and *Harpegnathos* genomes v3.5 with SOAP2 [47]. Cs in BS-seq reads that matched to Cs on the reference genome were counted as potential mCs. We calculated the false positive rate using a non-methylated control and utilized it to determine true positive mCs with a statistical model. We then assigned methylation rates to all mCs that passed statistical filtering. We defined “methylation level” as the sum of all methylation rates in the region of interest divided by the total number of covered Cs.

AS events were detected using junction reads identified by TOPHAT [48]. To ensure the accuracy and reliability of junction reads, at least 8 bp with no mismatch were required on each side of the exon junction. Sample specificity for methylated and unmethylated genes was calculated following Yanai *et al.* [25], with appropriate modifications for DNA methylation data.

Two-way analysis of variance (two-way ANOVA) was used to identify differentially methylated regions (DMRs) between two samples using a 200 bp sliding window with a step length of 100 bp. Given that DNA methylation may affect genes at the exon level, we reasoned that an exon-size window would have the best chances to detect DMRs. To ensure the power of statistical test, only windows with at least 6 (3 per strand) informative CpGs ($3\times$ coverage) in both replicates of the two comparable samples were considered.

To detect allele-specific methylation and expression, we determined SNPs with SOAPsnp [4, 49], then assigned BS-seq reads to one of two alleles for all heterozygous SNPs. Only regions with FDR-adjusted P -value < 0.05 were considered. SNPs linked to regions exhibiting ASM and overlapping exons of protein-coding genes were used to detect expression levels of hyper- and hypo-methylated alleles using RNA-seq.

Sequencing data generated for this study have been deposited in the National Center for Biotechnology Information as GSE31577. This comprises the sub-series GSE31344 (small RNA-seq), GSE31346 (RNA-seq), and GSE31576 (BS-seq). Previously published sequencing data analyzed in this publication are available in the GSE22680 super-series.

Additional details are described in the Supplemental Material and Methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank T. Bloss and K. Haight for ant colony maintenance; B. Kopenhaver for technical help; D. Beck, C. Desplan, D. Simola, and S. Tu for critical reading of the manuscript; H. Shao for help with the mC detection pipeline. R.B. was supported by a Helen Hay Whitney Foundation post-doctoral fellowship. This work was funded by a Howard Hughes Medical Institute Collaborative Innovation Award (#2009005) to S.L.B., D.R., and J.L.

References

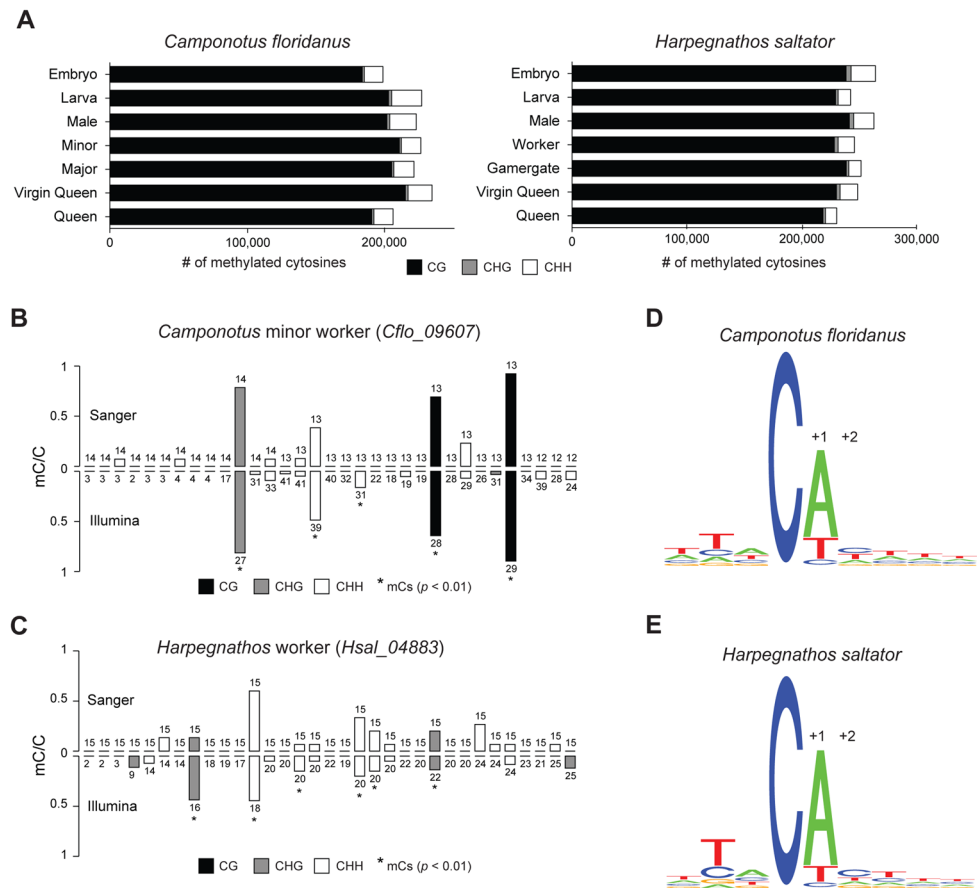
1. Hölldobler, B.; Wilson, EO. *The Ants*. Cambridge, MA: Harvard University Press; 1990.
2. Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. *Science*. 2010; 330:612–616. [PubMed: 21030644]
3. Bonasio R. Emerging topics in epigenetics: ants, brains, and noncoding RNAs. *Ann N Y Acad Sci*. 2012
4. Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, Qin N, Donahue G, Yang P, Li Q, Li C, et al. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science*. 2010; 329:1068–1071. [PubMed: 20798317]
5. Peeters C, Liebig J, Hölldobler B. Sexual reproduction by both queens and workers in the ponerine ant *Harpegnathos saltator*. *Insectes Sociaux*. 2000; 47:325–332.
6. Liebig J, Peeters C, Hölldobler B. Worker policing limits the number of reproductives in a ponerine ant. *Proc R Soc Lond B*. 1999; 266:1865–1870.
7. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*. 2010; 11:204–220. [PubMed: 20142834]
8. Suzuki M, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet*. 2008
9. Walsh TK, Brisson JA, Robertson HM, Gordon K, Jaubert-Possamai S, Tagu D, Edwards OR. A functional DNA methylation system in the pea aphid, *Acyrtosiphon pisum*. *Insect Mol Biol*. 2010; 19(Suppl 2):215–228. [PubMed: 20482652]
10. Xiang H, Zhu J, Chen Q, Dai F, Li X, Li M, Zhang H, Zhang G, Li D, Dong Y, et al. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat Biotechnol*. 2010; 28:516–520. [PubMed: 20436463]

11. Wang Y, Jorda M, Jones PL, Maleszka R, Ling X, Robertson HM, Mizzen CA, Peinado MA, Robinson GE. Functional CpG methylation system in a social insect. *Science*. 2006; 314:645–647. [PubMed: 17068262]
12. Kucharski R, Maleszka J, Foret S, Maleszka R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science*. 2008; 319:1827–1830. [PubMed: 18339900]
13. Lockett GA, Helliwell P, Maleszka R. Involvement of DNA methylation in memory processing in the honey bee. *Neuroreport*. 2010; 21:812–816. [PubMed: 20571459]
14. Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R. The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol*. 2010; 8:e1000506. [PubMed: 21072239]
15. Lyko F, Ramsahoye BH, Jaenisch R. DNA methylation in *Drosophila melanogaster*. *Nature*. 2000; 408:538–540. [PubMed: 11117732]
16. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev*. 2002; 16:6–21. [PubMed: 11782440]
17. Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A*. 2000; 97:5237–5242. [PubMed: 10805783]
18. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*. 2008; 454:766–770. [PubMed: 18600261]
19. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*. 2010; 328:916–919. [PubMed: 20395474]
20. Suzuki MM, Kerr ARW, De Sousa D, Bird A. CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Res*. 2007; 17:625–631. [PubMed: 17420183]
21. Lister R, Pelizzola M, Dowen R, Hawkins R, Hon, Tonti-Filippini J, Nery J, Lee L, Ye Z, Ngo Q, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009
22. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci USA*. 2002; 99:3740–3745. [PubMed: 11891299]
23. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, et al. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A*. 2010; 107:8689–8694. [PubMed: 20395551]
24. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell*. 2006; 126:1189–1201. [PubMed: 16949657]
25. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*. 2005; 21:650–659. [PubMed: 15388519]
26. Foret S, Kucharski R, Pittelkow Y, Lockett GA, Maleszka R. Epigenetic regulation of the honey bee transcriptome: unravelling the nature of methylated genes. *BMC Genomics*. 2009; 10:472. [PubMed: 19828049]
27. Foret S, Kucharski R, Pellegrini M, Feng S, Jacobsen SE, Robinson GE, Maleszka R. DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *P Natl Acad Sci USA*. 2012
28. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*. 2011
29. Tufail M, Takeda M. Insect vitellogenin/lipophorin receptors: molecular structures, role in oogenesis, and regulatory mechanisms. *Journal of insect physiology*. 2009; 55:87–103. [PubMed: 19071131]
30. Koshikawa S, Cornette R, Matsumoto T, Miura T. The homolog of Ciboulot in the termite (*Hodotermopsis sjostedti*): a multimeric beta-thymosin involved in soldier-specific morphogenesis. *BMC Dev Biol*. 2010; 10:63. [PubMed: 20529303]

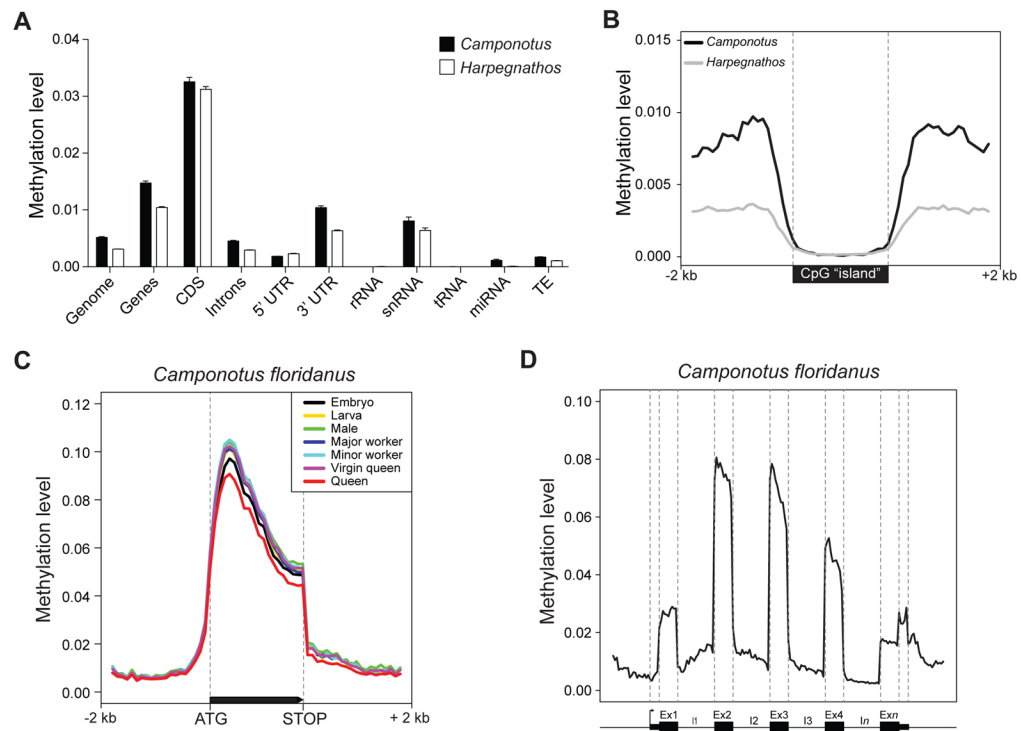
31. Payer B, Lee JT. X chromosome dosage compensation: how mammals keep the balance. *Annu Rev Genet.* 2008; 42:733–772. [PubMed: 18729722]
32. Barlow, DP.; Bartolomei, MS. Genomic Imprinting in Mammals. In: Allis, CD.; Jenuwein, T.; Reinberg, D., editors. *Epigenetics*. Cold Spring Harbor Laboratory Press; 2007.
33. Li Y, Zhu J, Tian G, Li N, Li Q, Ye M, Zheng H, Yu J, Wu H, Sun J, et al. The DNA methylome of human peripheral blood mononuclear cells. *PLoS Biol.* 2010; 8:e1000533. [PubMed: 21085693]
34. Schwander T, Lo N, Beekman M, Oldroyd BP, Keller L. Nature versus nurture in social insect caste differentiation. *Trends Ecol Evol.* 2010; 25:275–282. [PubMed: 20106547]
35. Lee TF, Zhai J, Meyers BC. Conservation and divergence in eukaryotic DNA methylation. *Proc Natl Acad Sci U S A.* 2010; 107:9027–9028. [PubMed: 20457928]
36. Zilberman D, Gehring M, Tran R, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature genetics.* 2006; 39:61–69. [PubMed: 17128275]
37. Beetsma J. The process of queen-worker differentiation in the honeybee. *Bee World.* 1979; 60:24–39.
38. Penick CA, Liebig J. Regulation of queen development through worker aggression in a predatory ant. *Behavioral Ecology.* 2012
39. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, et al. The *Polycomb* group protein EZH2 directly controls DNA methylation. *Nature.* 2006; 439:871–874. [PubMed: 16357870]
40. Schmitz KM, Mayer C, Postepska A, Grummt I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes & development.* 2010; 24:2264–2269. [PubMed: 20952535]
41. Queller DC, Strassmann JE. The many selves of social insects. *Science.* 2002; 296:311–313. [PubMed: 11951035]
42. Haig D. Intragenomic conflict and the evolution of eusociality. *Journal of theoretical biology.* 1992; 156:401–403. [PubMed: 1434666]
43. Queller DC. Theory of genomic imprinting conflict in social insects. *BMC Evol Biol.* 2003; 3:15. [PubMed: 12871603]
44. Dobata S, Tsuji K. Intragenomic conflict over queen determination favours genomic imprinting in eusocial Hymenoptera. *Proc Biol Sci.* 2012
45. Guzman-Novoa E, Hunt GJ, Page RE Jr, Uribe-Rubio JL, Prieto-Merlos D, Becerra-Guzman F. Paternal effects on the defensive behavior of honeybees. *J Hered.* 2005; 96:376–380. [PubMed: 15743904]
46. Hayatsu H, Tsuji K, Negishi K. Does urea promote the bisulfite-mediated deamination of cytosine in DNA? Investigation aiming at speeding-up the procedure for DNA methylation analysis. *Nucleic Acids Symp Ser (Oxf).* 2006:69–70.
47. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics.* 2009; 25:1966–1967. [PubMed: 19497933]
48. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009; 25:1105–1111. [PubMed: 19289445]
49. Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K. SNP detection for massively parallel whole-genome resequencing. *Genome Res.* 2009; 19:1124–1132. [PubMed: 19420381]

HIGHLIGHTS

- DNA methylation in ants accumulates at CpGs on exons of transcribed genes
- Changes in DNA methylation accompany alternative exon and splice site selection
- Differentially methylated genes regulate reproduction, telomeres, ncRNA metabolism
- Patches of monoallelic methylation correlate with monoallelic gene expression

**Figure 1.**

CpG and non-CpG methylation in the ant genome. A) Total number of mCGs (black), mCHGs (grey), and mCHHs (white) in the indicated sample from both ant species. B–C) Validation of CH methylation in adult individuals. Bars represent single, contiguous cytosines in the indicated locus. The fraction of clones (above the x axis) or Illumina reads (below the x axis) that support methylation is plotted on the y axis. Numbers indicate the total number of clones analyzed (top) or Illumina reads mapped to the site (below). Asterisks indicate sites that were determined to be mCs by BS-seq at an FDR-adjusted P -value < 0.01 . D–E) Sequence context of CH methylation; the first and second base after the methylated cytosine are indicated by numbers.

**Figure 2.**

Distribution of mCGs on coding and non-coding regions. A) The methylation level for each indicated genomic feature is plotted on the y axis. Values were calculated separately for each caste and developmental stage and the averages + s.e.m. are shown ($N = 7$). B) Methylation profile of genome regions with high GC content (>55%) and high CpG O/E (> 0.65). C) The average methylation level of CG sites along the body of all complete protein coding genes is plotted on the y axis for the indicated castes and developmental stages of *Camponotus*. Genes were divided in 20 bins and the methylation level was calculated for each bin of each gene and the average for all genes is shown. D) Methylation profile over the average gene body in *Camponotus*. Exons and introns are shown separately, as well as 5' and 3' UTR (thinner boxes).

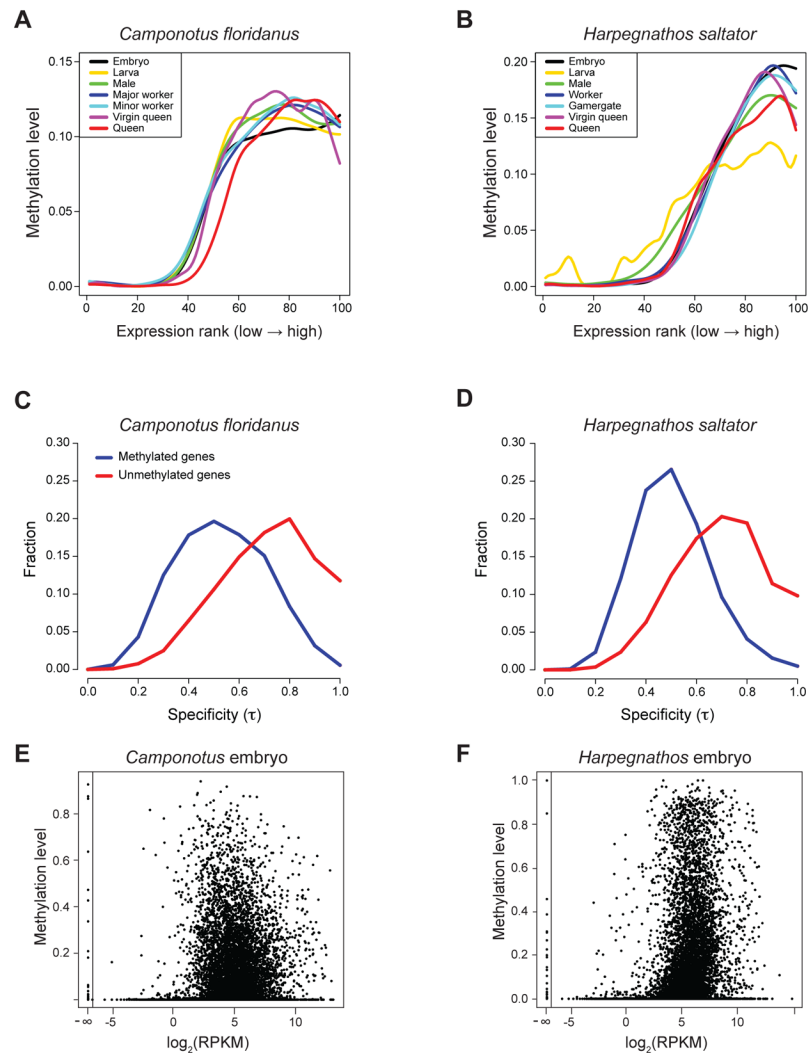


Figure 3.

Gene body methylation and gene expression. A–B) Genes were binned from 0 (least expressed) to 100 (most expressed), their expression rank plotted on the x axis and their methylation level plotted on the y axis. C–D) Specificity index (see text for details) for methylated and unmethylated genes. E–F) Scatter plot for methylation level versus RNA levels (log-transformed read per kilobase per million). Each point is an individual gene.

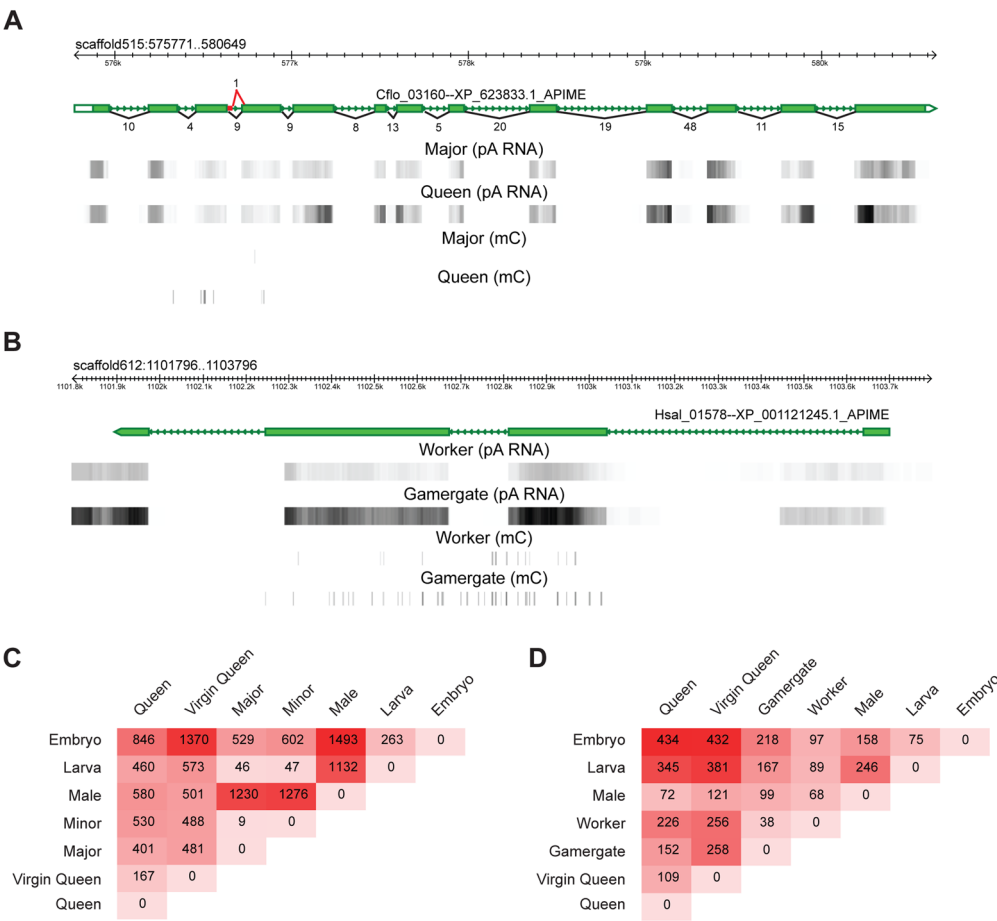
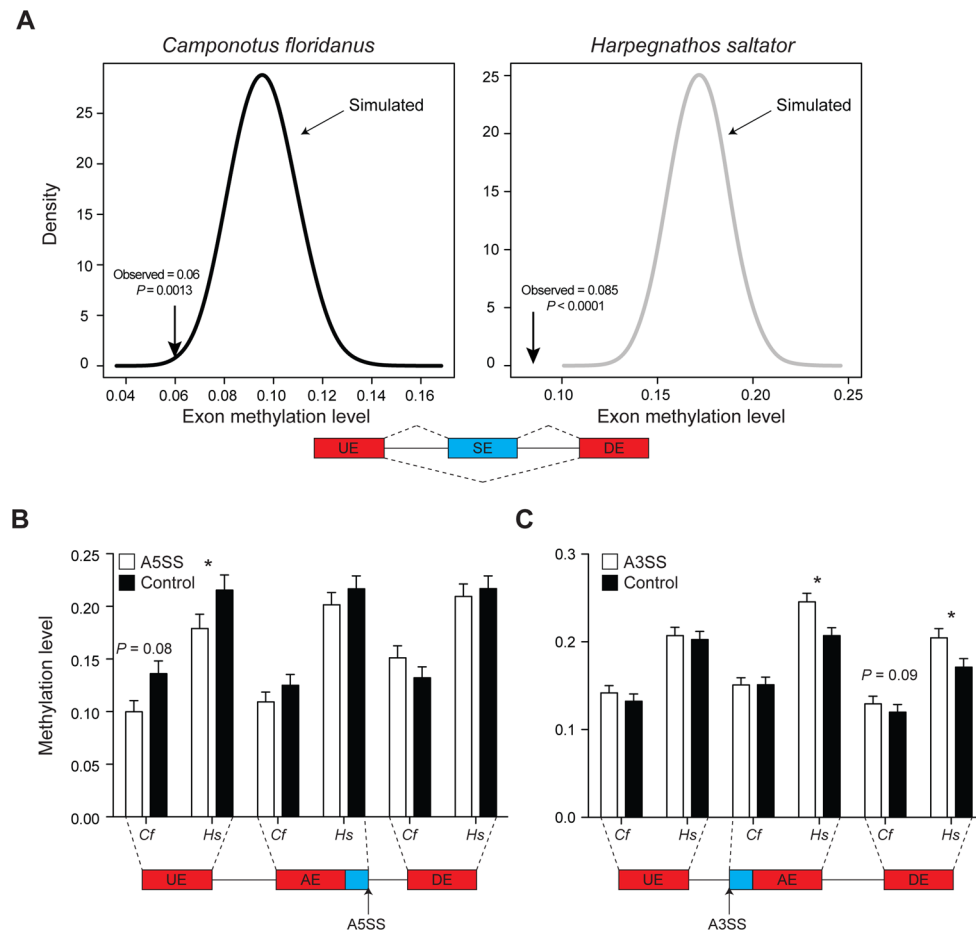
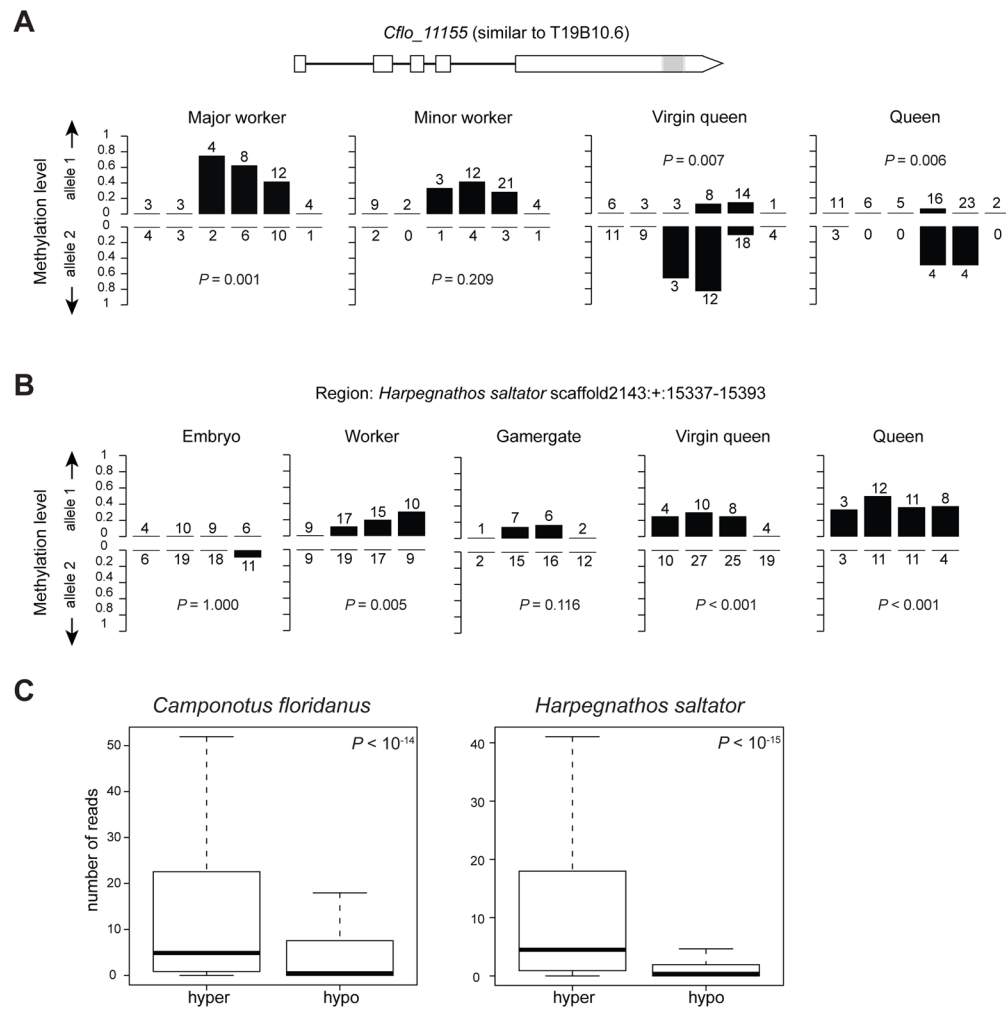


Figure 4. Differentially methylated genes. A) An example of a differentially methylated gene in *Camponotus*, where the differentially methylated region is very limited and coincides with the position of an alternative splicing event (thinner red box). The number of high-confidence junction reads from major worker RNA-seq is indicated in the gene model. B) An example of a differentially methylated gene in *Harpegnathos*, which shows an extended region hypermethylated in gamergates, where the gene is expressed at higher level. In both panels, green boxes indicate annotated exons; pA RNA tracks indicate normalized RNA-seq signal by intensity of the black coloring; mC tracks show methylated cytosines as vertical lines and their degree of methylation is indicated by the intensity of black coloring (black, all reads methylated; white, all reads unmethylated). C) Heatmap for the total number of differentially methylated genes containing at least one 200 bp region with 2-fold difference in methylation levels in each pairwise comparison in *Camponotus*. D) Differentially methylated genes heatmap for *Harpegnathos*.

**Figure 5.**

DNA methylation and alternative splicing in ants. A) The methylation level for randomly selected exons from the embryonic methylome of *Camponotus* (left) or *Harpegnathos* (right) was calculated for 10,000 simulations using randomly selected exons from comparably expressed genes and is shown as a bell-shaped curve. The average value for skipped exons (SE) is indicated by the arrow. B) Methylation level of upstream (UE), affected (AE) and downstream (DE) exons in regions with alternative 5' splice sites (A5SS) in *Camponotus* (Cf) and *Harpegnathos* (Hs) (white bars) compared to the methylation level of randomly selected exons (black bars). Bars show mean + s.e.m. *, $P < 0.05$. P -values higher than 0.05 but lower than 0.1 are indicated and were determined with a non-parametric Mann-Whitney test. C) Same as in B) but for alternative 3' splice sites (A3SS).

**Figure 6.**

Allele-specific methylation. A–B) Each bar represents a single CpG in the analyzed locus. The fraction of methylated reads belonging to an arbitrarily defined allele #1 is plotted above the x axis, and the fraction of methylated reads belonging to allele #2 is plotted below the x axis. The total number of informative reads covering each C on each allele is indicated above or below the bars. FDR-adjusted P -values for the null hypothesis (no allele-specific methylation) were calculated with a simulation process repeated for 100,000 cycles. A) *Camponotus* gene *Cflo_11155*, the region affected by ASM is shaded in gray. B) *Harpegnathos* scaffold 2143, position 15373. C) Boxplots showing the relationship between allele-specific methylation and allele-specific expression. The number of RNA-seq reads assigned to the hypermethylated allele (hyper) or the hypomethylated allele (hypo) is plotted on the y axis. P -values are from Wilcoxon signed-rank tests.