Current Biology 22, 1755–1764, October 9, 2012 ©2012 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2012.07.042

Article

Genome-wide and Caste-Specific DNA Methylomes of the Ants *Camponotus floridanus* and *Harpegnathos saltator*

Roberto Bonasio,[1,10] Qiye Li,[3,4,10] Jinmin Lian,[3] Navdeep S. Mutti,[5,11] Lijun Jin,[3] Hongmei Zhao,[3] Pei Zhang,[3] Ping Wen,[3] Hui Xiang,[6] Yun Ding,[6] Zonghui Jin,[3] Steven S. Shen,[1,2] Zongji Wang,[3,4] Wen Wang,[6] Jun Wang,[3,7,8] Shelley L. Berger,[9,*] Jürgen Liebig,[5,*] Guojie Zhang,[3,*] and Danny Reinberg[1,*]

[1]Howard Hughes Medical Institute and Department of Biochemistry

[2]Center for Health Informatics and Bioinformatics
New York University School of Medicine, New York, NY 10016

[3]BGI-Shenzhen, Shenzhen 518083, China

[4]School of Bioscience and Bioengineering, South China University of Technology, Guangzhou 510006, China

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

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

[7]Department of Biology

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

[9]University 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 methyltransferases 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 symmetric CG dinucleotides (CpG) and non-CpG contexts and are strongly enriched at exons of active genes. Changes in exonic DNA methylation correlate with 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 interspecies 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.

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 nonreproductive female workers, sometimes divided in distinct subcastes. 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.

[10]These authors contributed equally to this work

[11]Present address: DuPont Agricultural Biotechnology, DuPont Experimental Station, Wilmington, DE 19880, USA

*Correspondence: 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.)

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 seven 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 five 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

Methylcytosines in eukaryotes are typically found in symmetric CG dinucleotides (CpG), 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 did not detect or could not validate mCHs [10, 14], we confirmed their presence in ants by conventional sequencing of 15 loci (see Figure S1 available online). Manual verification confirmed that these regions contained mCHs, not only in embryos, where extensive de novo DNA methylation is expected, but also in adults (Figures 1B and 1C). Methylated CHs were mostly in the context of CpA dinucleotides (Figures 1D and 1E), 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 (Figures S2A and S2B), whereas mCHs were methylated to a much smaller degree (Figures S2C–S2F). 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.

Methylcytosines 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 and 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

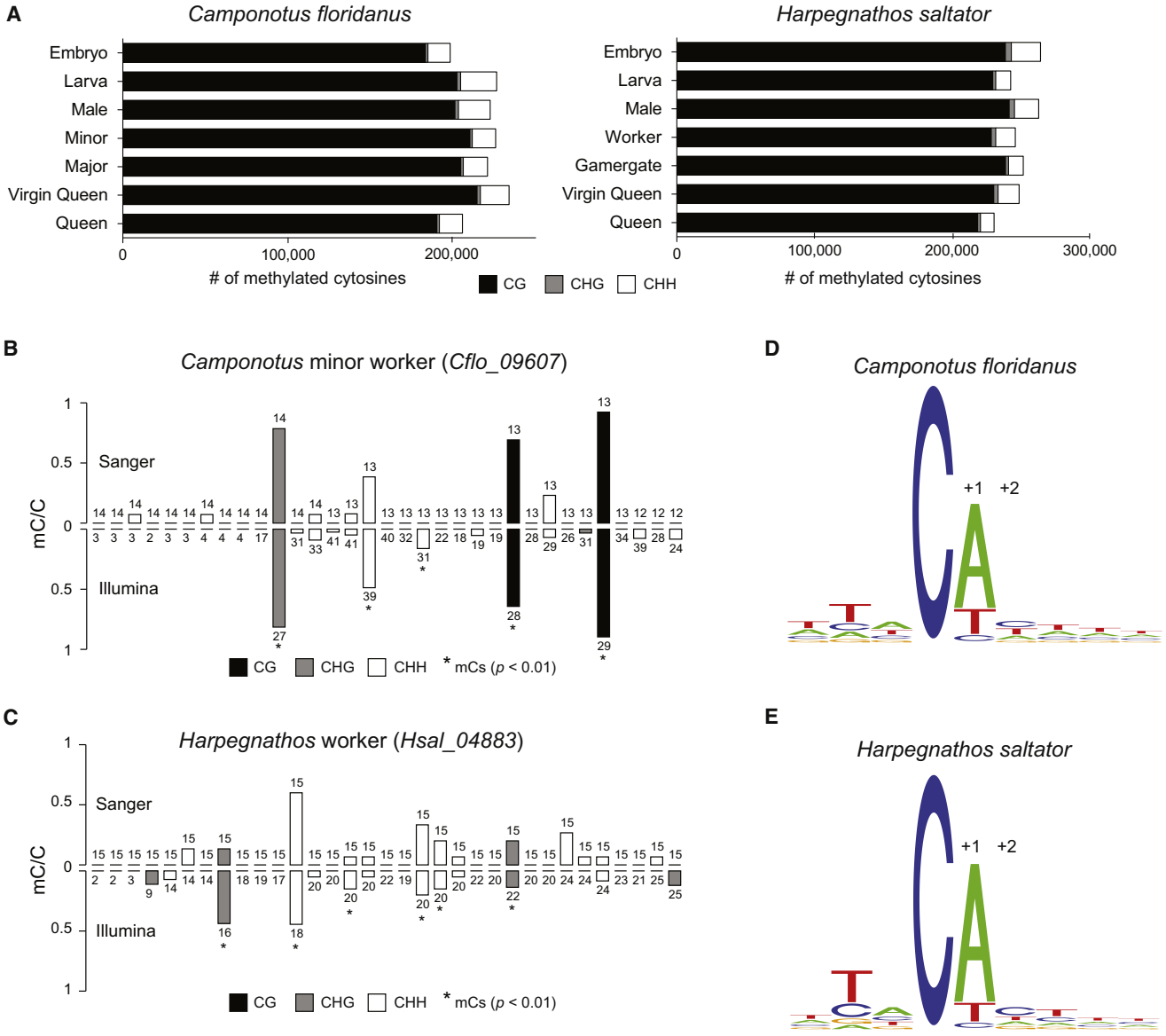

Figure 1. CpG and non-CpG Methylation in the Ant Genome

(A) Total number of mCGs (black), mCHGs (gray), and mCHHs (white) in the indicated sample from both ant species.

(B and 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 and E) Sequence context of CH methylation; the first and second base after the methylated cytosine are indicated by numbers.

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; Figure 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; Figure 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

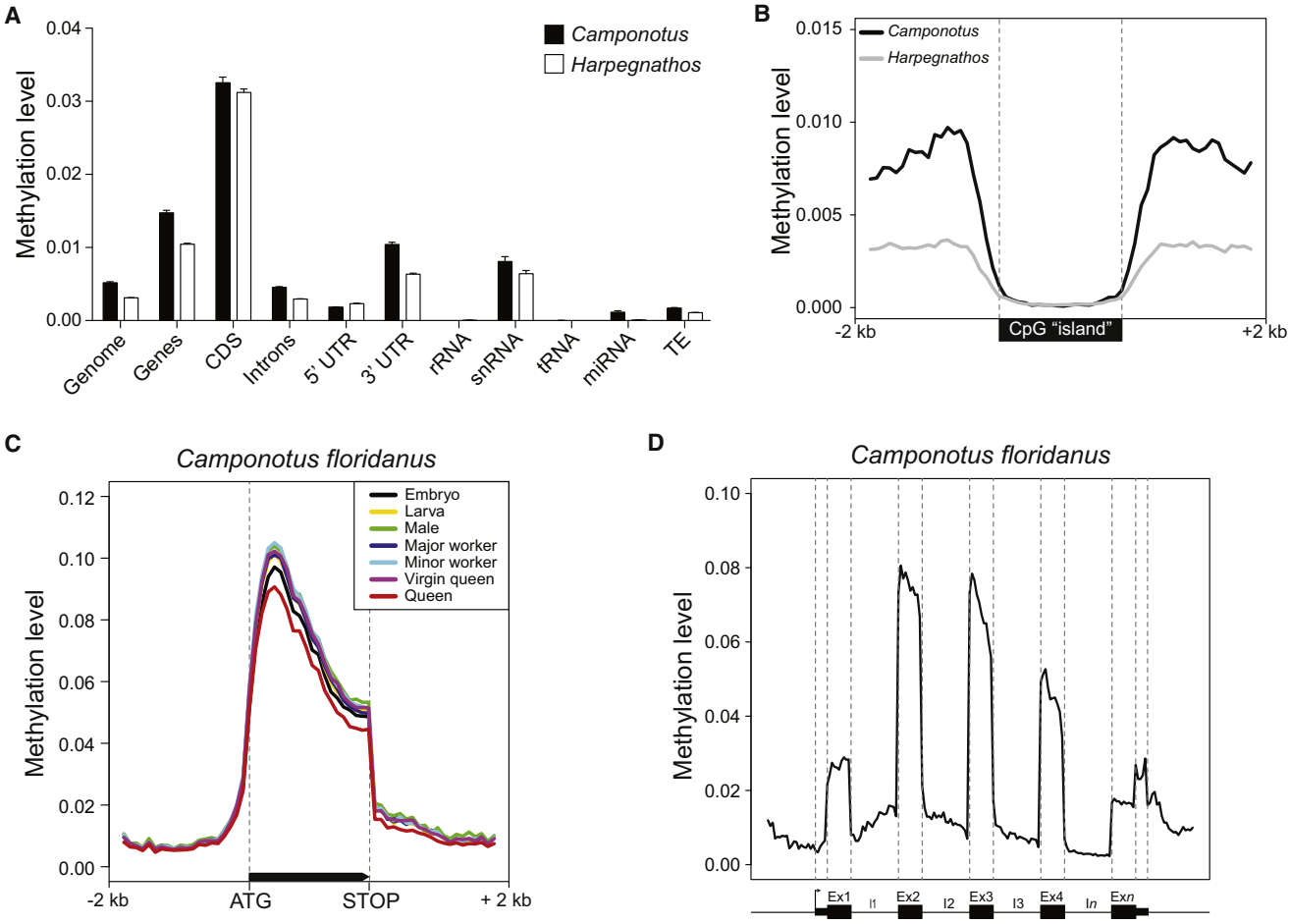

Figure 2. Distribution of mCGs on Coding and Noncoding Regions

(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 + SEM are shown (n = 7).

(B) Methylation profile of genome regions with high GC content (>55%) and high CpG O/E (>0.65) (CpG islands).

(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).

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 and/or splicing machinery (see below).

Relationship between Caste-Specific Methylomes and Transcriptomes

To compare DNA methylation and gene expression, we utilized our published RNA-seq data sets [4] and integrated them with newly generated data sets 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 (Figures 3A and 3B), as originally found in plants and other organisms [19, 24]. Methylated genes exhibited lower sample specificity [25] than unmethylated genes (Figures 3C and 3D), 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 (Figures 3E and 3F), 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 (Figures S5C and S5D), 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

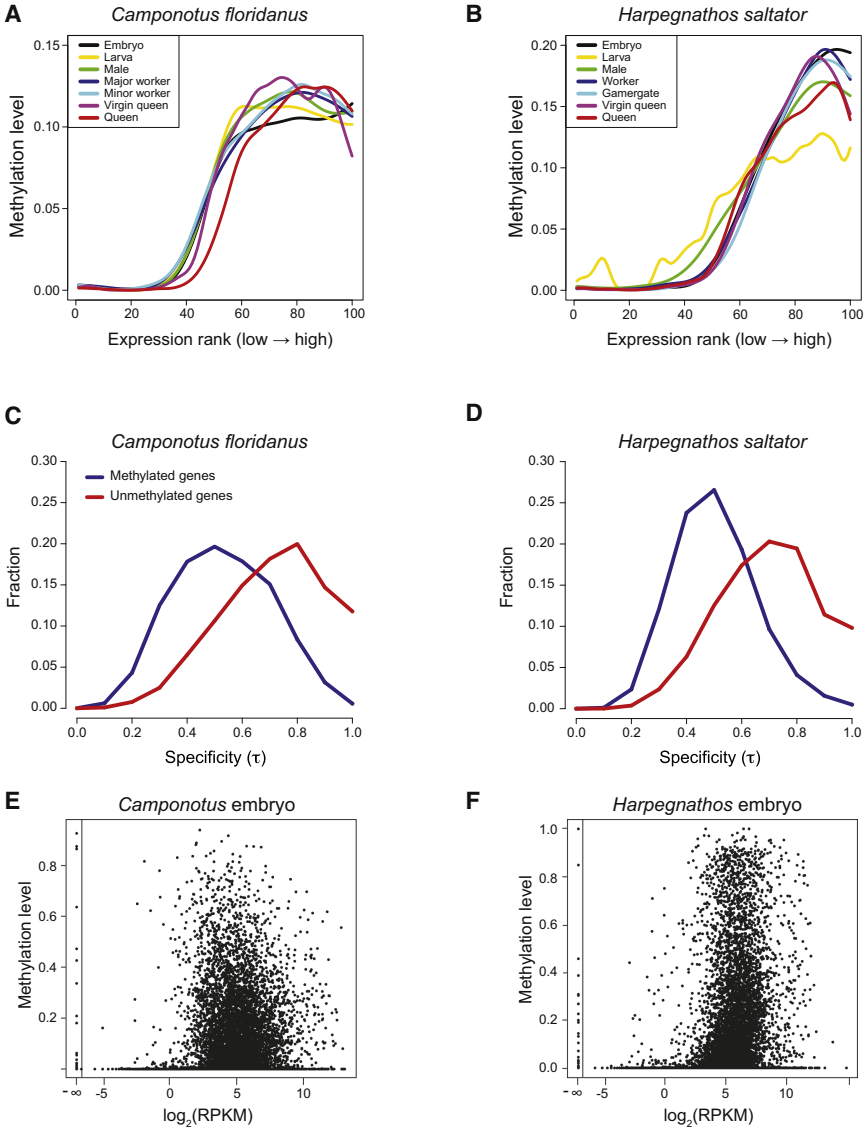

Figure 3. Gene Body Methylation and Gene Expression

(A and 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 and D) Specificity index (see text for details) for methylated and unmethylated genes.

(E and F) Scatterplot for methylation level versus RNA levels (log-transformed read per kilobase per million). Each point is an individual gene.

length (Figures 4A and 4B), 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 (Figures 4C and 4D) 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 (Figure 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 (Tables S1 and S2). The nine 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 nonreproductive 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 (Figures 5B and 5C). The fact that these changes were relatively minor suggests that the link between DNA methylation and alternative splice site

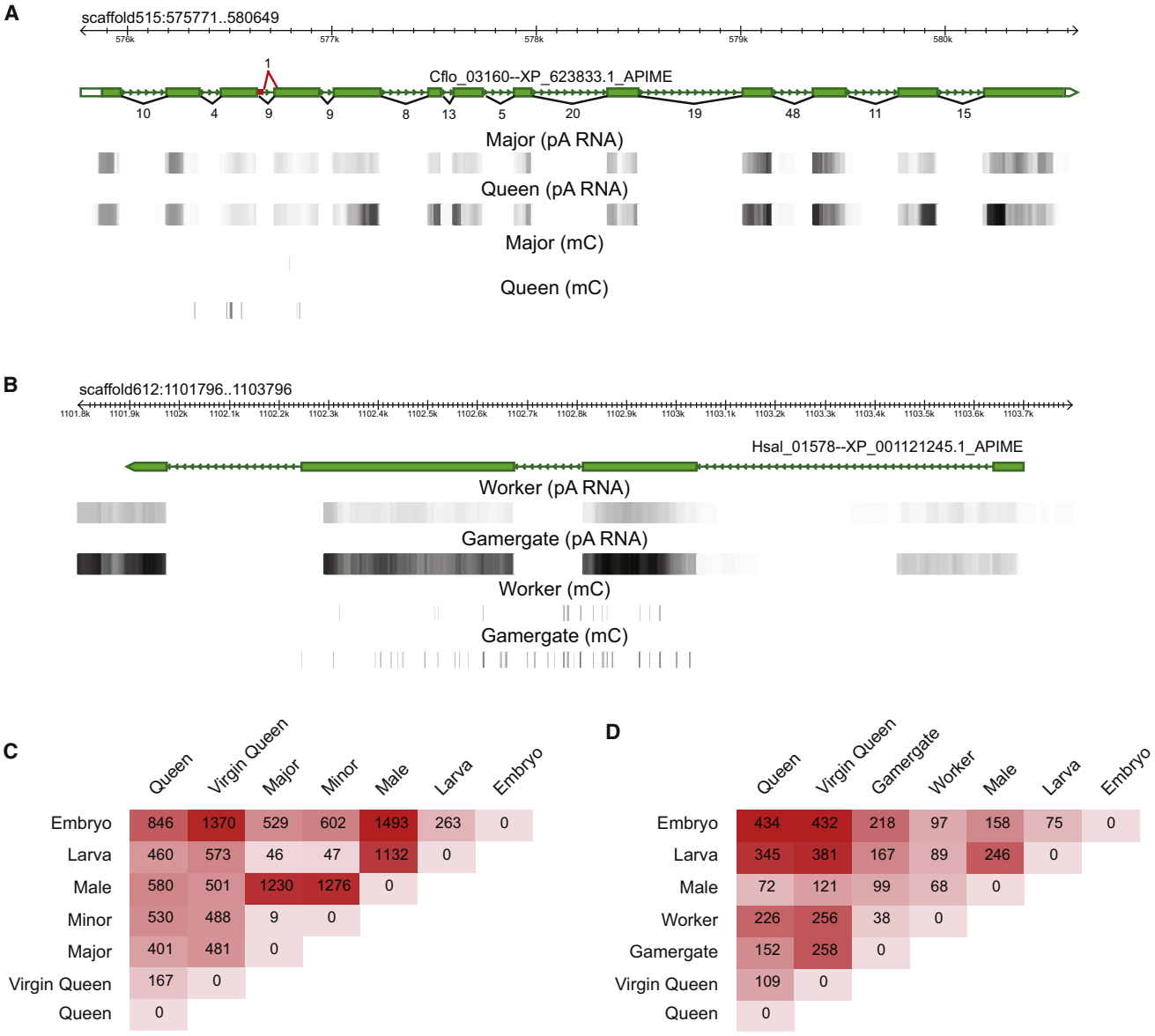

Figure 4. Differentially Methylated Genes

(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 $\geq$2-fold difference in methylation levels in each pairwise comparison in *Camponotus*.
(D) Differentially methylated genes heat map for *Harpegnathos*.

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 (Figures S6A and S6B), 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.

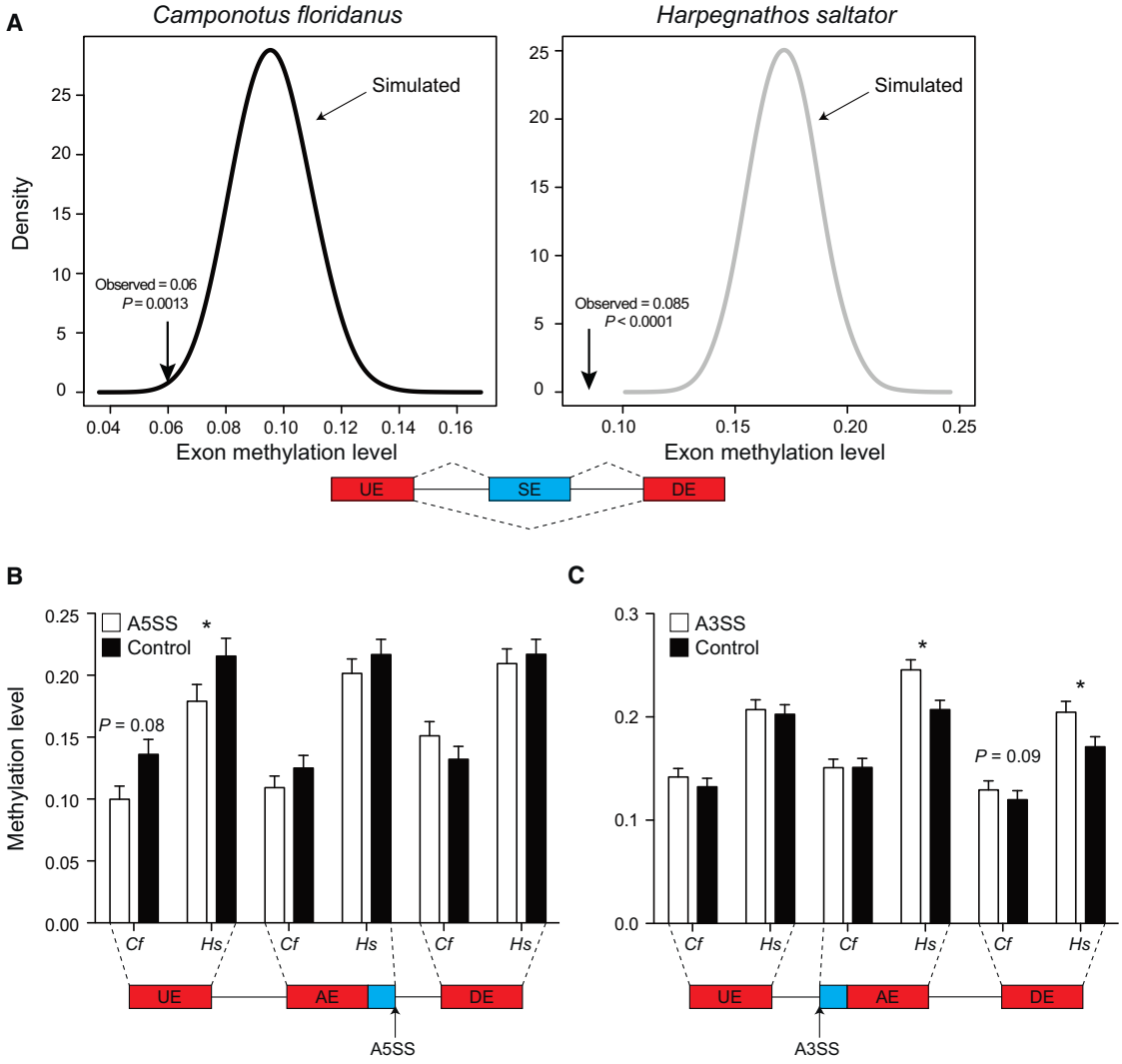

Figure 5. DNA Methylation and Alternative Splicing in Ants

(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 + SEM *p < 0.05. P values higher than 0.05 but lower than 0.1 are indicated and were determined with a nonparametric Mann-Whitney test.

(C) Same as in (B) but for alternative 3′ splice sites (A3SS).

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 nonreproductives 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 (Figures 6C and 6D), supporting a relationship between DNA methylation and gene expression in these regions.

Discussion

We obtained 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

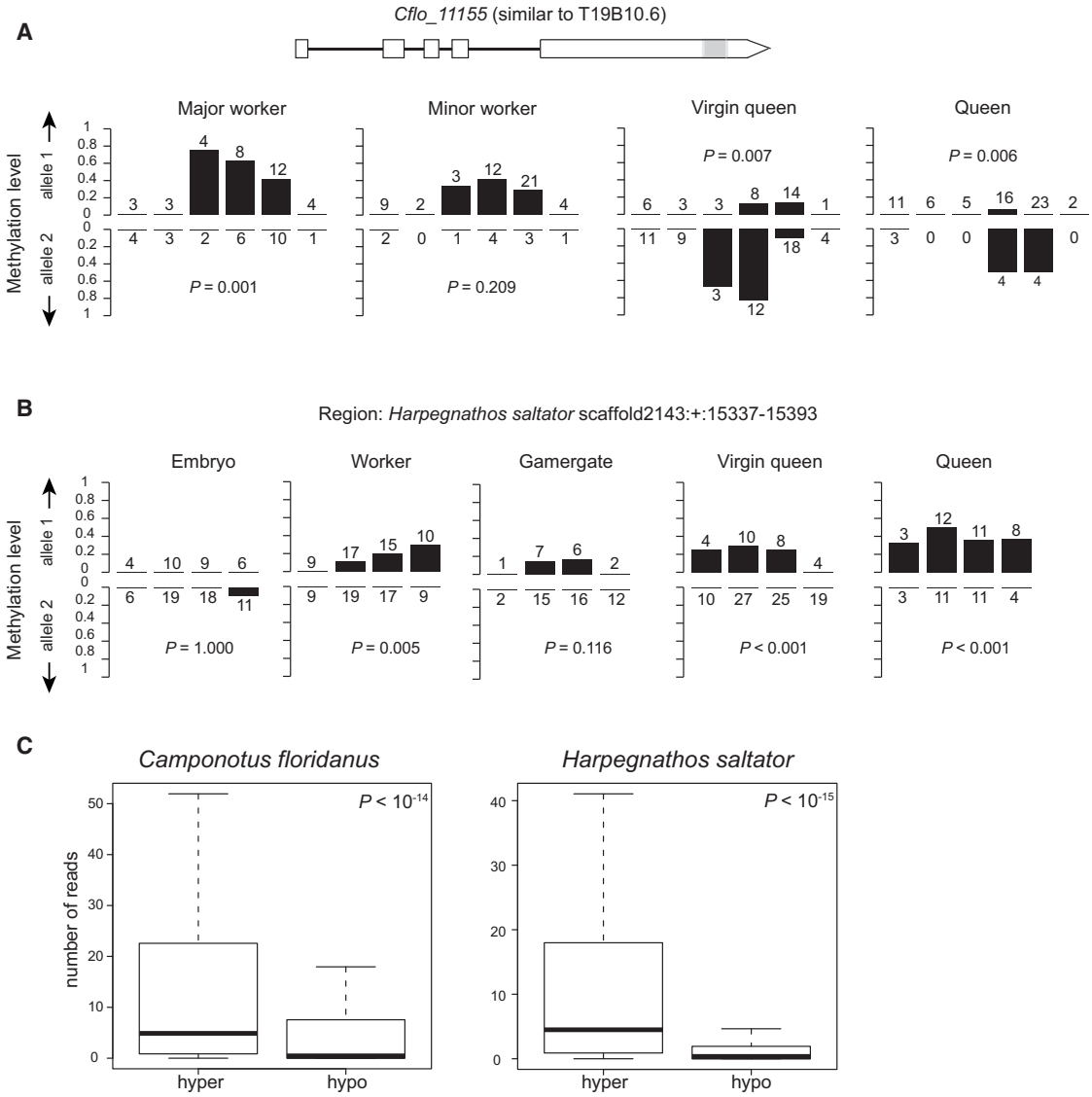

Figure 6. Allele-Specific Methylation

(A and 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 alelle-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. Boxplot whiskers indicate the most extreme data point within 1.5 times the interquartile range. (B) *Harpegnathos* scaffold 2143, position 15373.

(C) Boxplots showing the relationship between alelle-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.

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 resequencing efforts have failed to uncover any allelic bias between *Camponotus* worker castes.

By analyzing the DNA methylation maps of two developing stages and five 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, and (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 cells (ESCs) [17, 21]. It has been argued that mCH in animals is a byproduct of DNA methyltransferase 3A (DNMT3A), which has low sequence specificity, because they have been detected only in cells with strong de novo methyltransferase activity, such as ESCs [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 crosstalk 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 whether 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 orginates from imprinted DNA methylation.

Experimental Procedures

Camponotus and *Harpegnathos* colonies were housed and reared as described [4]. We isolated 20 μg of genomic DNA and 10 μg of RNA for each sample. Gut and poison glands 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 to the *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 nonmethylated 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.

Alternative splicing events were detected using junction reads identified by TOPHAT [48]. To ensure the accuracy and reliability of junction reads, we required at least 8 bp with no mismatch 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 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, we considered only windows with at least six (three per strand) informative CpGs ($\geq 3\times$ coverage) in both replicates of the two comparable samples.

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 hypomethylated alleles using RNA-seq.

Additional details are described in the Supplemental Experimental Procedures.

Accession Numbers

Sequencing data generated for this study have been deposited in the NCBI GEO database as GSE31577. This comprises the subseries GSE31344 (small RNA-seq), GSE31346 (RNA-seq), and GSE31576 (BS-seq). Previously published sequencing data analyzed in this publication are available in the GSE22680 superseries.

Supplemental Information

Supplemental Information includes six figures, six tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2012.07.042.

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; and H. Shao for help with the mC detection pipeline. R.B. was supported by a Helen Hay Whitney Foundation postdoctoral fellowship. This work was funded by a Howard Hughes Medical Institute Collaborative Innovation Award (#2009005) to S.L.B., D.R., and J.L.

Received: March 23, 2012
Revised: June 2, 2012
Accepted: July 18, 2012
Published online: August 9, 2012

References

1. Hölldobler, B., and Wilson, E.O. (1990). The Ants (Cambridge, MA: Harvard University Press).

2. Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. Science *330*, 612–616.
3. Bonasio, R. (2012). Emerging topics in epigenetics: ants, brains, and noncoding RNAs. Ann. N. Y. Acad. Sci. *1260*, 14–23.
4. Bonasio, R., Zhang, G., Ye, C., Mutti, N.S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., et al. (2010). Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science *329*, 1068–1071.
5. Peeters, C., Liebig, J., and Hölldobler, B. (2000). Sexual reproduction by both queens and workers in the ponerine ant *Harpegnathos saltator*. Insectes Soc. *47*, 325–332.
6. Liebig, J., Peeters, C., and Hölldobler, B. (1999). Worker policing limits the number of reproductives in a ponerine ant. Proc. Biol. Sci. *266*, 1865–1870.
7. Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220.
8. Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.
9. Walsh, T.K., Brisson, J.A., Robertson, H.M., Gordon, K., Jaubert-Possamai, S., Tagu, D., and Edwards, O.R. (2010). A functional DNA methylation system in the pea aphid, *Acyrthosiphon pisum*. Insect Mol. Biol. *19* (*Suppl 2*), 215–228.
10. Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., Zhang, H., Zhang, G., Li, D., Dong, Y., et al. (2010). Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. Nat. Biotechnol. *28*, 516–520.
11. Wang, Y., Jorda, M., Jones, P.L., Maleszka, R., Ling, X., Robertson, H.M., Mizzen, C.A., Peinado, M.A., and Robinson, G.E. (2006). Functional CpG methylation system in a social insect. Science *314*, 645–647.
12. Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. Science *319*, 1827–1830.
13. Lockett, G.A., Helliwell, P., and Maleszka, R. (2010). Involvement of DNA methylation in memory processing in the honey bee. Neuroreport *21*, 812–816.
14. Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol. *8*, e1000506.
15. Lyko, F., Ramsahoye, B.H., and Jaenisch, R. (2000). DNA methylation in *Drosophila melanogaster*. Nature *408*, 538–540.
16. Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.
17. Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. USA *97*, 5237–5242.
18. Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.
19. Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science *328*, 916–919.
20. Suzuki, M.M., Kerr, A.R.W., De Sousa, D., and Bird, A. (2007). CpG methylation is targeted to transcription units in an invertebrate genome. Genome Res. *17*, 625–631.
21. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.
22. Takai, D., and Jones, P.A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc. Natl. Acad. Sci. USA *99*, 3740–3745.
23. Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. Proc. Natl. Acad. Sci. USA *107*, 8689–8694.
24. Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., and Ecker, J.R. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in *arabidopsis*. Cell *126*, 1189–1201.
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. (2005). Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics *21*, 650–659.
26. Foret, S., Kucharski, R., Pittelkow, Y., Lockett, G.A., and Maleszka, R. (2009). Epigenetic regulation of the honey bee transcriptome: unravelling the nature of methylated genes. BMC Genomics 2009 *10*, 472.
27. Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S.E., Robinson, G.E., and Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. Proc. Natl. Acad. Sci. USA *109*, 4968–4973.
28. Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature *479*, 74–79.
29. Tufail, M., and Takeda, M. (2009). Insect vitellogenin/lipophorin receptors: molecular structures, role in oogenesis, and regulatory mechanisms. J. Insect Physiol. *55*, 87–103.
30. Koshikawa, S., Cornette, R., Matsumoto, T., and Miura, T. (2010). The homolog of Ciboulot in the termite (*Hodotermopsis sjostedti*): a multimeric beta-thymosin involved in soldier-specific morphogenesis. BMC Dev. Biol. *10*, 63.
31. Payer, B., and Lee, J.T. (2008). X chromosome dosage compensation: how mammals keep the balance. Annu. Rev. Genet. *42*, 733–772.
32. Barlow, D.P., and Bartolomei, M.S. (2007). Genomic Imprinting in Mammals. In Epigenetics, C.D. Allis, T. Jenuwein, and D. Reinberg, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
33. Li, Y., Zhu, J., Tian, G., Li, N., Li, Q., Ye, M., Zheng, H., Yu, J., Wu, H., Sun, J., et al. (2010). The DNA methylome of human peripheral blood mononuclear cells. PLoS Biol. *8*, e1000533.
34. Schwander, T., Lo, N., Beekman, M., Oldroyd, B.P., and Keller, L. (2010). Nature versus nurture in social insect caste differentiation. Trends Ecol. Evol. *25*, 275–282.
35. Lee, T.F., Zhai, J., and Meyers, B.C. (2010). Conservation and divergence in eukaryotic DNA methylation. Proc. Natl. Acad. Sci. USA *107*, 9027–9028.
36. Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61–69.
37. Beetsma, J. (1979). The process of queen-worker differentiation in the honeybee. Bee World *60*, 24–39.
38. Penick, C.A., and Liebig, J. (2012). Regulation of queen development through worker aggression in a predatory ant. Behav. Ecol. *23*, 141–152.
39. Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., et al. (2006). The *Polycomb* group protein EZH2 directly controls DNA methylation. Nature *439*, 871–874.
40. Schmitz, K.M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev. *24*, 2264–2269.
41. Queller, D.C., and Strassmann, J.E. (2002). The many selves of social insects. Science *296*, 311–313.
42. Haig, D. (1992). Intragenomic conflict and the evolution of eusociality. J. Theor. Biol. *156*, 401–403.
43. Queller, D.C. (2003). Theory of genomic imprinting conflict in social insects. BMC Evol. Biol. *3*, 15.
44. Dobata, S., and Tsuji, K. (2012). Intragenomic conflict over queen determination favours genomic imprinting in eusocial Hymenoptera. Proc. Biol. Sci. *279*, 2553–2560.
45. Guzman-Novoa, E., Hunt, G.J., Page, R.E., Jr., Uribe-Rubio, J.L., Prieto-Merlos, D., and Becerra-Guzman, F. (2005). Paternal effects on the defensive behavior of honeybees. J. Hered. *96*, 376–380.
46. Hayatsu, H., Tsuji, K., and Negishi, K. (2006). 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) *50*, 69–70.
47. Li, R., Yu, C., Li, Y., Lam, T.W., Yiu, S.M., Kristiansen, K., and Wang, J. (2009). SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics *25*, 1966–1967.
48. Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.
49. Li, R., Li, Y., Fang, X., Yang, H., Wang, J., Kristiansen, K., and Wang, J. (2009). SNP detection for massively parallel whole-genome resequencing. Genome Res. *19*, 1124–1132.