Reversible switching between epigenetic states in honeybee behavioral subcastes

Brian R Herb[1,2,8], Florian Wolschin[3,4,8], Kasper D Hansen[1,5], Martin J Aryee[1,6], Ben Langmead[5], Rafael Irizarry[5], Gro V Amdam[3,4] & Andrew P Feinberg[1,2,5,7]

In honeybee societies, distinct caste phenotypes are created from the same genotype, suggesting a role for epigenetics in deriving these behaviorally different phenotypes. We found no differences in DNA methylation between irreversible worker and queen castes, but substantial differences between nurses and forager subcastes. Reverting foragers back to nurses reestablished methylation levels for a majority of genes and provides, to the best of our knowledge, the first evidence in any organism of reversible epigenetic changes associated with behavior.

Epigenetic changes are thought to underlie lineage-specific differentiation, as the pattern of gene expression is stably changed, but the DNA sequence remains the same. Recently, the epigenome of a specific differentiation pathway was mapped, defining hundreds of differentially methylated regions (DMRs) that define lineage commitment in mouse hematopoietic progenitors. DNA methylation appears to be critical in that system for lineage specificity, as lymphoid cells show greater global DNA methylation than myeloid cells[1]. However, the roles of the epigenome in global changes in organismal remodeling or in behavior have not been defined[2].

The honeybee *Apis mellifera* is an ideal model organism for such studies[3], as it organizes social structures from distinct individual forms that can emerge from one genome. A female embryo may become a queen by receiving a diet of royal jelly and commit her life to egg-laying (germline) or become a sterile helper 'worker' (soma)[4]. Workers follow a rich behavioral program of nursing and later undergo a transition to foraging that involves extensive gene expression changes in the brain[5]. In contrast with queens, worker behavior is remarkably flexible: age-matched workers can nurse or forage, and foragers may revert to nursing tasks[6].

To investigate the potential role of DNA methylation in defining honeybee caste phenotypes, we compared the methylomes of sister queens versus workers and sister nurses versus foragers by whole genome bisulfite sequencing (WGBS) and comprehensive high-throughput array-based relative methylation (CHARM) analysis[1]. CHARM covers 85% and WGBS covers 92% of the CpGs in the 270-Mb genome, both revealing sparse methylation throughout the bee genome (**Supplementary Fig. 1**).

We first compared five biological replicates of queens and workers, both collected within 4 h of adult emergence from the pupal stage (**Fig. 1a**). Brain was selected because of its influence on behavior and, unlike ovary, is similar in size between queens and workers. CHARM analysis found no significant DMRs (at a false discovery rate (FDR) of 5%, permutation test) between queens and workers. WGBS of the same samples found no differences using single CpG t tests corrected for multiple testing. In addition, we tested the top-ranked differences by CHARM, albeit statistically insignificant (FDR cutoff of 5%), using bisulfite pyrosequencing, an independent measure of DNA methylation at the single-base level, and found no caste-specific differences (**Supplementary Fig. 2a–c**).

Given these negative results, we then compared subcastes of workers. Initially, most workers are nurses that care for the queen and larvae inside the hive. About 2–3 weeks later, the majority switches to foraging and collect pollen, nectar and water outside[5]. Using CHARM, we identified 155 DMRs that distinguished nurses from age-matched foragers (**Figs. 1b** and **2a,b**, **Table 1**, **Supplementary Table 1** and **Supplementary Fig. 3a–e**). Approximately 70% of DMRs overlapped exons (**Table 1** and **Supplementary Table 1**), similar to previous findings[7,8]. The genes associated with the 155 nurse-to-forager DMRs appeared to be enriched for gene regulation and development through transcriptional control and chromatin remodeling. Many histone modification writers, including LOC412350 (a histone deacetylase similar to Hdac3), JIL-1 (a histone phosphotransferase) and LOC411070 (a histone H3 methyltransferase[9]), increased in methylation during the nurse to forager transition. In addition, DEAD-box helicase genes *Iswi* and *spn-E* have chromatin remodeling capacity and are involved in morphogenesis[10]. *Iswi* in particular is involved in dendrite morphogenesis[11] and may contribute to noted changes to the nurse brain before foraging[5].

To determine whether the DMRs that we observed during the nurse to forager transition are linked to phenotype, and not simply the result of the transition, we induced the reversion of foragers back to nurses using a strategy of hive trickery. To initiate reversion, we set up foragers to return to a hive in which only queen and larvae are present (**Fig. 1b**). The foragers will then segregate into reverted nurses that pick up caregiver tasks and continuing foragers that do not change behavior[6]. Reversion separates changes caused by nervous system development, maturation and foraging experience that are shared between reverted nurses and foragers, but not nurses, from changes robustly linked to current behavior that are shared between reverted nurses and nurses, but not foragers.

[1]Center for Epigenetics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [2]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [3]Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Aas, Norway. [4]School of Life Sciences, Arizona State University, Tempe, Arizona, USA. [5]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. [6]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [7]Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [8]These authors contributed equally to this work. Correspondence should be addressed to G.V.A. (gro.amdam@asu.edu) or A.P.F. (afeinberg@jhu.edu).
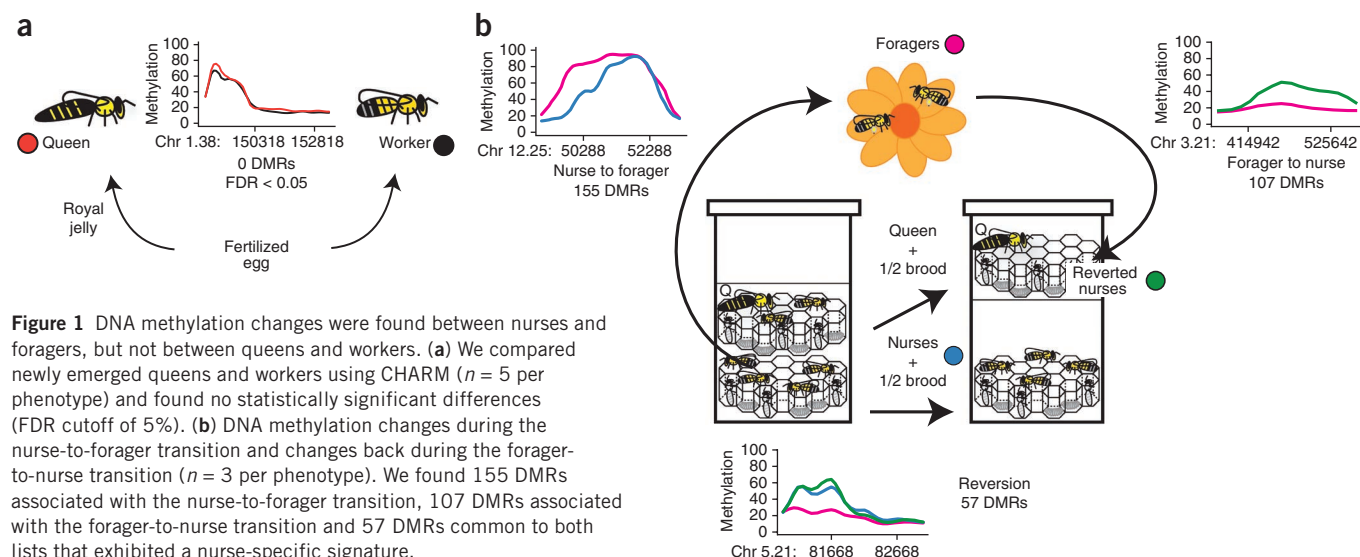
Figure 1 DNA methylation changes were found between nurses and foragers, but not between queens and workers. (**a**) We compared newly emerged queens and workers using CHARM (*n* = 5 per phenotype) and found no statistically significant differences (FDR cutoff of 5%). (**b**) DNA methylation changes during the nurse-to-forager transition and changes back during the forager-to-nurse transition (*n* = 3 per phenotype). We found 155 DMRs associated with the nurse-to-forager transition, 107 DMRs associated with the forager-to-nurse transition and 57 DMRs common to both lists that exhibited a nurse-specific signature.

With CHARM, we found 107 DMRs for the forager-to–reverted nurse transition. The genes associated with these CHARM DMRs appeared to be enriched in transcription factors and DEAD-box helicases, as seen in the nurse-to-forager CHARM DMRs (**Fig. 2a,b**, **Supplementary Table 2** and **Supplementary Fig. 4**). Of these 107 CHARM DMRs, 57 overlapped with CHARM DMRs associated with the nurse-to-forager transition, a markedly close concordance ($P < 2.2 \times 10^{-16}$ by Fisher's test, $P < 10^{-3}$ based on 1,000 permutations; **Supplementary Fig. 5**). This subset of epigenetically reversible genes showed enrichment for development, ATP binding and nuclear pore formation (**Supplementary Table 2**). These genes include the ortholog to kismet, LOC726524, which regulates developmental genes such as hedgehog and affects learning and axon migration in

Drosophila[12,13], and might explain observed differences in learning[14] between nurses and foragers. In addition, DEAD-box helicase genes LOC725306 (ref. 15) and LOC726524 both have roles in transcription, whereas LOC411989 is involved in translation[16].

To independently validate this result, we replicated the reversion experiment and created six new pools of six brains for both foragers and reverted nurses. We performed WGBS on these 12 samples and found that 45 of 57 reversion DMRs showed the same direction of change in methylation between CHARM and WGBS (**Supplementary Fig. 6**). This overlap of DMRs between replicated experiments was highly significant ($P = 3.3 \times 10^{-6}$). Furthermore, the 45 WGBS-correlated genes showed enrichment for ATP binding and nuclear pore formation (**Supplementary Table 3**), consistent with our analysis of the 57 CHARM
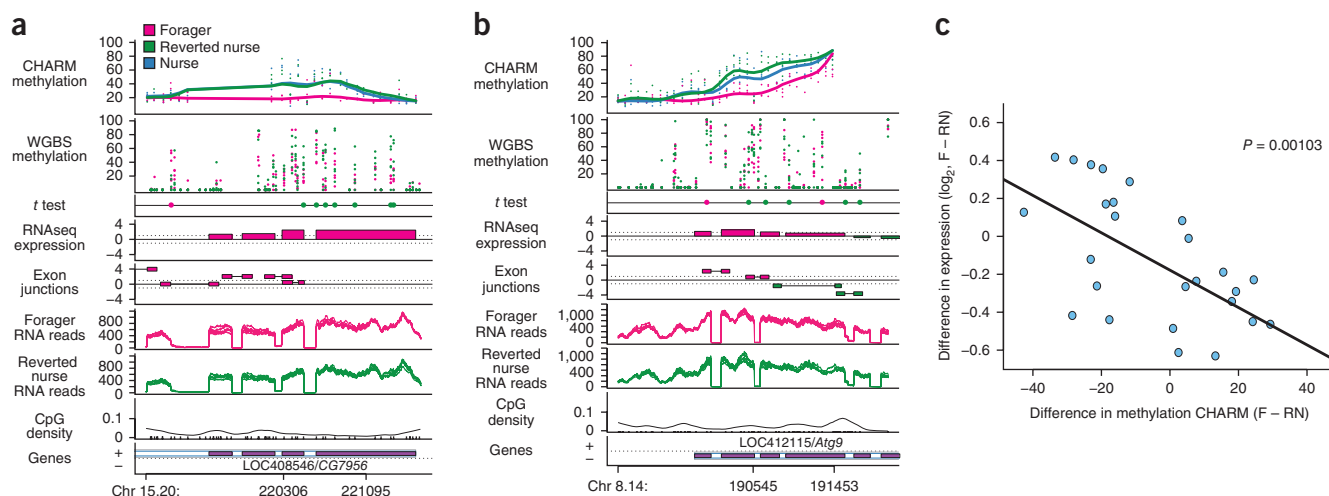

Figure 2 DNA methylation distinguishes nurses, foragers and reverted nurses. (**a,b**) Two examples of CHARM DMRs. Top, percent methylation for both CHARM and WGBS data sets, with points representing individual samples and the smoothed lines representing the average for the phenotype. The *t* test panel displays the top 1% differentially methylated CpGs by *t* test. The color of the point indicates which phenotype had greater methylation at that CpG (*n* = 6 per phenotype). The RNAseq expression panel is a *t* statistic based on the number or reads detected in the annotated exons, with the color indicating the higher expressed phenotype. The exon junctions panel is a *t* statistic based on the number of reads detected spanning the exon junctions, as predicted by the TopHat program, with the color indicating the higher expressed phenotype. Switching between higher expressed nurse and forager exon junctions is indicative of alternative splicing events. The RNA reads panels indicate the number of reads per phenotype as compiled by TopHat program (*n* = 6 per phenotype). The bottom two panels show the CpG density and the relative position of the gene. (**c**) Plot of relative gene expression comparing foragers (F) to reverted nurses (RN). We tested 26 genes associated with DMRs for expression differences by real-time PCR (*n* = 12 per phenotype). The plot depicts the difference in average log$_2$ expression versus average difference in methylation as determined by CHARM. Correlation analysis results in a *P* value of 0.001.

Table 1 Summary of DMRs

DMR type	Number of DMRs	Percentage of DMRs overlapping exons	GO category examples	Example genes
Nurse-to-forager transition	155	72.2%	Helicase activity, chromatin remodeling, neuron development	*alpha-Cat, Stat92E, Dhc64C, alpha-Spec, hts, spn-E*
Forager-to-nurse transition	107	72.9%	Nucleoside binding, helicase activity, cytoskeleton organization	*Hel89B, Hsc70Cb, CG7177, Upf1, CG2017, eIF-4a, kis, Dhc64C*
Overlap	57	73.7%	Nuclear import, cell differentiation, ATP binding	*Ranbp21, Fs(2)Ket, Mtor eIF-4a, kis, poe, bur, BicD*

reversion DMRs. These results provide evidence for a nurse-specific methylome that needs to be reestablished during the reversion.

To determine the relevance of these reversible DMRs, we performed transcriptome sequencing (RNAseq) on six pools each of foragers and reverted nurses. We then used the TopHat program to analyze the RNAseq data to predict the location of annotated and unannotated exons and to determine the prevalence of exon skipping. This analysis revealed that 22 of the 45 WGBS-correlated reversion DMRs colocalized with alternative splicing events (**Fig. 2b** and **Supplementary Fig. 7a–f**). These data suggest a high incidence of alternative splicing events in DMRs and strengthen the potential role of DNA methylation in regulating alternative splicing[8]. We also found a negative correlation between gene expression and levels of DNA methylation between foragers and reverted nurses for 26 genes by real-time PCR ($P = 0.00103$; **Fig. 2c** and **Supplementary Fig. 8a–f**).

Our data suggest a strong link between reversible DNA methylation and nurse-forager transition and reversion, but no relationship to queen-worker segregation. These data stand in contrast with a study comparing 2.5-week-old mated queens and 8-d-old foraging-capable workers[8], which is likely explained by the difference in timing of the data, that is, newly emerged queens and workers. Although DNA methylation may be involved in distinguishing queens from workers during development[3], our data clearly indicate that the queen and worker brain methylomes are the same at the time of emergence, despite differences in body morphology.

In summary, we found substantial DNA methylation changes that accompany phenotype switching in honeybee subcastes. Genes associated with these DMRs can potentially influence global gene expression patterns by altering chromatin structure or regulating transcriptional machinery. Profound phenotype shifts between nurses and foragers may be orchestrated by a subset of genes, which are themselves regulated by DNA methylation. Key regulatory genes may either be differentially expressed or differentially spliced, which we correlate with changes in DNA methylation. For example, the eIF-4a homolog LOC411989, which is critical for translation initiation[17], exhibits alternative splicing in an exon that codes for RNA binding (**Supplementary Fig. 7c**). Different isoforms of eIF-4a may bind to RNA with greater affinity, thereby globally affecting the rate or regulation of translation. As this differentially spliced exon is in a DMR, methylation might be used to remember which isoform to express in nurses or foragers.

We found that DNA methylation was able to revert, concomitant with experimental reversion of foragers back to nurses, which we demonstrated in replicated experiments. This suggests a subcaste-specific methylation signature that assists in forming subcaste phenotypes. Although studies in rodents found methylation changes associated with learning, these changes disappear over several hours and do not establish a stable phenotype[18,19]. Similarly, nurturing can induce long lasting methylation marks in rodents[20]. Our results are, to the best of

our knowledge, the first to show reversible DNA methylation corresponding to a reversible behavioral phenotype in any species.

METHODS

Methods and any associated references are available in the online version of the paper.

AUTHOR CONTRIBUTIONS
B.R.H. performed genome-scale, gene-specific DNA methylation analysis and performed gene expression analysis. F.W. raised bees and manipulated hives for reversion experiment, and collected bees and dissected brains. R.I., M.J.A., K.D.H., B.L. and B.R.H. performed statistical analysis. B.R.H. and M.J.A. generated microarray data sets. B.R.H., B.L. and K.D.H. generated WGBS and RNAseq data sets. A.P.F. and G.V.A. conceived, designed and oversaw the experiments. A.P.F., B.R.H. and G.V.A. wrote the paper with the assistance of K.D.H., M.J.A., B.L. and F.W.

1. Ji, H. *et al. Nature* **467**, 338–342 (2010).
2. Wolf, C. & Linden, D.E. *Genes Brain Behav.* **11**, 3–28 (2012).
3. Wang, Y. *et al. Science* **314**, 645–647 (2006).
4. Winston, M.L. *The Biology of the Honey Bee* (Harvard University Press, 1987).
5. Whitfield, C.W. *Science* **302**, 296–299 (2003).
6. Robinson, G.E., Page, R.E., Strambi, C. & Strambi, A. *Ethology* **90**, 336–348 (1992).
7. Elango, N., Hunt, B.G., Goodisman, M.A. & Yi, S.V. *Proc. Natl. Acad. Sci. USA* **106**, 11206–11211 (2009).
8. Lyko, F. *et al. PLoS Biol.* **8**, e1000506 (2010).
9. Suganuma, T. *et al. Nat. Struct. Mol. Biol.* **15**, 364–372 (2008).
10. Tsukiyama, T., Daniel, C., Tamkun, J. & Wu, C. *Cell* **83**, 1021–1026 (1995).
11. Barak, O. *et al. EMBO J.* **22**, 6089–6100 (2003).
12. Terriente-Félix, A., Molnar, C., Gómez-Skarmeta, J.L. & de Celis, J.F. *Dev. Biol.* **350**, 382–392 (2011).
13. Melicharek, D.J., Ramirez, L.C., Singh, S., Rhea, T. & Marenda, D.R. *Hum. Mol. Genet.* **19**, 4253–4264 (2010).
14. Amdam, G.V. *Aging Cell* **10**, 18–27 (2011).
15. Goldman-Levi, R., Miller, C., Bogoch, J. & Zak, N.B. *Nucleic Acids Res.* **24**, 3121–3128 (1996).
16. Linder, P. *Nucleic Acids Res.* **34**, 4168–4180 (2006).
17. Parsyan, A. *et al. Nat. Rev. Mol. Cell Biol.* **12**, 235–245 (2011).
18. Guo, J.U. *et al. Nat. Neurosci.* **14**, 1345–1351 (2011).
19. Miller, C.A. & Sweatt, J.D. *Neuron* **53**, 857–869 (2007).
20. Weaver, I.C. *et al. Nat. Neurosci.* **7**, 847–854 (2004).

ONLINE METHODS

Bee preparation. For each replicate, two colonies were initially prepared for nurse and forager rearing. They each consisted of 6,000–7,000 newly emerged (0–24 h old) workers born to five sister queens of our standard research stock and a wild-type queen of Californian commercial origin. All individuals were paint-marked (Testors) on the thorax; after a solid foraging pattern was established, foragers were paint-marked on the abdomen for the purpose of tracking their life history. The behavioral reversion was carried out essentially as described before[21]. This reversion resulted in two forager-derived colonies and two nurse-derived colonies. Young brood of Californian commercial origin was provided to the colonies as incentive for reversal from foraging to nursing behavior. For the replicate used in the CHARM and bisulfite pyrosequencing assays, bees of three groups were collected 12–14 d after the reversion: continuous foragers (forager bees that had not reverted to nursing behavior), reverted nurse bees (forager bees that exhibited nurse-like behavior (head in brood cells, sluggish response to a flight challenge) in the brood area) and continuous nurses (bees that had never been observed foraging and exhibited at least one nurse-like behavioral trait at the time of collection). For the replicate used in the WGBS and RNAseq assays, continuous foragers and reverted nurse bees were collected for the validation of the CHARM results.

Queens and workers were derived from eggs produced by a single drone–inseminated honeybee queen that belonged to a standard research stock with restricted genetic background[22] and were left to develop into 2-d-old larvae in the hive. Subsequently, larvae at this crucial point in the divergent development of queen and worker traits were either allowed to develop into worker pupae or, for queen pupae development, they were manually grafted into queen cells and raised as previously described[23]. On the day before emergence (the day of sample collection), queen and worker pupae were transferred into an incubator at 33 °C, 65–70% relative humidity.

For all bees, individuals from both experiments were collected directly into 2 ml of 80% ice-cold ethanol and stored at 4 °C until the central brains were dissected (<48 h). Following dissections, central brains were immediately transferred into liquid N_2 and stored at −80 °C until further use.

CHARM DNA methylation analysis. Genomic DNA was isolated from brains using the Masterpure kit from Epicentre. Brains from the nurse/forager study were pooled in groups of seven for each biological replicate, three replicates per phenotype. Brains from the queen/worker study were pooled in groups of eight for each biological replicate, five replicates per phenotype. Genome-wide methylation was assessed by CHARM[24,25] and performed as previously described[1]. A custom-designed Nimblegen 2.1 million feature array was designed for the honeybee genome, which covers approximately 200 Mb of non-repeat genomic sequence and includes ~8.7 million CpG sites covering ~85% of the ~10.2 million CpG sites in the genome.

DMRs for the nurse/forager study were determined by a cutoff of at least 10% methylation difference[24], and DMRs with an average methylation close to the baseline of 20% were eliminated. DMRs from pairwise comparisons were combined to determine the relative methylation of all three phenotypes. Methylation differences between continuous nurses and both reverted nurses and continuous foragers were determined, and these differences were plotted. Clustering analysis identified three classes of DMRs (**Supplementary Fig. 9**). GO analysis[26] of each class was performed by first determining the closest *Apis mellifera* gene to each DMR. The orthologous gene in *Drosphila melanogaster* was found for each *Apis mellifera* gene, and this new gene list was used for GO analysis using the web tools available on the DAVID bioinformatics database at http://david.abcc.ncifcrf.gov/. DMRs for the queen/worker study were determined by calculating a FDR score for each potential DMR.

Bisulfite pyrosequencing. Approximately 400 ng of pooled genomic DNA that was also used for the CHARM analysis was bisulfite converted using the Zymo DNA-Methylation Gold kit. We used nested PCR to amplify regions of interest in DMRs and quantified the level of methylation using the Biotage PSQ HS96 pyrosequencer. The percent methylation for every CpG in our target region was calculated using the Q-CpG methylation software (Biotage). Control DNA was prepared using the Repli-G kit (Qiagen) from genomic DNA. Repli-G amplified DNA served as a 0% methylated control. The 100% methylated control was created by treating the Repli-G amplified DNA with SssI

methyltransferase (NEB), which methylates every CpG site. The 25%, 50% and 75% methylated controls were created by mixing 0% and 100% controls. All controls were bisulfite treated with same Zymo kit as test samples. Primer sequences are listed in **Supplementary Table 4**.

Real-time quantitative RT-PCR. RNA was extracted from single brains by first lysing cells in Chaos buffer (4.5 M guanidinium thiocyanate, 2% *N*-lauroylsarcosine (wt/vol), 50 mM EDTA, 25 mM Tris-HCl, 0.1 M β-mercaptoethanol) followed by phenol, chloroform, then purified with Qiagen RNeasy columns. cDNA was synthesized using Quantitect Reverse Transcription Kit (Qiagen) and 1 ng of cDNA was used for each real-time PCR reaction. Fast Sybr green (Applied Biosystems) was used for real-time PCR reaction and quantified by 7900HT (Applied Biosystems). Primer sequences are listed in **Supplementary Table 5**.

WGBS. Libraries of queen/worker and reverted nurse/forager samples were created using TruSeq DNA library preparation kits (Illumina) with some modification to the standard protocol. Genomic DNA samples were prepared by homogenizing pools of whole brains. For queen/worker samples, five pools of eight brains per pool were prepared. For reverted nurse/forager samples, six pools of six brains per pool were prepared. Genomic DNA was extracted from homogenized brains using either Masterpure kit from Epicentre (queen/worker), or lysing cells in Chaos buffer (4.5 M guanidinium thiocyanate, 2% *N*-lauroylsarcosine, 50 mM EDTA, 25 mM Tris-HCl, 0.1M β-mercaptoethanol) and purified with Qiagen DNeasy columns (reverted nurse/forager). For all samples, genomic DNA was sheared to an average size of 350 bp using a Covaris sonicator with the following settings: duty cycle = 10%, intensity = 5.0, bursts per second = 200, duration = 60 s. Blunt ends were created on the DNA fragments using a unique protocol to eliminate the introduction of non-genomic cytosines into the fragments, which would be falsely interpreted as unmethylated cytosines during subsequent analysis. To achieve this, we only used A, G and T nucleotides with a mixture of the enzymes T4 DNA polymerase, Klenow DNA polymerase and T4 PNK(NEB) to perform end repair of fragments. Illumina adapters were then ligated to the fragments after the addition of a single A, per TruSeq protocol. Libraries were then size selected by cutting a 400–500-bp fragment from an agarose gel (BioRad-Certified Low Range Ultra Agarose, NEB 100-bp DNA Ladder, Invitrogen-SYBR Gold nucleic acid gel stain) and purified using Qiagen MinElute Gel Extraction Kit. The purified libraries were then bisulfite converted and purified using Zymo EZ DNA Methylation Gold. Libraries were then amplified using a mixture of Uracil insensitive polymerases; Denville Choice Taq and Agilent Turbo Pfu. Queen/worker samples were amplified for 12 cycles and reverted nurse/forager for 15 cycles using the TruSeq PCR conditions.

RNA sequencing. RNA samples for the six pools of six brains of reverted nurse and foragers were derived from the same lysate that was used to create the reverted nurse and forager WGBS libraries. RNAseq libraries were created using the Illumina TruSeq RNA sample prep kit with no modifications to the standard protocol. This kit enriches for mRNA by using beads bound with a poly-T oligo to bind to the poly-A tails of mRNA.

Data analysis of queen and worker WGBS. We ran the Bsmooth[27] bisulfite alignment pipeline (version 0.4.5-beta) on the 100 × 100 nucleotide paired-end HiSeq 2000 sequencing reads obtained for each queen and worker pool. We used Bsmooth's Bowtie 2–based alignment pipeline, which employs a version of the unbiased and efficient *in silico* bisulfite conversion approach introduced previously[28]. We used Bowtie 2 version 2.0.0-beta5 (ref. 29). We aligned to a reference index consisting of the Baylor Human Genome Sequencing Center *A. mellifera* assembly version 4.0, the *A. mellifera* mitochondrial sequence and the lambda phage genome. **Supplementary Table 6** summarizes alignment results.

We then used Bsmooth to extract read-level measurements. One read-level measurement corresponds to one instance in which an aligned read overlapped a CpG in the reference genome. The measurement records the genomic position of the CpG, the allele observed in the read, its base quality, the alignment's mapping quality and other related measures.

Using Bsmooth, we filtered read-level measurements in three ways. First, we removed read-level measurements from alignments with mapping quality less than 10. Second, we remove read-level measurements in which the allele in the alignment was neither C nor T. Third, we removed read-level measurements

from sequencing cycles that we deemed unreliable after visually inspecting the M-bias plot. That is, we plotted the fraction of methylated read-level measurements versus sequencing cycle. Ideally, this plot should be flat and horizontal, indicating no strong relationship between sequencing cycle and fraction of methylated read-level measurements. In practice, we found peaks and troughs at the extremes of both mates. We filtered out measurements from the affected cycles. In the case of this data set, we filtered out read-level measurements from the 5′-most eight nucleotides of mate 1, the 3′-most four nucleotides of mate 1 and the 5′-most eight nucleotides of mate 2.

After filtering, we used Bsmooth to sort read-level measurements by genome coordinate and compile them into a table summarizing methylation measurements at each CpG in the reference genome. We used the summarized methylation measurements over the lambda genome, which we assume is entirely unmethylated, to estimate the bisulfite conversion rate. **Supplementary Table 7** summarizes the read-level measurements obtained, how they were filtered and the estimated bisulfite conversion rates. After filtering we only included evidence with a quality score greater than or equal to 20 for a particular CpG.

We then used Bsmooth to smooth the data and determine the correlation in methylation between the WGBS data and CHARM data. Given that ~85% of CpGs that have greater than 25% methylation are located in genes, we compared the methylation levels between WGBS and CHARM by segmenting genes into 1,000-bp windows and found the average smoothed methylation value in each window. Each window was required to contain at least four CHARM probes and eight CpGs. This analysis was performed for each queen and worker sample, and correlation values range from 0.691 to 0.807, mean = 0.755 (**Supplementary Fig. 1a,b**). Methylation profiles were compared to determine the reproducibility of detecting regions of methylation (**Supplementary Fig. 1c–f**).

To assess whether there was any difference between queens and workers, we carried out the following analysis. First, we only analyzed CpGs with a coverage across the ten samples of greater than 10. For each of these CpGs, we performed a t test for difference in methylation mean between the two groups. The t test allows us to measure biological variability. The P values were corrected for multiple testing using the Benjamini-Horchberg procedure and no CpGs were significant at a 0.05 false discovery rate.

Data analysis of forager and reverted nurse WGBS. The same methods were used to analyze the forager and reverted nurse sequencing data as were used to analyze the queen and worker sequencing data (see above). **Supplementary Table 8** summarizes alignment results and **Supplementary Table 9** summarizes the read-level measurements obtained, how they were filtered and the estimated bisulfite conversion rates.

To determine whether the differences in methylation between foragers and reverted nurses that were discovered using CHARM also exist in WGBS data set, we performed t tests on individual CpGs within 500 bp of CHARM reversion DMRs. The average difference of the top three CpGs ranked by

significance in the DMR was calculated and compared with the average CHARM methylation in the DMR. These results are presented as a scatter plot in **Supplementary Figure 6**.

To asses the significance of the overlap in direction of change between CHARM and WGBS, we used the following test. Assuming that there is no correlation between CHARM and WGBS there should be a 50% chance that the direction of change is the same. Thus, we calculate the probability of observing a more extreme statistic by $P(X \geq 45) + P(X \leq 11)$ with X being binomially distributed with 57 trials and 0.5 chance of success.

Data analysis of forager and reverted nurse RNA sequencing. We used TopHat[30] v1.3.3 to align the 100×100 nucleotide paired-end HiSeq 2000 sequencing reads obtained for each nurse and reverted forager pool. We aligned to a reference index consisting of the Baylor Human Genome Sequencing Center *A. mellifera* assembly version 4.0 and the *A. mellifera* mitochondrial sequence. **Supplementary Table 10** summarizes alignment results.

TopHat alignments were overlapped with annotated exons obtained from NCBI to form a table of overlap counts per gene and per sample. An alignment was said to overlap an exon if there was any reference position covered both by the exon and by the alignment. A spliced alignment that spanned an exon without either mate overlapping the exon did not count as overlapping the exon.

Junctions and junction counts emitted by TopHat were combined to form a table of counts per junction and per sample. Two junctions that were not identical, but where their boundaries differed by no more than five nucleotides on one or both sides, were considered to be identical. This was necessary because there is some variability in where exactly TopHat will place junction boundaries. Junction counts were used to determine differentially expressed exon junctions between foragers and reverted nurses. We define alternative splice events as the presence of at least two distinct exon junctions that have opposite expression in the same gene or DMR. To determine the frequency of alternative splicing events in DMRs, we expanded the DMR 500 bp on each side and checked for the presence of exon junctions that were more expressed in foragers colocalizing with exon junctions that were more expressed in reverted nurses. Examples are presented in **Figure 2b** and **Supplementary Figure 7**.

21. Amdam, G.V. *et al. Exp. Gerontol.* **40**, 939–947 (2005).
22. Page, R.E. & Fondrk, M.K. *Behav. Ecol. Sociobiol.* **36**, 135–144 (1995).
23. Laidlaw, H.H. & Page, R.E. *Queen Rearing and Bee Breeding* (Wicwas Press, 1997).
24. Aryee, M.J. *et al. Biostatistics* **12**, 197–210 (2011).
25. Irizarry, R.A. *et al. Genome Res.* **18**, 780–790 (2008).
26. Huang, D.W., Sherman, B.T. & Lempicki, R.A. *Nat. Protoc.* **4**, 44–57 (2009).
27. Hansen, K.D. *et al. Nat. Genet.* **43**, 768–775 (2011).
28. Lister, R. *et al. Nature* **462**, 315–322 (2009).
29. Langmead, B. & Salzberg, S.L. *Nat. Methods* **9**, 357–359 (2012).
30. Trapnell, C., Pachter, L. & Salzberg, S.L. *Bioinformatics* **25**, 1105–1111 (2009).