Report

Current Biology

Robust DNA Methylation in the Clonal Raider Ant Brain

Highlights

- A large proportion of brain DNA methylation in the clonal raider ant is robust
- Genes with robust methylation show high and stable expression
- DNA methylation is not associated with different reproductive and behavioral states
- Evidence for caste-specific DNA methylation in social insects is weak

Authors

Romain Libbrecht, Peter Robert Oxley, Laurent Keller, Daniel Jan Christoph Kronauer

Correspondence

romain.libbrecht@gmail.com

In Brief

Libbrecht et al. show that in the clonal raider ant, brain DNA methylation is robust, particularly in genes with high and stable expression, and is not associated with different reproductive and behavioral states. They also report that currently there is little evidence of differential DNA methylation between the female castes of social insects.

Libbrecht et al., 2016, Current Biology *26*, 1–5
February 8, 2016
http://dx.doi.org/10.1016/j.cub.2015.12.040

CellPress

Please cite this article in press as: Libbrecht et al., Robust DNA Methylation in the Clonal Raider Ant Brain, Current Biology (2016), http://dx.doi.org/10.1016/j.cub.2015.12.040

Current Biology
Report

Robust DNA Methylation in the Clonal Raider Ant Brain

Romain Libbrecht,[1,2,3,*] Peter Robert Oxley,[1,3] Laurent Keller,[2] and Daniel Jan Christoph Kronauer[1]
[1]Laboratory of Social Evolution and Behavior, The Rockefeller University, New York, NY 10065, USA
[2]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
[3]Co-first author
*Correspondence: romain.libbrecht@gmail.com
http://dx.doi.org/10.1016/j.cub.2015.12.040

SUMMARY

Social insects are promising model systems for epigenetics due to their immense morphological and behavioral plasticity. Reports that DNA methylation differs between the queen and worker castes in social insects [1–4] have implied a role for DNA methylation in regulating division of labor. To better understand the function of DNA methylation in social insects, we performed whole-genome bisulfite sequencing on brains of the clonal raider ant *Cerapachys biroi*, whose colonies alternate between reproductive (queen-like) and brood care (worker-like) phases [5]. Many cytosines were methylated in all replicates (on average 29.5% of the methylated cytosines in a given replicate), indicating that a large proportion of the *C. biroi* brain methylome is robust. Robust DNA methylation occurred preferentially in exonic CpGs of highly and stably expressed genes involved in core functions. Our analyses did not detect any differences in DNA methylation between the queen-like and worker-like phases, suggesting that DNA methylation is not associated with changes in reproduction and behavior in *C. biroi*. Finally, many cytosines were methylated in one sample only, due to either biological or experimental variation. By applying the statistical methods used in previous studies [1–4, 6] to our data, we show that such sample-specific DNA methylation may underlie the previous findings of queen- and worker-specific methylation. We argue that there is currently no evidence that genome-wide variation in DNA methylation is associated with the queen and worker castes in social insects, and we call for a more careful interpretation of the available data.

RESULTS AND DISCUSSION

The clonal raider ant *Cerapachys biroi* provides a good system to investigate insect DNA methylation, because age-matched individuals that are genetically identical can be collected easily [7]. *C. biroi* has no distinct queen and worker castes. Instead, all ants in a colony produce female offspring by parthenogenesis [8], and colonies undergo stereotypical cycles alternating between queen-like reproductive phases (ants lay eggs inside the nest) and worker-like brood care phases (ants do not lay eggs but nurse the brood and forage for food) [5]. To characterize the brain methylome of *C. biroi*, we sequenced eight samples of bisulfite-treated DNA extracted from pools of 20 brains dissected from age-matched ants collected in the reproductive phase (four samples) and in the brood care phase (four samples) from four source colonies belonging to two different clonal lineages (Experimental Procedures).

The average proportion of methylated cytosines across the eight samples was 2.1% ± 0.1% (mean ± SD), which is substantially higher than what has been reported for the honeybee (0.1%) [1] and other ant species (0.3% in *Camponotus floridanus* and 0.2% in *Harpegnathos saltator*) [2]. Methylation-sensitive AFLP on additional samples confirmed higher levels of methylation in *C. biroi* than in other social insects (Table S1; Supplemental Experimental Procedures). DNA methylation was found primarily in CpG dinucleotides (66.3% ± 1% of the methylated cytosines) and within genes (82.5% ± 0.6%), especially in exons (57% ± 0.9%). Such exonic CpG methylation has been reported in other insect species and in mammals, and it may affect gene function through histone modifications [9], nucleosome stability [10], and/or alternative splicing [1, 2, 11]. As previously shown in other ant species [2], levels of DNA methylation in *C. biroi* were associated with patterns of alternative splicing (Figure S1; Supplemental Experimental Procedures), and transposable elements were hypomethylated compared to the genome baseline (Wilcoxon rank-sum test, $W = 64$, $p = 0.0002$; Table S2; Supplemental Experimental Procedures).

Robust DNA Methylation Is Associated with Highly Expressed Genes Involved in Core Functions

On average, 29.5% ± 1.7% of the methylated cytosines in a given sample showed robust methylation, as they were methylated in all eight samples, despite behavioral, reproductive, and genotypic differences among samples. Additionally, the percentage of sequencing reads indicating methylation was higher for the cytosines that were methylated in all samples (58.2% ± 0.4%) than for those that were methylated in only a subset of samples (17.4% ± 1.9%). Strikingly, 99.3% ± 0.1% of the cytosines with more than 60% reads indicating methylation were methylated in all samples (Figure S2). This suggests that DNA methylation is not only robust across samples but also within samples, hence across individual brains. However, to more definitively assess variation in DNA methylation across

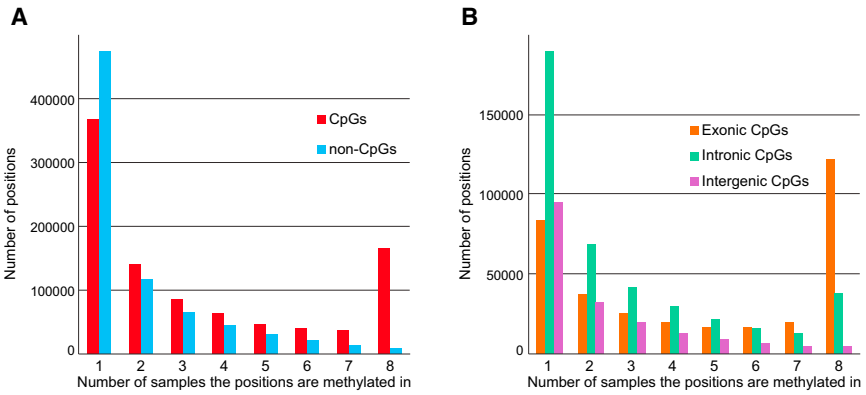


Figure 1. Robust Methylation Is Context and Location Dependent

The graphs show the number of methylated cytosines that are methylated in different numbers of samples (from one to eight) for CpG and non-CpG contexts (A) and for exonic, intronic, and intergenic CpGs (B). Most methylated cytosines are methylated in one sample only (random or sample-specific methylation) or in all eight samples (robust methylation). See also Tables S1 and S3 and Figure S2.

(A) Levels of robust methylation differ between CpG and non-CpG contexts, as illustrated by the sharp increase observed between seven and eight samples for CpGs, but not for non-CpGs.

(B) Levels of robust methylation differ across genomic locations: DNA methylation is more robust in exons compared to introns (sharper increase between seven and eight samples for exonic CpGs than intronic CpGs) and in introns compared to intergenic regions (increase between seven and eight samples for intronic CpGs, but not intergenic CpGs).

individuals would require very deep sequencing coverage from single brains.

The degree of robust DNA methylation differed between CpG and non-CpG contexts and across genomic locations. While 164,258 CpG positions (41.3% ± 2.2% of the methylated CpGs) were methylated in all eight samples, only 9,047 non-CpG positions (4.8% ± 0.4% of the methylated non-CpGs) were methylated in all samples, revealing that CpG methylation is more robust than non-CpG methylation (Figure 1A). Similarly, while 121,858 exonic CpGs (60.9% ± 3.8% of the methylated exonic CpGs) were methylated in all eight samples, only 38,036 intronic CpGs (26.2% ± 1.5% of the methylated intronic CpGs) and 4,364 intergenic CpGs (8.3% ± 0.5% of the methylated intergenic CpGs) were methylated in all samples, revealing that DNA methylation is more robust in exons compared to introns and in genic (exons and introns) compared to intergenic regions (Figure 1B).

The comparison between genes with and without robust methylation revealed that genes with robust methylation (i.e., with at least one cytosine methylated in all eight samples) were significantly enriched for gene ontology (GO) terms related to core processes, such as DNA repair; RNA binding and processing; and protein translation, folding, transport, and binding (Table S3). Genes with robust methylation also were more expressed than genes without robust methylation (Wilcoxon rank-sum test, $W = 5{,}216{,}694$, $p < 0.0001$). More generally, there was a positive relationship between the level of expression and the level of methylation (Spearman rank-correlation test, $rho = 0.59$, $p < 0.0001$; Figure 2A). DNA methylation may preferentially target highly expressed genes and/or DNA methylation may enhance gene expression.

DNA Methylation Is Not Associated with Reproduction and Behavior

To determine whether parts of the *C. biroi* methylome are associated with reproduction and behavior, we performed two analyses to investigate whether DNA methylation differs between brains of age-matched ants in the reproductive phase and in the brood care phase. First, we compared the proportion of methylated reads between the two phases for each CpG. There was no CpG for which the proportion of methylated reads significantly differed between phases after correcting for multiple testing (all p values > 0.22). Second, we used the methylation status of each CpG (methylated or not methylated) to calculate the number of CpGs that were methylated in all four samples from one phase and not methylated in all four samples from the other phase. Then we determined whether such a number of differentially methylated CpGs could be expected by chance by repeating the analysis for all possible sample randomizations. We found 1,560 differentially methylated CpGs between the reproductive phase and the brood care phase, while random comparisons returned an average of 1,727 ± 222 differentially methylated CpGs (median = 1,705; ranging from 1,418 to 2,115; Figure S3). This suggests that the 1,560 apparently differentially methylated CpGs were false positives. Therefore, our analyses did not detect any significant differences in DNA methylation between brains of ants in the reproductive phase and brains of ants in the brood care phase.

In line with the finding that DNA methylation is not associated with reproduction and behavior in the context of colony cycles in *C. biroi*, there was a strong negative relationship between the level of DNA methylation and the level of differential gene expression. Genes that were differentially expressed between the reproductive phase and the brood care phase had fewer methylated sites, while genes with a stable expression between phases tended to be more methylated (Spearman rank-correlation test, $rho = -0.32$, $p < 0.0001$; Figure 2B). Because our analyses did not detect differentially methylated CpGs and DNA methylation is less likely to be found in genes that are differentially expressed between phases, it is unlikely that DNA methylation is involved in the regulation of the clonal raider ant colony cycles.

Re-evaluating the Evidence for Caste-Specific DNA Methylation

Our finding that DNA methylation is robust and not associated with changes in reproduction and behavior in *C. biroi* seems to contradict previous studies that reported DNA methylation differences between the queen and worker castes in four social insect species. Although the findings of caste-specific DNA methylation have been reviewed extensively in the

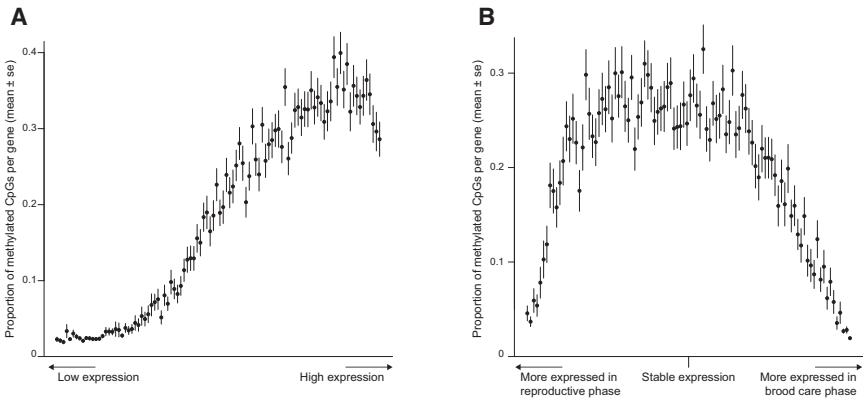


Figure 2. Relationship between DNA Methylation and Gene Expression and between DNA Methylation and Proportional Change in Gene Expression between the Phases of the Colony Cycle

(A) There is a positive relationship between the proportion of methylated CpGs per gene and gene expression. Genes were ranked according to their mean expression across the eight samples before being divided into 100 bins. For each bin, we plotted the mean ± SE proportion of methylated CpGs per gene.

(B) Genes with stable expression between phases tend to be more methylated than genes with differential expression. Genes were ranked depending on how differential their expression was before being divided into 100 bins: in the center are genes with stable expression, on the left those that are more expressed in the reproductive phase compared to the brood care phase, and on the right those that are more expressed in the brood care phase compared to the reproductive phase. For each bin, we plotted the mean ± SE proportion of methylated CpGs per gene. See also Table S2 and Figure S1.

literature [12–27], there are only four empirical studies that used whole-genome bisulfite sequencing to report such differences in ants and bees [1–4]. All those studies investigated differential methylation using the same statistical method, which does not require biological replicates but is prone to producing false positives stemming from sample-specific DNA methylation.

We used the *C. biroi* methylome to assess the validity of the statistical method used in previous studies. First, we investigated whether sample-specific DNA methylation occurred in *C. biroi* by comparing DNA methylation across the eight samples. We found that, on average, 105,321 ± 18,935 cytosines (17.8% ± 2.7% of the methylated cytosines) and 46,027 ± 6,453 CpGs (11.5% ± 1.3% of the methylated CpGs) showed sample-specific DNA methylation. Second, we applied the statistical method used in previous studies to our own data (Supplemental Experimental Procedures). Instead of performing one analysis with four replicates, we performed four separate analyses, each comparing the reproductive phase and the brood care phase of one source colony. We found several hundred differentially methylated exons between the phases for all four source colonies (Figure 3), which is in striking contrast to our combined analysis of the four replicates. However, overlapping the results from the four separate comparisons revealed no exon that was consistently significantly differentially methylated between the two phases in all four analyses (Figure 3). This shows that the lists of differentially methylated exons generated by the statistical method used in previous studies are random or colony specific, and they likely stem from sample-specific DNA methylation.

To our knowledge there are only two empirical genome-wide studies of DNA methylation in social insects that used a replicated experimental design to test whether methylation differs between queens and workers in honeybees [28], *Dinoponera* ants, and *Polistes* wasps [29]. Neither of the two studies detected significant differences in DNA methylation between queen and worker brains (Supplemental Experimental Procedures), which is consistent with our finding that brain DNA methylation does not differ between the reproductive and brood care phases in the clonal raider ant.

Conclusions

The use of biological replicates allowed us to conduct a proper study of the brain methylome of the clonal raider ant *C. biroi*. Our analysis reveals that a large proportion of methylation is robust both across and within samples, especially in exonic CpGs of highly expressed genes involved in general processes. We also report that DNA methylation is unlikely to be involved in regulating the reproductive and behavioral dynamics of the *C. biroi* colony cycle. Finally, evaluating the statistical method used in previous studies with our data indicates that there currently is no empirical evidence for genome-wide variation in DNA methylation associated with the queen and worker castes in other social insect species. Such a lack of well-supported evidence does not necessarily imply that caste-specific methylation does not exist, but rather calls for more controlled and carefully replicated studies of DNA methylation in insect societies.

EXPERIMENTAL PROCEDURES

Sample Preparation

In *C. biroi*, the presence or absence of larvae triggers the switch between the phases of the colony cycle [30]. We used this effect of the larvae to prepare the samples for our study. We first collected 500 callow (recently eclosed) workers, which are light-colored age-matched ants, from a source colony in the brood care phase. We split those callows into two subcolonies, from one of which we removed all the larvae. The subcolony with the larvae remained in the brood care phase, while the other entered a new reproductive phase. We then waited a complete cycle (circa 34 days) until the two subcolonies were again at opposite ends of the cycle. The subcolony in the brood care phase was flash frozen 6 days after the ants started foraging, while the subcolony in the reproductive phase was flash frozen when the first eggs were laid. Thus, the ants collected in the brood care phase and in the reproductive phase were the same age, and they were morphologically and genetically identical (all came from the same source colony, i.e., the same clonal genotype).

For each subcolony, we dissected the brains of 30 individuals with two ovarioles [8], pooled 20 brains to extract DNA for whole-genome bisulfite sequencing, and pooled ten brains to extract RNA for RNA sequencing (see the Supplemental Experimental Procedures for DNA and RNA extraction protocols). We repeated this entire process four times using four different source colonies spanning two clonal lineages: source colonies A1 and A2 (C1 and C16 from clonal lineage A or MLL1 in [31]), and B1 and B2 (STC1 and STC6 from clonal lineage B or MLL4 in [31]). This resulted in eight DNA samples and eight

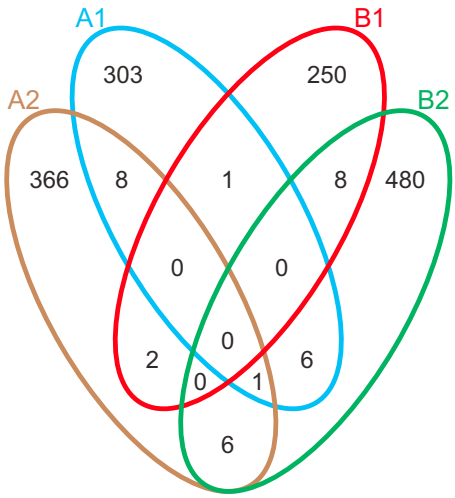


Figure 3. The Lists of Differentially Methylated Exons Returned by the Statistical Method Used in Previous Studies without Biological Replicates Are Random or Colony-Specific Lists of Exons

This graph shows the number of differentially methylated exons between the reproductive phase and the brood care phase for each source colony: 319 in colony A1, 383 in colony A2, 261 in colony B1, and 501 in colony B2 (see details in the Supplemental Experimental Procedures). There was no exon that was consistently differentially methylated between phases in all four source colonies. This shows that the statistical method used in previous studies, especially when used without biological replicates [1–4, 6], is prone to return random or colony-specific lists of exons. See also Figure S3.

RNA samples (four in the reproductive phase and four in the brood care phase for both DNA and RNA).

Library Preparation and Sequencing

Library preparation for whole-genome bisulfite sequencing and RNA sequencing, sequencing, and post-processing of the raw data were performed at the Epigenomics Core at Weill Cornell Medical College (see the Supplemental Experimental Procedures for details). Each phase and each clonal lineage was equally represented in each of the two batches of library preparation and sequencing.

Methylated Cytosines

For each position with coverage $\geq$ 10 in each sample (on average 63.6% ± 4.6% of the cytosines had a coverage $\geq$ 10), the methylation status (methylated or not methylated) was determined by comparing the proportion of sequencing reads indicating methylation (methylated reads) to a binomial distribution, where the number of trials is the number of reads (coverage), the number of successes is the number of methylated reads, and the probability of success is the conversion rate of the bisulfite sequencing treatments. If the proportion of methylated reads could not be explained by chance ($p < 0.05$ after correcting for multiple testing [32]), the position was considered methylated. If it could, the position was considered unmethylated.

Differentially Methylated CpGs

Quantitative Method

For each CpG with coverage $\geq$ 10 in all samples, we performed a paired t test to compare the proportion of methylated reads between the reproductive phase (four replicates) and the brood care phase (four replicates), and then we corrected the p values for multiple testing [32].

Permutation Method

We counted the number of CpGs with coverage $\geq$ 10 in all samples that were methylated in the four samples of one phase but unmethylated in the four samples of the other phase. We then compared this number to the numbers for all possible combinations of four and four samples to assess the number of differentially methylated CpGs that could be expected by chance.

ACCESSION NUMBERS

The accession number for the sequencing data reported in this paper is NCBI BioProject: PRJNA304722.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2015.12.040.

AUTHOR CONTRIBUTIONS

R.L., P.R.O., L.K., and D.J.C.K. designed the experiments. R.L. conducted the experiments. R.L. and P.R.O. analyzed the data. R.L., P.R.O., L.K., and D.J.C.K. wrote the manuscript.

ACKNOWLEDGMENTS

This work was supported by a Marie Curie International Outgoing Fellowship (PIOF-GA-2012-327992) to R.L., as well as NIH grant 1DP2GM105454-01, a Klingenstein-Simons Fellowship Award in the Neurosciences, and an Irma T. Hirschl/Monique Weill-Caulier Trusts Research Award to D.J.C.K. P.R.O. was supported by a grant from the Leon Levy Foundation for Mind, Brain, and Behavior. We thank the Epigenomics Core at Weill Cornell Medical College for technical support. Ryszard Maleszka, Sylvain Foret, and two anonymous reviewers provided helpful comments on an earlier version of the manuscript.

Received: August 29, 2015
Revised: October 30, 2015
Accepted: December 9, 2015
Published: January 21, 2016

REFERENCES

1. 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.
2. Bonasio, R., Li, Q., Lian, J., Mutti, N.S., Jin, L., Zhao, H., Zhang, P., Wen, P., Xiang, H., Ding, Y., et al. (2012). Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Curr. Biol. *22*, 1755–1764.
3. Glastad, K.M., Hunt, B.G., Yi, S.V., and Goodisman, M.A. (2014). Epigenetic inheritance and genome regulation: is DNA methylation linked to ploidy in haplodiploid insects? Proc. Biol. Sci. *281*, 20140411.
4. 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.
5. Ravary, F., and Jaisson, P. (2002). The reproductive cycle of thelytokous colonies of *Cerapachys biroi* Forel (Formicidae, Cerapachyinae). Insectes Soc. *49*, 114–119.
6. Drewell, R.A., Bush, E.C., Remnant, E.J., Wong, G.T., Beeler, S.M., Stringham, J.L., Lim, J., and Oldroyd, B.P. (2014). The dynamic DNA methylation cycle from egg to sperm in the honey bee *Apis mellifera*. Development *141*, 2702–2711.
7. Oxley, P.R., Ji, L., Fetter-Pruneda, I., McKenzie, S.K., Li, C., Hu, H., Zhang, G., and Kronauer, D.J. (2014). The genome of the clonal raider ant *Cerapachys biroi*. Curr. Biol. *24*, 451–458.
8. Ravary, F., and Jaisson, P. (2004). Absence of individual sterility in thelytokous colonies of the ant *Cerapachys biroi* Forel (Formicidae, Cerapachyinae). Insectes Soc. *51*, 67–73.
9. Singer, M., Kosti, I., Pachter, L., and Mandel-Gutfreund, Y. (2015). A diverse epigenetic landscape at human exons with implication for expression. Nucleic Acids Res. *43*, 3498–3508.

10. Hunt, B.G., Glastad, K.M., Yi, S.V., and Goodisman, M.A. (2013). The function of intragenic DNA methylation: insights from insect epigenomes. Integr. Comp. Biol. *53*, 319–328.

11. Li-Byarlay, H., Li, Y., Stroud, H., Feng, S., Newman, T.C., Kaneda, M., Hou, K.K., Worley, K.C., Elsik, C.G., Wickline, S.A., et al. (2013). RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. Proc. Natl. Acad. Sci. USA *110*, 12750–12755.

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

13. Isles, A.R. (2015). Neural and behavioral epigenetics; what it is, and what is hype. Genes Brain Behav. *14*, 64–72.

14. Bonasio, R. (2015). The expanding epigenetic landscape of non-model organisms. J. Exp. Biol. *218*, 114–122.

15. Welch, M., and Lister, R. (2014). Epigenomics and the control of fate, form and function in social insects. Curr. Opin. Insect Sci. *1*, 31–38.

16. Yan, H., Simola, D.F., Bonasio, R., Liebig, J., Berger, S.L., and Reinberg, D. (2014). Eusocial insects as emerging models for behavioural epigenetics. Nat. Rev. Genet. *15*, 677–688.

17. Herb, B.R. (2014). Epigenetics as an answer to Darwin's "special difficulty". Front. Genet. *5*, 321.

18. Maleszka, R. (2014). The social honey bee in biomedical research: realities and expectations. Drug Discov. Today Dis. Models *12*, 7–13.

19. Alvarado, S., Fernald, R.D., Storey, K.B., and Szyf, M. (2014). The dynamic nature of DNA methylation: a role in response to social and seasonal variation. Integr. Comp. Biol. *54*, 68–76.

20. Bonasio, R. (2014). The role of chromatin and epigenetics in the polyphenisms of ant castes. Brief. Funct. Genomics *13*, 235–245.

21. Dolezal, A.G., and Toth, A.L. (2014). Honey bee sociogenomics: a genome-scale perspective on bee social behavior and health. Apidologie *45*, 375–395.

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

23. Weiner, S.A., and Toth, A.L. (2012). Epigenetics in social insects: a new direction for understanding the evolution of castes. Genet. Res. Int. *2012*, 609810.

24. Drewell, R.A., Lo, N., Oxley, P.R., and Oldroyd, B.P. (2012). Kin conflict in insect societies: a new epigenetic perspective. Trends Ecol. Evol. *27*, 367–373.

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

26. Glastad, K.M., Hunt, B.G., Yi, S.V., and Goodisman, M.A. (2011). DNA methylation in insects: on the brink of the epigenomic era. Insect Mol. Biol. *20*, 553–565.

27. Lyko, F., and Maleszka, R. (2011). Insects as innovative models for functional studies of DNA methylation. Trends Genet. *27*, 127–131.

28. Herb, B.R., Wolschin, F., Hansen, K.D., Aryee, M.J., Langmead, B., Irizarry, R., Amdam, G.V., and Feinberg, A.P. (2012). Reversible switching between epigenetic states in honeybee behavioral subcastes. Nat. Neurosci. *15*, 1371–1373.

29. Patalano, S., Vlasova, A., Wyatt, C., Ewels, P., Camara, F., Ferreira, P.G., Asher, C.L., Jurkowski, T.P., Segonds-Pichon, A., Bachman, M., et al. (2015). Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. Proc. Natl. Acad. Sci. USA *112*, 13970–13975.

30. Ravary, F., Jahyny, B., and Jaisson, P. (2006). Brood stimulation controls the phasic reproductive cycle of the parthenogenetic ant *Cerapachys biroi*. Insectes Soc. *53*, 20–26.

31. Kronauer, D.J., Pierce, N.E., and Keller, L. (2012). Asexual reproduction in introduced and native populations of the ant *Cerapachys biroi*. Mol. Ecol. *21*, 5221–5235.

32. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Stat. Methodol. *57*, 289–300.

Current Biology

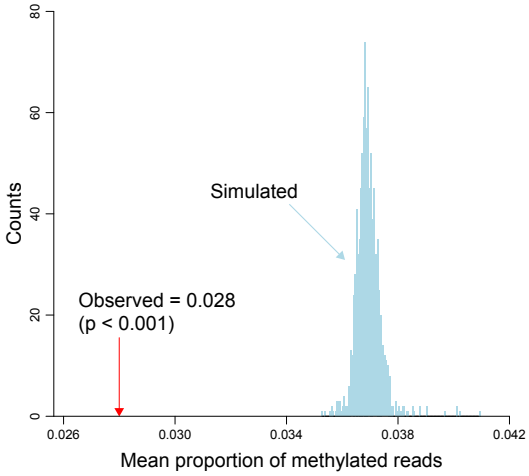
Supplemental Information

Robust DNA Methylation in the Clonal Raider Ant Brain

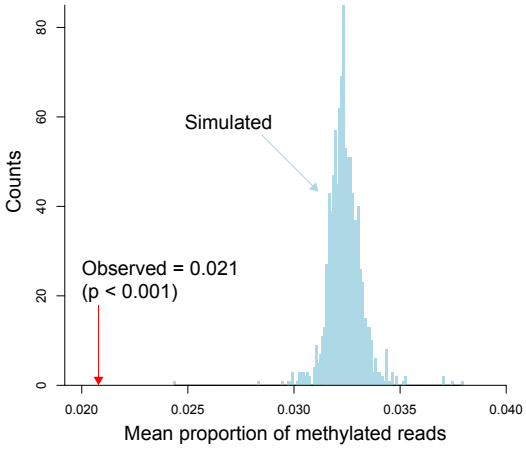
Romain Libbrecht, Peter Robert Oxley, Laurent Keller, and Daniel Jan Christoph Kronauer

Figure S1 – Related to Figure 2

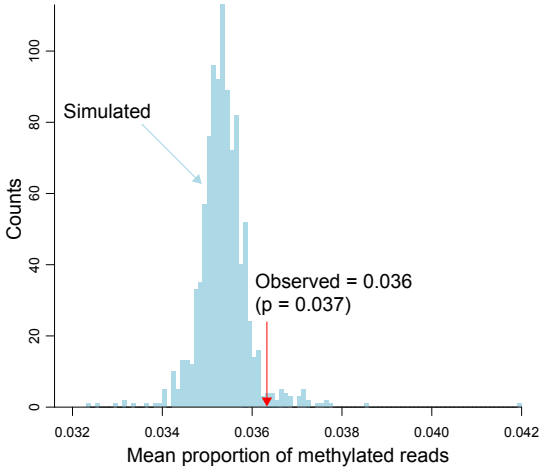
A/ Skipped exons



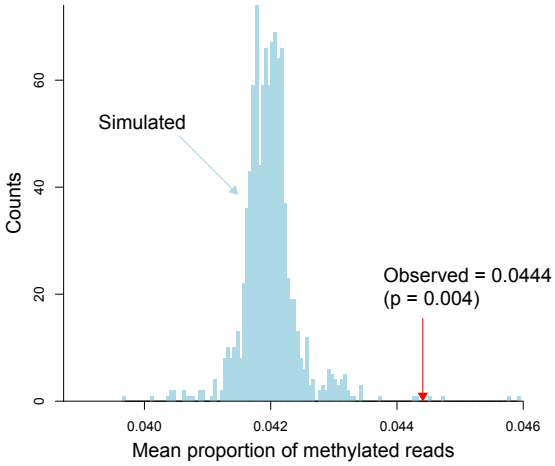
B/ Mutually exclusive exons



C/ Retained introns



D/ Alternative 3' splice sites



E/ Alternative 5' splice sites

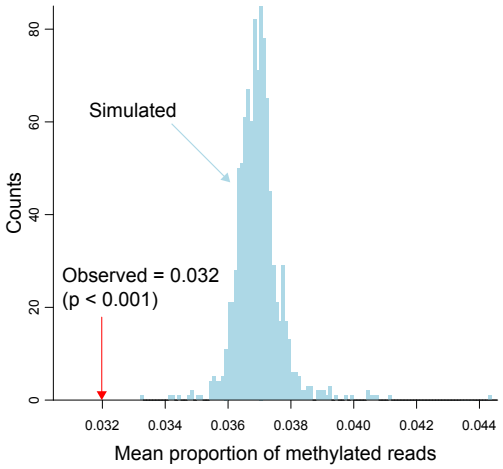


Figure S1 – Alternative splicing is associated with altered levels of DNA methylation.

For each of five types of alternative splicing (skipped exons, mutually exclusive exons, retained introns, alternative 3' splice sites and alternative 5' splice sites), we generated a list of all alternatively spliced exons (Supplemental Experimental Procedures). For each exon, we randomly selected another exon in the genome that had the same position in a gene with similar expression (Supplemental Experimental Procedures). We repeated this process to generate 1,000 random lists of exons for each type of alternative splicing. Then we compared the mean proportion of methylated reads per exon calculated from the empirical list of exons to the random distribution generated from the random lists of exons. The analyses revealed that skipped exons (A), mutually exclusive exons (B) and alternative 5' splice sites (E) were associated with lower levels of DNA methylation than expected, while retained introns (C) and alternative 3' splice sites (D) were associated with higher levels of DNA methylation than expected.

Figure S2 – Related to Figure 1

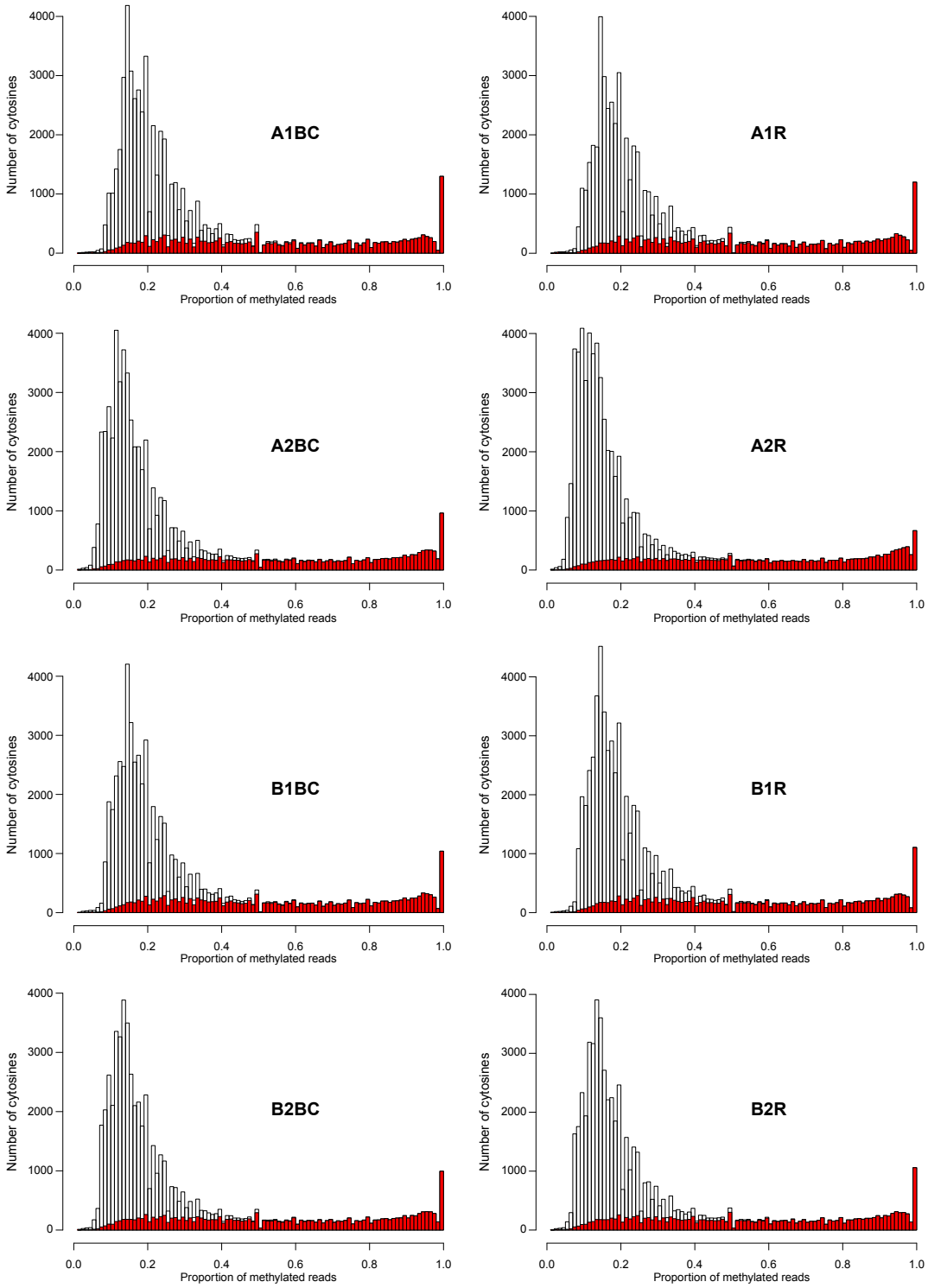


Figure S2 – Cytosines that were robustly methylated across samples also showed robust methylation within samples.

This graph shows for each sample the distribution of the proportion of methylated reads for the methylated cytosines. Cytosines that were methylated in all eight samples (in red) had a higher proportion of methylated reads compared to cytosines that were only methylated in a subset of samples (in white). Almost all cytosines with more than 60% methylated reads were methylated in all eight samples. Robust methylation across samples is thus associated with robust methylation within samples. For each graph, the first two symbols (“A1”, “A2”, “B1”, “B2”) indicate the source colony, “BC” stands for brood care phase, and “R” for reproductive phase. White and red bars are stacked.

Figure S3 – Related to Figure 3

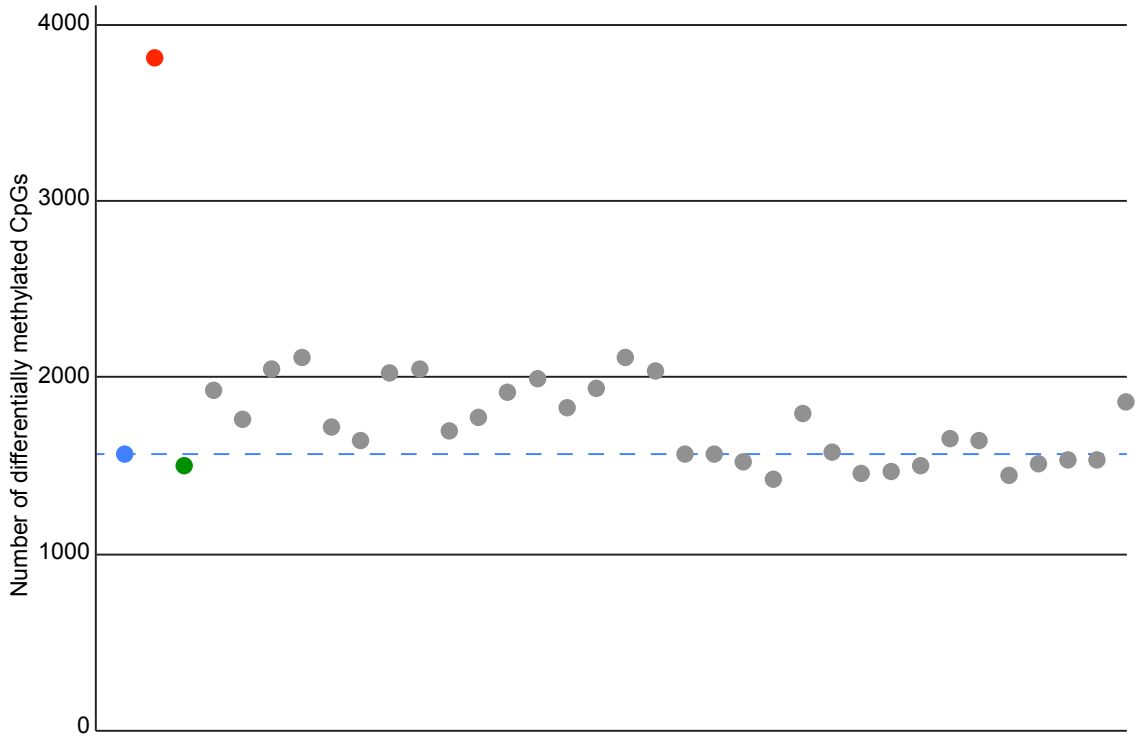


Figure S3 – The comparison between the reproductive phase and the brood care phase did not return more differentially methylated CpGs than expected by chance

This graph shows the number of differentially methylated CpGs obtained when comparing a set of four samples (e.g., the four samples collected in the reproductive phase) to another set of four samples (e.g., the four samples collected in the brood care phase). A given CpG was considered differentially methylated if it was methylated in all the samples of one set but unmethylated in all the samples of the other set. Each dot corresponds to one comparison of two sets of samples, and there are as many dots as there are possible combinations of four and four samples. The blue dot is the comparison between the reproductive phase and the brood care phase. The red dot is the comparison between the first and the second batch of sequencing. The green dot is the comparison between the clonal lineage A and the clonal lineage B. The grey dots are the thirty-two random comparisons that do not have any biological basis.

Neither the comparison between the reproductive phase and the brood care phase (blue dot) nor the comparison between clonal lineage A and clonal lineage B (green dot) had more differentially methylated CpGs than the random comparisons (grey dots), i.e. what could be expected by chance.

The finding of more differentially methylated CpGs when comparing the two batches of sequencing (red dot) compared to random comparisons (grey dots) shows that there is a batch effect in our bisulfite sequencing data, and suggests that differential methylation could in fact have been detected by this analysis if it had existed at appreciable levels.

We performed the same analysis to look for differential methylation in non-CpGs. We found 1,053 non-CpGs that were differentially methylated between the two phases, while random comparisons returned an average of 1,196 differentially methylated non-CpGs (median = 1,189; ranging from 942 to 1,482). Thus, as in CpGs, we did not detect significant differential methylation in non-CpGs between the two phases of the *C. biroi* colony cycle.

Table S1 – Related to Figure 1

	Species name	ms-AFLP methylation rate
Ants	***Cerapachys biroi***	**0.76**
	Aphaenogaster albisetosa	0.04
	Camponotus festinatus	0.07
	Messor pergandei	0.06
	Pogonomyrmex barbatus	0.38
	Pheidole obtusospinosa	0.09
Bees	*Apis mellifera*	0.05
	Melipona bicolor	0.08
	Trigona spinipes	0.01
Wasps	*Liostenogaster flavolineata*	0.03
	Metapolybia cingulata	0.06
	Polistes dominulus	0.19
	Polybia sericea	0.11
	Vespula pensylvanica	0.11
Termite	*Coptotermes lacteus*	0.07

Table S1 – Methylation sensitive AFLP (ms-AFLP) analysis is consistent with the genome of *C. biroi* being more methylated than the genomes of other previously studied social insects [S2-S4]. The methods, enzymes and primers used to perform ms-AFLP in *C. biroi* were the same as in [S3]. The ms-AFLP methylation rate is the estimated percentage of methylated CCGG sites (see [S3] for details). We performed ms-AFLP on eight DNA samples each extracted from a pool of eight heads. On average the ms-AFLP methylation rate was 0.76 ± 0.03 (mean ± sd) in *C. biroi*.

Table S2 – Related to Figure 2

		Mean proportion of methylated cytosines	SE
	Genome	**0.0206**	**0.0010**
LTR Retrotransposons	BEL ***	0.0119	0.0003
	Copia *	0.0158	0.0002
	Gypsy **	0.0141	0.0002
	Others ***	0.0130	0.0005
Non-LTR Retrotransposons	CR1 *	0.0155	0.0006
	Jockey	0.0169	0.0008
	R1 *	0.0142	0.0004
	SINE	0.0201	0.0054
	Others **	0.0147	0.0004
DNA transposons	hAT ***	0.0099	0.0003
	Helitron *	0.0154	0.0006
	Marine Tc1 *	0.0154	0.0003
	Sola	0.0185	0.0005
	Transib ***	0.0126	0.0008
	Others ***	0.0117	0.0002

Table S2 – Transposable elements were hypomethylated compared to the genome baseline.

Eighty percent (12 out of 15) of the transposable element classes included in the analysis had a significantly lower proportion of methylated cytosines compared to the whole genome (*** for $P < 0.001$, ** for $P < 0.01$, * for $P < 0.05$; Supplemental Experimental Procedures).

Table S3 – Related to Figure 1

GO term	Ontology	Description	q value
GO:0003723	MF	RNA binding	< 0.0001
GO:0005515	MF	protein binding	< 0.0001
GO:0003735	MF	structural constituent of ribosome	< 0.0001
GO:0005488	MF	binding	< 0.0001
GO:0005524	MF	ATP binding	< 0.0001
GO:0000166	MF	nucleotide binding	< 0.0001
GO:0004812	MF	aminoacyl-tRNA ligase activity	< 0.0001
GO:0004672	MF	protein kinase activity	< 0.0001
GO:0001104	MF	RNA polymerase II transcription cofactor activity	< 0.001
GO:0008026	MF	ATP-dependent helicase activity	< 0.0001
GO:0005737	CC	cytoplasm	< 0.0001
GO:0005634	CC	nucleus	< 0.0001
GO:0005622	CC	intracellular	< 0.0001
GO:0005840	CC	ribosome	< 0.0001
GO:0016592	CC	mediator complex	< 0.0001
GO:0006886	BP	intracellular protein transport	< 0.0001
GO:0006396	BP	RNA processing	< 0.0001
GO:0006397	BP	mRNA processing	< 0.0001
GO:0008033	BP	tRNA processing	< 0.0001
GO:0006281	BP	DNA repair	< 0.001
GO:0016192	BP	vesicle-mediated transport	< 0.0001
GO:0006511	BP	ubiquitin-dependent protein catabolic process	< 0.0001
GO:0006412	BP	translation	< 0.0001
GO:0006418	BP	tRNA aminoacylation for protein translation	< 0.0001
GO:0006468	BP	protein phosphorylation	< 0.0001
GO:0006457	BP	protein folding	< 0.0001

Table S3 – List of Gene Ontology (GO) terms significantly enriched in genes with robust methylation (n = 6929) compared to genes without robust methylation (n = 3502).

Twenty-six GO terms were significantly enriched in genes with robust methylation. The three ontology categories are molecular function (MF), cellular component (CC) and biological processes (BP). The q values were obtained by correcting the p values for multiple testing [S1]. To determine whether such GO term enrichment could be expected by chance, we randomly generated 10,000 lists of 6929 genes. No GO term was significantly enriched in any of those random lists (all q values > 0.05).

Supplemental Experimental Procedures

DNA extraction

DNA was extracted from pools of 20 brains using the standard protocol of the QIAamp® DNA Micro Kit (Qiagen) with a final elution in 40 μl of buffer AE.

RNA extraction

RNA was extracted using a modified Trizol/phenol chloroform protocol. RNA was extracted using Trizol (Invitrogen) followed by RNeasy (Qiagen) purification with DNAse I (Qiagen) on-column digestion.

Library preparation and sequencing

Library preparation, whole-genome bisulfite sequencing and RNA sequencing were performed at the Epigenomics Core at Weill Cornell Medical College as follows:

Whole-Genome Bisulfite Sequencing (WGBS)

Briefly, 100 ng of DNA were bisulfite converted using the EZ DNA Methylation-Gold Kit (cat # D5005, Zymo Research Corporation, 17062 Murphy Ave. Irvine, CA 92614). The single stranded DNA obtained was processed for library construction using the EpiGenome Methyl-Seq kit EGMK81324 as per the manufacturer's protocols (Illumina Madison, 5602 Research Park Blvd., Suite 200 Madison, WI 53719). 5' tagged random hexamers were annealed to single-stranded DNA and subsequently 3' tagged with a terminal-tagging oligo. The di-tagged DNA was enriched using 10 cycles of PCR, with PCR primers compatible with Illumina sequencing. Each library was made with a unique index sequence and each batch of four libraries was pooled together. The pools were clustered at 7 pM on a paired-end read flow cell and sequenced for 100 cycles on an Illumina HiSeq 2500.

RNA-Seq

RNA-Seq library preparations were done using established Illumina methods for mRNA-Seq (Part #RS-122-2001). Briefly, poly A+ RNA was purified from 200 ng of total RNA with oligo-dT beads. Purified mRNA was fragmented with divalent cations at elevated temperature to ~200bp. First strand cDNA synthesis was performed with random hexamer priming and reverse transcriptase. Second strand cDNA synthesis was performed using RNAseH and DNA PolI. Following dscDNA synthesis, the double stranded products were end repaired, followed by addition of a single 'A' base and ligation to the Illumina TruSeq adaptors. The resulting product was amplified with 15 cycles of PCR. Each library was made with a unique index sequence and each batch of four libraries was pooled together. The pools were clustered at 6.5pM on a paired-end read flow cell and sequenced for 100 cycles on an Illumina HiSeq 2500.

Data processing

Primary processing of sequencing images was done using Illumina's Real Time Analysis software (RTA). CASAVA 1.8.2 software was then used to demultiplex samples and generate raw reads and corresponding quality scores. The WGBS raw data was quality filtered, adapter trimmed, aligned to the *Cerapachys biroi* genome (Official Gene Set version 2.0.1), and methylation calls were generated using the in-house bisulfite sequencing analysis pipeline in the Epigenomics Core at Weill Cornell Medical College [S5]. RNA-Seq reads passing Illumina's purity filter were adapter trimmed and aligned to the *Cerapachys biroi* genome using STAR aligner [S6]. Aligned read counts for each gene were calculated using HTSeq, and a variance-stabilized transformation applied using DESeq2.

Alternative splicing

The *Cerapachys biroi* genome was annotated using NCBI's eukaryotic genome annotation pipeline, identifying 5,112 genes with more than one isoform. These isoforms were searched for skipped exons, mutually exclusive exons, retained introns, and alternative 5' and 3' splice sites. Alternative splicing events were classified using gff_make_annotation.py from the rnaseqlib package (http://yarden.github.com/rnaseqlib), using the "commonshortest" flanking rule. For each alternatively spliced exon, we identified a list of equivalent exons in the *Cerapachys biroi* genome. Using the variance-stabilized transformed expression level from the brain RNA-Seq data, we identified genes expressed between 0.8 and 1.2 times the expression level of the alternatively spliced genes. We then removed those

genes that had alternative splicing in the same manner as the target exon. In the remaining genes, we selected the exon in the same position in the gene (first exon, second exon, etc.) as the target exon. For each of the five types of alternative splicing events, we generated 1,000 lists of exons drawn randomly from the list of equivalent exons, such that each list contained a single non-alternatively spliced equivalent exon for each of the alternatively spliced exons. For each type of alternative splicing, we calculated the mean proportion of methylated reads for the empirical list of exons and for the random lists of exons, and then compared the observed value to the expected distribution.

Transposable elements

Transposable elements in the *C. biroi* genome assembly were identified using RepeatModeler (http://www.repeatmasker.org) and the RepBase [S7] database of repeat elements. All elements that had more than 400 fragments identified in the genome were grouped independently, while all remaining elements were categorized as "others" in the LTR retrotransposon, Non-LTR retrotransposon or DNA transposon classes. SINEs were also grouped independently to enable comparisons with the literature. The RepeatModeler GFF output was used to define the positions of all transposable elements for methylation analysis.

For each transposable element class in each sample, we calculated the proportion of methylated cytosines with a minimum coverage of 10x in the focal sample (Table S2). We built a linear model to compare the mean proportion of methylated cytosines in the transposable element classes to the genome, using the eight samples as replicates (Table S2). Additionally, we used a Wilcoxon rank sum test to compare the mean proportion of methylated cytosines across all classes of transposable elements to the genome (Main text).

Testing the statistical method used in previous studies to detect differentially methylated genomic regions without biological replicates

Most studies that have reported caste-specific differential methylation in social insects used the same statistical method to detect such differences [S8-S11]. We applied this method to our data. To make our analysis comparable to these previous studies, instead of performing one analysis with four replicates, we performed four analyses with one replicate each. Each analysis compared the reproductive phase and the brood care phase for one source colony. For each exon with more than three CpGs and less than 100 CpGs, we built a generalized linear model (binomial family) that explained the proportion of methylated reads by the caste and the position. We used all the CpGs (minimum coverage = 3x) in the focal exon as replicates in the model, as did previous studies that used this method. We analyzed each exon separately and then corrected the p values for multiple testing [S1].

As presented in the main text, the analyses detected several hundred differentially methylated exons ($p < 0.05$) for the four source colonies, but none of those exons were differentially methylated in all four comparisons, revealing that the lists of exons reported by the analyses were actually random or colony-specific lists of exons. The main problem of this statistical method is that it does not involve biological replicates, but uses the different CpGs in a given genomic region as replicates in the model. This makes this statistical method prone to false positives arising from individual variation or experimental noise.

Experimental procedures for the study of DNA methylation in the social insect literature

In the main text we focus on previous studies that performed whole-genome bisulfite sequencing (WGBS) to compare DNA methylation between the queen and worker castes in social insect species. WGBS is the only method providing the genome-wide single nucleotide resolution of DNA methylation comparable to the data we collected in *C. biroi*. In this part of the supplement, we provide a wider review of the experimental procedures used to study DNA methylation in social insects.

Using the whole body to perform whole-genome bisulfite sequencing is problematic

As detailed in the main text and above, most social insect studies that have compared DNA methylation between castes using WGBS used the same statistical method, which does not require biological replicates but is prone to producing false positives stemming from sample-specific DNA methylation [S8-S11]. Additionally, two of those studies used the whole body to extract DNA [S9, S10], which is problematic when comparing social insect castes that differ in morphology and allometry, such as queens and workers [S8]. If tissues that show between-caste differences in their relative proportion to the whole body (e.g., the abdomen in queens and workers) have specific patterns of DNA methylation, this would result in apparent

differences in DNA methylation between castes. In fact, comparing honeybee queen and worker brains [S8, S12] may also be problematic, because a queen brain is structurally different from a worker brain, thus not directly comparable. In our study, the use of morphologically and genetically identical individuals in the reproductive phase and in the brood care phase allowed us to circumvent this problem.

Whole-genome bisulfite sequencing to compare nurses and foragers

To our knowledge there is only one WGBS study of DNA methylation that uses a replicated design to test whether methylation differs between honeybee queens and workers, and between nurses and foragers [S12]. No significant differences in DNA methylation were detected between queen and worker brains, in contrast with a previous study that did report differential methylation [S8]. Those two studies differed in the age and maturity of the individuals used for WGBS, which might help account for some of the discrepancies. Arguably more importantly, however, the two studies also differed with respect to the use of biological replication. Interestingly, Herb et al (2012) reported an association between differences in DNA methylation and behavioral differences within the worker caste, as differentially methylated regions were detected when comparing honeybee nurses and foragers [S12]. Removing nurses prompted some of the foragers to revert to nursing, and differentially methylated regions were also detected when comparing such reverted nurses to the workers that remained foragers. Because there was a significant overlap between the two comparisons, Herb et al (2012) concluded that their data provided evidence for dynamic DNA methylation that can switch back and forth depending on behavior. However, such dynamic patterns of DNA methylation are not required to explain their data, as stable individual differences in DNA methylation may have affected the response threshold to forage, and thus the likelihood of switching back from foraging to nursing tasks when all the nurses were removed. In addition, the workers in each replicate were produced by at least five different queens, resulting in individual genetic differences that could have independently affected DNA methylation [S13, S14] and the transitions between nursing and foraging [S15, S16].

Candidate gene approach

A recent study using a candidate gene approach revealed that larval DNA methylation in the gene *Egfr* is associated with worker size variation in the ant *Camponotus floridanus* [S17]. Another study reported that downregulating *Dnmt3* (coding for the enzyme responsible for *de novo* DNA methylation) in honeybee larvae triggered the development of queens rather than workers [S18]. This finding still awaits confirmation, as the decrease of DNA methylation after *Dnmt3* knockdown was not investigated genome-wide but in only ten CpGs spanning three exons of a single gene [S18].

Methylation sensitive AFLP

Methylation sensitive AFLP provides an estimation of the proportion of methylated sites at the level of the genome. It has been used to assess how common DNA methylation is in a variety of social insect species [S2], but only two studies used it to compare DNA methylation between castes: one in ants [S4] and one in termites [S3]. In *Pogonomyrmex* ants, the analysis revealed a higher proportion of methylated loci in virgin queens than in workers [S4]. In *Coptotermes* termites, the proportion of methylated loci did not differ significantly between sexes or between castes [S3].

Supplemental References

S1. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. 289-300.

S2. Kronforst, M., Gilley, D., Strassmann, J., and Queller, D. (2008). DNA methylation is widespread across social Hymenoptera. Curr. Biol. *18*, 287-288.

S3. Lo, N., Li, B., and Ujvari, B. (2012). DNA methylation in the termite *Coptotermes lacteus*. Insectes soc. *59*, 257-261.

S4. Smith, C.R., Mutti, N.S., Jasper, W.C., Naidu, A., Smith, C.D., and Gadau, J. (2012). Patterns of DNA methylation in development, division of labor and hybridization in an ant with genetic caste determination. PLoS ONE *7*, e42433.

S5. Garrett-Bakelman, F.E., Sheridan, C.K., Kacmarczyk, T.J., Ishii, J., Betel, D., Alonso, A., Mason, C.E., Figueroa, M.E., and Melnick, A.M. (2015). Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. J. Vis. Exp. *96*, 52246.

S6. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinfo. *29*, 15-21.

S7. Jurka, J. (2000). Repbase update: a database and an electronic journal of repetitive elements. Trends Genet. *9*, 418-420.

S8. 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.

S9. Glastad, K.M., Hunt, B.G., Soojin, V.Y., and Goodisman, M.A. (2014). Epigenetic inheritance and genome regulation: is DNA methylation linked to ploidy in haplodiploid insects? Proc. R. Soc. B *281*, 20140411.

S10. Bonasio, R., Li, Q., Lian, J., Mutti, N.S., Jin, L., Zhao, H., Zhang, P., Wen, P., Xiang, H., Ding, Y., et al. (2012). Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Curr. Biol. *22*, 1755-1764.

S11. 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.

S12. Herb, B.R., Wolschin, F., Hansen, K.D., Aryee, M.J., Langmead, B., Irizarry, R., Amdam, G.V., and Feinberg, A.P. (2012). Reversible switching between epigenetic states in honeybee behavioral subcastes. Nat. Neurosci. *15*, 1371-1373.

S13. Gutierrez-Arcelus, M., Lappalainen, T., Montgomery, S.B., Buil, A., Ongen, H., Yurovsky, A., Bryois, J., Giger, T., Romano, L., Planchon, A., et al. (2013). Passive and active DNA methylation and the interplay with genetic variation in gene regulation. Elife *2*, e00523.

S14. Wedd, L., Kucharski, R., and Maleszka, R. (2015). Differentially methylated obligatory epialleles modulate context-dependent LAM gene expression in the honey bee *Apis mellifera*. Epigenetics. *27*, online first.

S15. Calderone, N.W., and Page, R.E. (1988). Genotypic variability in age polyethism and task specialization in the honey bee, *Apis mellifera* (Hymenoptera: Apidae). Behav. Ecol. Sociobiol. *22*, 17-25.

S16. Siegel, A.J., Fondrk, M.K., Amdam, G.V., and Page Jr, R.E. (2013). In-hive patterns of temporal polyethism in strains of honey bees (*Apis mellifera*) with distinct genetic backgrounds. Behav. Ecol. Sociobiol. *67*, 1623-1632.

S17. Alvarado, S., Rajakumar, R., Abouheif, E., and Szyf, M. (2015). Epigenetic variation in the *Egfr* gene generates quantitative variation in a complex trait in ants. Nat. Commun. *6*, 6513.

S18. Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. Science *319*, 1827.