bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # DNA methylation is not a driver of gene expression reprogramming
2 # in young honey bee workers

3 Carlos A. M. Cardoso-Junior[1,2,*], Boris Yagound[2], Isobel Ronai[2], Emily J. Remnant[2], Klaus
4 Hartfelder[1], Benjamin P. Oldroyd[2,3]
5

6 [1] Departamento de Biologia Celular e Bioagentes Patogênicos, Faculdade de Medicina de
7 Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brasil.
[2] Behaviour, Ecology and Evolution (BEE) Laboratory, Ecology and Evolution, School of Life and Environmental Sciences A12, University of Sydney, Sydney NSW 2006, Australia.
8 [3] Wissenschaftskolleg zu Berlin, Wallotstrasse 19, 14193 Berlin, Germany.

9 [*] **Corresponding author**: carloscardoso.bio@alumni.usp.br
10

11 **Contact Information**
12 Carlos A. M. Cardoso-Júnior – carloscardoso.bio@alumni.usp.br; Phone: +55 16 3315 3063
13 Boris Yagound - boris.yagound@sydney.edu.au
14 Isobel Ronai – isobel.ronai@sydney.edu.au
15 Emily Remnant - emily.remnant@sydney.edu.au
16 Klaus Hartfelder - klaus@fmrp.usp.br
17 Benjamin Oldroyd – benjamin.oldroyd@sydney.edu.au
18

19 **Author contributions**
20 CAMCJ conceived the study, performed the field work, carried out the dissections, laboratory
21 molecular work and bioinformatics analyses, and wrote a first draft of the manuscript and
22 revised the subsequent versions; BY contributed to the design of the study, helped with
23 bioinformatic analyses, performed field work and statistical analyses. ER participated in the
24 design of the study and performed bioinformatics analyses of amplicon sequencing. IR helped
25 with the dissections and participated in the gene expression analyses. KH conceived the study
26 and supervised the experiments. BPO participated in the design the study, performed the
27 artificial inseminations and field work, supervised the experiments and contributed to the
28 statistical analyses. All authors made critical contributions during the writing of the manuscript
29 and gave final approval for publication.
30

31 **Disclosure declaration:** There are no competing interests.
32

33 **Running title:** Social environment does not affect DNA methylation
34 **Keywords:** honeybees, whole genome bisulfite sequencing, amplicon sequencing, DNMT,
35 *Apis mellifera*
36

37 **This PDF file includes:**
38 Main Text
39 Figures 1 to 5

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

40 **Abstract**

41 Intragenic DNA methylation, also called gene body methylation, is an evolutionarily-
42 conserved epigenetic mechanism in animals and plants. In social insects, gene body
43 methylation is thought to contribute to behavioral plasticity, for example between foragers and
44 nurse workers, by modulating gene expression. However, recent studies have suggested that
45 the majority of DNA methylation is sequence-specific, and therefore cannot act as a flexible
46 mediator between environmental cues and gene expression. To address this paradox, we
47 examined whole-genome methylation patterns in the brains and ovaries of young honey bee
48 workers that had been subjected to divergent social contexts: the presence or absence of the
49 queen. Although these social contexts are known to bring about extreme changes in behavioral
50 and reproductive traits through differential gene expression, we found no significant
51 differences between the methylomes of workers from queenright and queenless colonies. In
52 contrast, thousands of regions were differentially methylated between colonies, and these
53 differences were not associated with differential gene expression in a subset of genes examined.
54 Methylation patterns were highly similar between brain and ovary tissues and only differed in
55 nine regions. These results strongly indicate that DNA methylation is not a driver of differential
56 gene expression between tissues or behavioral morphs. Finally, despite the lack of difference
57 in methylation patterns, queen presence affected the expression of all four DNA
58 methyltransferase genes, suggesting that these enzymes have roles beyond DNA methylation.
59 Therefore, the functional role of DNA methylation in social insect genomes remains an open
60 question.

61

62

63

64

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

65 **Introduction**

66 DNA methylation is a reversible chemical modification of DNA whereby methyl groups are
67 added to cytosines in CpG dinucleotides by enzymes of the DNA methyltransferase family
68 (DNMTs). In mammals DNA methylation plays important roles in regulating gene expression,
69 including inactivation of the X chromosome in females, transposon suppression and genomic
70 imprinting, a mechanism by which parents influence gene expression in offspring (1). The role
71 of DNA methylation in genomic regulation of invertebrates is less clear, in part because the
72 two main model species, the fruit fly *Drosophila melanogaster* and the nematode
73 *Caenorhabditis elegans,* either lack methylation (*C. elegans*) or have extremely low and
74 transient levels of methylation (*D. melanogaster*) (2, 3). In contrast, the honey bee (*Apis*
75 *mellifera*) genome project revealed a functional epigenetic system comparable to that of
76 vertebrates (4, 5). Since this discovery the honey bee has emerged as a model species for
77 epigenetics studies in invertebrates, as, in contrast to mammals, the methylation marks in honey
78 bee genomes are sparse and mostly restricted to gene bodies (intragenic DNA methylation) (2,
79 6, 7). This sparseness brings a technical advantage to researchers, as it facilitates the study of
80 gene body methylation and its potential roles in the regulation of gene expression without noise
81 from methylated cytosines of other genomic compartments, such as promoters and transposons.
82 Mechanistically, it has been proposed that gene body methylation modulates the affinity of
83 cofactor binding in regulatory DNA methylation-dependent regions in order to regulate RNA
84 polymerase II activity (8–11).

85 In honey bees and other social insects, gene body methylation has been associated with
86 behavioral and phenotypic plasticity (reviewed in (11, 12)). For example, methylome
87 differences have been found between queens and workers in honey bees (6, 13) and ants (7).
88 Differences in gene body methylation were also associated with division of labor among ant
89 (7, 14) and bee (15) workers. Furthermore, DNA methylation has been associated with several

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

90 biological processes, including aging, reproduction, aggressiveness, response to social stimuli,
91 memory and learning, and the haplodiploid sex determination system (15–25).
92 In support of the hypothesis that gene body methylation of honey bees modulates
93 transcription, the RNA interference-mediated knockdown of the *DNA methyltransferase 3*
94 (*Dnmt3*) gene, which codes for the enzyme responsible for *de novo* DNA methylation, affected
95 14% of the honey bee worker transcriptome (9). RNA splicing was also affected, corroborating
96 previous *in-silico* predictions that associated DNA methylation with alternative splicing (6, 7,
97 13, 15, 26–29). Even more spectacularly, the knockdown of *Dnmt3* in young female larvae
98 resulted in adults with a queen-like phenotype, mimicking the transcriptional program induced
99 by a royal jelly diet (16).
100 More recent studies, however, now cast doubt on the role of gene body DNA
101 methylation as a flexible regulator of gene expression in social insects, and an emerging
102 consensus is that DNA methylation is genotype-specific (30–41). For example, a reanalysis of
103 the *Dnmt3* knockdown data of Li-Byarlay et al. (2013) suggested that the original analysis had
104 overestimated the number of regulated genes, and that the *Dnmt3* knockdown had in fact only
105 a minor effect on the honey bee gene body methylation pattern, and hence, on gene expression
106 (37). Furthermore, recent studies have not provided support for the proposed association
107 between differential methylation and alternative splicing in honey bees (38) and other social
108 insects (34, 42–44).
109 Particularly compelling evidence that DNA methylation tends to be sequence-specific
110 rather than representing a flexible gene regulatory mechanism comes from the clonal raider ant
111 (*Ooceraea biroi*) (31). *O. biroi* reproduces asexually, and therefore allows experiments in a
112 uniform genetic background (45). Libbrecht et al. (2016) showed that DNA methylation is not
113 associated with reproductive and asexual stages in *O. biroi* workers, casting doubt on its role
114 in other ants and honey bees. These authors proposed that the previously reported differential

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

115 methylation patterns seen between experimental groups may have be an artefact arising from
116 combinations of colony-specific methylation patterns.
117 In this study, we test the hypothesis that differential methylation is associated with
118 differential gene expression in response to environmental change, versus the alternative
119 hypothesis that differentially methylated regions are a colony and/or individual-specific
120 character that does not vary in response to environmental change. To do this we compared the
121 gene expression and methylation patterns in the brains and ovaries of young honey bee workers
122 as they matured in the presence or absence of their queen. Honey bee queens, through their
123 mandibular gland pheromones, influence the behavioral maturation and reproductive capacity
124 of workers (46–48), and this process involves changes in the expression of hundreds of genes
125 in the brain and ovary (49–52), including the *Dnmt* genes (22, 53). Furthermore, instead of
126 generating data from whole body methylomes, we compared methylation patterns in the tissues
127 (brain and ovary) most likely to respond to the absence/presence of a queen. This allowed us
128 to further examine whether differential methylation is related to tissue function, social context
129 or to genotype.
130
131 **Results**
132 **Social context does not affect the brain and ovary methylomes of young honey bee**
133 **workers**
134 We sequenced and analyzed at single base-pair resolution the complete methylomes of brains
135 and ovaries of honey bee workers reared in queenright or queenless colonies. After removing
136 adaptors and reads of low quality, and aligning the methylomes to the honey bee reference
137 genome (54) we obtained high coverage in all sequenced samples (Supplemental Table S1).
138 The conversion rate of bisulfite treatment was above 99.4% (Supplemental Table S1),
139 indicating a low frequency of false-positive methylated CpGs. The observed frequency of

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

140 methylated CpGs (~1% - Supplemental Table S2) is consistent with previous studies on honey
141 bees and other holometabolous insects (6, 7, 15, 55).
142 With these data we were able to ask whether the methylomes of workers differ when
143 they are in a queenright (QR) or queenless (QL) social context, especially in the two tissues
144 that are known to respond transcriptionally to the presence of the queen: the brain and the ovary
145 (49–51). Comparing three pairs of QR colonies with three pairs of QL colonies we found 879
146 differently methylated regions (DMRs) for the brains of QL and QR workers (Supplemental
147 Table S3) and 376 DMRs for the ovaries (Supplemental Table S4). However, we found that
148 the number of DMRs shared by two or more colonies was very low for both tissues (Fig. 1A).
149 This indicates that the majority of candidate DMRs previously identified (e.g., Supplemental
150 Table S3 and S4) were driven by between-colony variability and do not reflect a reproductible
151 effect of 'social context'.
152 Hierarchical clustering of worker methylomes showed that samples from the same
153 colony cluster together, irrespective of social context (Fig. 1B). When comparing the number
154 of DMRs in the methylomes of the three colonies regardless of social context (i.e., A *vs.* B, A
155 *vs.* C, and B *vs.* C), we found over 10,000 significant DMRs in the pairwise colony comparisons
156 (Supplemental Fig. 1). Therefore, the effect of 'colony' on the worker methylomes is of much
157 greater magnitude than the effect of 'social environment', if there is such an effect.
158 For a more in-depth analysis, we next selected three top-ranked candidate DMRs for
159 both tissues from the QR *vs.* QL comparison for amplicon sequencing (Fig. 1C, Supplemental
160 Table S5). Bisulfite-treated DNA extracts from the same colonies as those used for WGBS
161 (colonies A-C) were PCR-amplified and sequenced in a high-throughput platform. To increase
162 sample size and check for data reproducibility we also added samples from three new pairs of
163 colonies (colonies D-F) to this analysis. The high-coverage obtained by amplicon sequencing
164 (average >10,000 reads/region, range 672-30,823, Supplemental Table S5) suggests that the

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

165 majority of cell types present in the brain and ovary tissues are represented in this dataset.
166 Amplicon sequencing revealed that the methylation differences (QR *vs.* QL) observed in the
167 original WGBS data (Fig. 1C – grey bars) did not reach the differential methylation threshold
168 set at 10%, despite the increased coverage for the original samples (Fig. 1C – blue bars,
169 Supplemental Fig. 2; Supplemental Table S5), or samples from the three new independent
170 colonies (Fig. 1C – yellow bars). Importantly, the differential methylation pattern for the
171 majority of regions (five out of six) from colonies D-F was in the opposite direction to that of
172 colonies A-C. Together, these comparisons provide strong support for the hypothesis that the
173 differences seen in the WGBS comparisons of the QR *vs.* QL workers were indeed a result of
174 strong differences in the methylomes of colonies A-C, and not a consistent consequence of
175 social context, either in the original colonies, nor in the three new colonies.
176

177 **Differential methylation is not associated with differential gene expression when**
178 **comparing different colonies and social context**

179 To further investigate whether DNA methylation-mediated transcriptional responses are
180 triggered by queen exposure, we determined the expression patterns of nine genes displaying
181 DMRs found in the contrast of the QR *vs.* QL methylomes (Supplemental Tables S3, S4). We
182 found that only three differentially methylated genes (*GB42836*, *GB49839* and *GB54664*),
183 were differentially expressed (Fig. 2, Supplemental Table S6), and there was no differential
184 expression for the other six genes in the respective tissue displaying differential methylation.
185 This lack of correlation between differential methylation and differential expression was also
186 statistically confirmed (Supplemental Fig. S3, Pearson correlation $r = 0.37$, $p = 0.23$, $n = 12$).
187 Taken together, although major transcriptional responses to the presence or absence of a queen
188 in a colony are regularly observed and reported for the brain and ovary of workers (49–52), our

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

189 data indicate that there is no direct association between the DNA methylation status of a gene
190 and its differential expression, even in the context of a very dramatic contrast in social context.
191 Given that the genome methylation differences appeared to be much more related to
192 colony genotype than to social context (Fig. 1 and Supplemental Fig. 1), we next asked whether
193 colony-specific methylation patterns are associated with differential gene expression between
194 colonies. To do so, we first estimated the "colony effect" by ascertaining how many regions
195 showed different methylation levels in the brain and ovary of workers, regardless of the social
196 context (Supplemental Fig. S1). We found large numbers of DMRs in both the brain (n = 3,603)
197 and ovary (n = 1,997) that differed in their methylation level by at least 10% when comparing
198 the same genomic region pairwise among the colonies (Fig. 3A). This indicates that there is a
199 core set of genomic regions that are hypervariable with respect to their methylation status.
200 Interestingly, these regions are not chromosome specific (Supplemental Fig. S4), suggesting a
201 colony-specific methylation fingerprint across all chromosomes. Also, almost 60% of the
202 hypervariable regions (n = 1,194) seen in the ovaries were the same as those found in the brains
203 (Fig. 3B), suggesting that these colony-specific signatures are independent of tissue.
204 Next, we asked whether changes in DNA methylation were associated with differential
205 gene expression between colonies. For these analyses we selected 11 genes that were either
206 hypermethylated or hypomethylated in one source colony, but methylated in the opposite
207 direction in the other two source colonies (e.g., hypermethylated in colony A but
208 hypomethylated in colonies B and C). We confirmed the methylation pattern of these regions
209 by amplicon sequencing of colonies A-C (Fig. 3C, Supplemental Table S5) and D-F
210 (Supplemental Fig. S5, Supplemental Table S5). We noticed that the methylation patterns of a
211 given genomic window were reasonably stable across different social contexts and tissues
212 (Supplemental Figs. S2, S6). These analyses revealed strong colony-specific methylation
213 patterns for all six analyzed colonies. Thus, if DNA methylation plays a role in regulating gene

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

214 expression, these genes would be strong candidates for differential gene expression, as their
215 differential methylation across colonies was much greater than the differences we observed for
216 social context.

217 Only one gene (*GB50283*) out of the 12 genes assessed, which was strongly methylated
218 in colony C compared to the other two colonies (Fig. 3C), turned out to be differently expressed
219 among colonies (Fig. 3D, $p < 0.01$, Supplemental Table S7). For the other genes assessed, we
220 found that even clear differences in DNA methylation profiles between colonies (Fig. 3C) did
221 not affect their expression (Fig. 3D). After all these comparisons between different colonies
222 and radically different social contexts, we conclude that there is little evidence in support of
223 the hypothesis that differences in DNA methylation drive alterations in gene expression in the
224 two tissues that are most likely to respond to social context. Hence, the brain and ovary
225 methylomes of young honey bee workers are primarily a manifestation of colony identity rather
226 than a mediator between social environmental changes and gene expression.
227

228 **Brain and ovary methylomes are highly similar, despite the functional differences**
229 **between the two tissues**

230 After showing that DNA methylation of young honey bee workers is colony specific rather
231 than influenced by the social context, we investigated whether there may be tissue-specific
232 differences between the brain and ovary methylomes. Comparing the methylomes of tissues
233 with completely distinct biological functions, we expected to identify hundreds of DMRs
234 associated with the transcriptomic differences between these two tissues. Surprisingly, we
235 found only nine DMRs across all three source colonies when comparing the methylomes of the
236 two tissues, irrespective of social context (Fig. 4A). These regions are associated with only
237 four genes: *GB47277*, *GB51802*, *GB55278*, and *GB50784*. The amplicon sequencing analysis

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

238 performed for two of these DMRs confirmed the differences originally seen in the WGBS data
239 for these two tissues (Fig. 4B and Supplemental Fig. S7).
240 When testing whether alterations in gene expression between the brain and ovaries were
241 associated with differences in DNA methylation levels we found that three of the four
242 differentially methylated genes were also differentially expressed (Fig. 4C, Supplemental
243 Table S6). However, differential expression between tissues were identified for 11 out of the
244 12 genes previously analyzed for colony specificity (Fig. 3D, Supplemental Table S7). Thus,
245 it is not yet clear whether the alterations in methylation level promotes causative alterations in
246 gene expression between the two tissues. From an overall perspective the methylomes of brains
247 and ovaries were strikingly similar, with significant differential methylation seen for only four
248 genes.
249
250 **Expression of the four *Dnmt* genes is influenced by social context**
251 As our data strongly indicate that the presence or absence of a queen does not drive significant
252 alterations in the gene body methylation patterns in the brain and ovaries of young honey bee
253 workers (Fig. 1) we wondered what the expression patterns of the genes encoding the four
254 honey bee DNA methyltransferases would look like, especially since previous studies have
255 used the expression of *Dnmt* genes as an indicator of epigenetic events in social insects (16,
256 18, 20, 23, 25, 56) . Furthermore, we have recently reported that the queen mandibular
257 pheromone upregulated the expression of three *Dnmt* genes (*Dnmt1b*, *Dnmt2* and *Dnmt3*) in
258 the brain of caged honey bee workers (53). Here we report that the expression of all four *Dnmt*
259 genes predicted in the *A. mellifera* genome is affected by the presence or absence of a queen in
260 the colony (Fig. 5, Supplemental Table S6), either in the brain or the ovary. Interestingly, the
261 same social cue (presence/absence of a queen) that upregulated two *Dnmt* genes (*Dnmt1b* and
262 *Dnmt3*) in the brain of young workers had an opposite effect on the expression of three *Dnmt*
263 genes (*Dnmt1a*, *Dnmt2* and *Dnmt3*) in the ovary. This suggests that expression of the *Dnmt*

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

264 genes responds to social cues in both tissues, even though their differential expression
265 apparently is not associated with alterations in the methylomes of honey bee workers.
266
267 **Discussion**
268
269 With the presence or absence of a queen being one of the strongest contrasts in the social
270 context of a colony, and with DNA methylation as a strong candidate for mediating adaptive
271 responses in the individual colony members, we had expected to see signatures of this social
272 context condition in the methylomes of young honey bee workers. However, all the evidence
273 we obtained suggests the opposite: that gene body methylation is not a major regulator of gene
274 expression reprogramming in young honey bees. First, the contrast between the presence *vs.*
275 absence of a queen was not reflected in significant alterations in the respective methylomes,
276 indicating that differential DNA methylation is not a mediator for the transcriptional
277 consequences of queen presence. Second, despite strong differences in gene expression
278 between brain and ovarian tissues (57), we found that their methylomes are very similar. Third,
279 thousands of DMRs were identified when comparing individuals from different colonies.
280 However, even when differences in methylation were greater than 50%, there was no
281 association between DNA methylation and gene expression, except for one gene (*GB50283*,
282 Fig. 3D). Therefore, these data provide strong support for the hypothesis that intragenic CpG
283 methylation is not a driver of gene expression reprograming in the brain and ovary of young
284 honey bees. This result is actually in line with and likely generalizable for other social insects
285 (31, 32, 34, 41, 43, 44, 58).

286 In contrast to our findings, two previous studies have suggested that honey bee workers
287 kept under different social conditions exhibit a low number of DMRs, and that these are
288 associated with differential gene expression (15, 17). A plausible explanation (32, 39) for these
289 contrasting results is that previously-observed differences in the methylomes of honey bee
290 workers in response to divergent social stimuli arose as an artifact of genotype-associated

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

291 methylation variants (33). Indeed, when the colony genetic background was standardized, no
292 differences were found between the methylomes of newly-emerged queens and newly-emerged
293 workers (15). Furthermore, ultra-deep analyses on the dynamics of brain DNA methylation of
294 *Dynactin*, *Nadrin* and *PKCbp1* genes, indicated that their methylation content are not altered
295 over time, and thereby not implicated in honey bee workers' aging and behavioral maturation
296 (59). Thus, we conclude that the methylomes of honey bee workers do not respond dynamically
297 to changes in the social milieu (e.g., loss of the queen), whereas gene expression levels do so
298 (49–53).

299 Our results also show that DNMT-encoding genes are differentially expressed, and thus
300 affected by the presence or absence of a queen. This finding is consistent with previous honey
301 bee studies that showed an effect of queen mandibular pheromones on the expression of *Dnmt*
302 genes (22, 53). However, these alterations seen in the expression of the *Dnmt* genes apparently
303 does not promote a "real-time" epigenetic reprogramming process (this study; Harris et al.
304 2019), indicating that the expression levels of *Dnmt* genes should not be used as a proxy of
305 possible differences in DNA methylation patterns or levels. This result further highlights the
306 weak association between the expression of *Dnmt* genes, DNA methylation, and gene
307 expression reprogramming seen in honey bees and other insects. For instance, even though the
308 genome of the red flour beetle *Tribolium castaneum* is not methylated (60, 61), the knockdown
309 of a *Dnmt1* ortholog resulted in offspring showing developmental arrest and high mortality
310 (61). Similar developmental defects in response to a knockdown of *Dnmt1* were observed in
311 the reproductive system of females of the milkweed bug *Oncopeltus fasciatus* (37, 62). These
312 phenotypic alterations are not a consequence of transcriptomic changes, even though the
313 *Dnmt1* knockdown successfully reduced the global levels of DNA methylation (37).
314 Combined, these results imply that DNMTs have other, not yet understood functions, that go
315 beyond DNA methylation and its maintenance.

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

316 After an unequivocal demonstration that colony-specific methylation patterns are major
317 signatures in the methylomes of honey bee workers (Fig. 3), an interesting question emerges:
318 why do individuals from different colonies exhibit such contrasting differences between their
319 methylomes? While with our current data we are not yet able to fully solve this puzzle, we
320 propose that colony-specific methylation patterns are both functional and genotype-associated;
321 with a minimum effect on transcription (37, 38). Our results show that colony-specific
322 methylation patters are predominantly chromosome and tissue independent (Fig. 3B, Fig. 4.
323 and Fig. S4). Therefore, the functional role of intragenic methylation in honey bees is likely to
324 be associated with basal processes necessary for cell viability, e.g., cell division, genome
325 integrity, and/or cell cycle progression (37, 61–63), rather than representing as a marker of
326 tissue identity (Fig. 4), or acting as a dynamic sensor of changes in the social environment (Fig.
327 1). The finding that the knockdown of the *Dnmt3* gene function, a key orchestrator of honey
328 bee development, is lethal at early development (16) suggests that DNA methylation is
329 essential to honey bee viability. Furthermore, the colony-specific methylation patterns are, at
330 least in part, sequence-specific (32, 39, 40). Ultra-deep amplicon sequencing analyses
331 identified that a honey bee brain displays only a small repertoire of methylation patterns. This
332 may reflect the small number of maternal and paternal epialleles predicted for polyandrous
333 insects (59). Interestingly, the colony-specific methylation patterns are inherited from drones
334 to their worker daughters (40), suggesting a role for the paternally inherited epialleles to the
335 composition of colony-specific worker methylomes.

336 In conclusion, all evidence generated in this study goes against the hypothesis that gene
337 body methylation is a driver of gene expression re-programming in adult honey bee workers.
338 With our experimental design, using several colonies of standardized within-colony genetic
339 backgrounds and high-throughput genomic sequencing approaches we had originally expected
340 to provide evidence in favor of this hypothesis and to see defined differentially-methylated

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

341 gene sets between queen states, with correlated expression patterns. As it has turned out, the
342 conclusion to be drawn is that gene body methylation patterns are essentially colony specific
343 and unaffected by even radical changes in social context. Nonetheless, it is possible that
344 intragenic DNA methylation plays a role in gene expression during other stages of a worker's
345 life cycle. If so, this is likely be restricted to a small set of genes under specific contexts (16,
346 35, 37, 38), including the *GB51802*, *GB55278* and *GB50784* genes that show both tissue-
347 specific methylation patterns and differential expression (Fig. 4C). Our results make it clear
348 that, whatever may be the specific function of gene body DNA methylation in the honey bee,
349 it is now undeniable that there is a strong genotype effect shaping the workers' methylomes,
350 and that future studies need to consider the possible effect of genotype variants (colony and
351 patriline genotypes) in their experimental design. In this sense, we believe that our work makes
352 a significant contribution to the understanding of the meaning of the enigmatic, but
353 evolutionarily-conserved intragenic DNA methylation in an important invertebrate epigenetic
354 model species, the honey bee. As it stands now, we must say that the functional role of gene
355 body methylation in social insects is still an unsolved issue.
356

357 **Material and methods**

358 **Bees and manipulation of the social environment**

359 We used workers from six source colonies. These colonies were headed by queens of standard
360 commercial stock (mostly *A. m. ligustica*) each of which had been instrumentally inseminated
361 with sperm from a single male (64). Hence, the workers within each colony pairs were all full
362 sisters, with a genetic relatedness of 0.75, thus minimizing genotypic heterogeneity within
363 colonies.

364 Queenright and queenless colonies were prepared by splitting each of the six host
365 colonies into two (n = 12), thus generating colony pairs of the same genetic background but

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

366 differing completely in terms of social context. Brood frames were equalized among the colony
367 pairs to ensure a similar condition in terms of brood presence, despite the presence or absence
368 of the queen. After splitting the host colonies into queenless and queenright halves at a remote
369 apiary, we moved the splits to the apiary at the University of Sydney, thus reducing the
370 tendency of the queenless workers moving back to the respective unit with the queen. The
371 experiments were conducted during the southern hemisphere summers of 2017 (source
372 colonies A-C) and 2018 (source colonies D-F).

373 Newly-emerged workers from source colonies A-F were obtained by placing sealed
374 brood frames overnight in separate boxes, in an incubator at 34 °C on the day of colony
375 splitting. The newly-emerged workers were divided into two groups (n = 200 bees per group),
376 paint marked with different colors, and then introduced into their respective pair of queenright
377 and queenless host colonies. After four days, the marked workers were collected, snap frozen
378 on dry ice, and stored at -80 °C. At this age workers are known to respond transcriptionally to
379 the presence or absence of a queen (49, 53). As expected, the queen was seen in all QR host
380 colonies at the time of sampling, but never in the QL host colonies. Samples from six of these
381 colonies (colonies queenright A-C and queenless A-C) were initially used for methylome
382 sequencing and subsequently also for the region-specific analyses, whereas the other six
383 colonies (colonies queenright D-F and queenless D-F) were used for the region-specific
384 analyses only, which served to validate and confirm the reproducibility of the methylome
385 results. The brains and ovaries of the workers were dissected as described elsewhere (53, 65).
386

387 **Whole Genome Bisulfite Sequencing (WGBS) and Amplicon Sequencing**

388 We first performed Whole Genome Bisulfite Sequencing (WGBS) to identify which genes
389 were differentially methylated in the samples from the first three queenright and queenless
390 colonies (colonies A-C). For each sample we pooled either eight brains or 20 pairs of non-

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

391 activated ovaries. Pooling was necessary to obtain a sufficient DNA yield for high-throughput
392 sequencing. This resulted in a total of 12 samples, represented by six brain and six ovary
393 samples for each of the two social contexts (three queenright and three queenless). DNA was
394 extracted with the DNeasy Blood and Tissue Kit^{TM} (Qiagen) and quantified with a Qubit 2.0
395 Fluorometer system (Invitrogen). Before sequencing, 0.01% (w/w) of unmethylated Lambda
396 DNA (Promega) was added to each sample to be later used for calculating the bisulfite
397 conversation efficiency. For WGBS, the DNA samples were sent to the Beijing Genomics
398 Institute (China), where library construction, bisulfite treatment, and sequencing were
399 performed. Bisulfite treatment was done with the EZ-DNA methylation kit (Zymo Research),
400 and paired-end WGBS was performed on an Illumina NovaSeq platform. Each library was
401 sequenced twice in two separate lanes for high coverage. In the bioinformatics analysis, the
402 data from the two lanes were then merged, as we did not detect relevant differences between
403 the two runs. Data on coverage for each sample are given in Supplemental Table S1.

404 After processing the WGBS data and identification of DMRs, we performed a paired-
405 end ultra-deep amplicon sequencing analysis (59, 66) for 13 of the DMRs revealed by WGBS,
406 using a total of six colony pairs, the original three pairs (source colonies A-C) plus three new
407 colony pairs (source colonies D-F). DNA extractions, bisulfite conversion, and Lambda DNA
408 spiking was performed as described above. Columns were eluted with 20 μL of ultrapure water,
409 and 1 μL of the eluted solution was used as template in the PCR assays. Bisulfite PCR primers
410 were designed to amplify fragments between 140-300 bp of the forward strand of each
411 differentially methylated region and the control Lambda spike (Supplemental Table S8).
412 Bisulfite-treated DNA was amplified using the KAPA HiFi Uracil^{+} Kit^{TM} (Roche). PCR assays
413 were set up with 5 μL of Kapa Master Mix, 0.3 μL of each primer (forward and reverse – 3
414 pmol/reaction), 1 μL of bisulfite-treated DNA and 3.4 μL of water for a total reaction volume
415 of 10 μL, and performed with the annealing temperatures of the respective primers

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

416 (Supplemental Table S8). For multiplexing of the samples into two libraries, Nextera barcodes
417 were added to the 5' ends of all primers. Amplicons from different primers were pooled to
418 generate a separate library for each sample (24 samples in total: queenright and queenless
419 brains and ovaries from each of the source colonies A-F). Samples were purified, and Nextera
420 paired-end libraries were constructed at the Australian Genome Research Facility (AGRF).
421 Libraries were constructed in duplicate for each sample and sequenced (150 bp paired end) in
422 a single flow cell of an Illumina MiSeq platform. Data from the two libraries were merged for
423 downstream analyses.
424
425 **Differential methylation analyses from WGBS data**
426 Quality of the raw data was assessed by FastQC 0.11.8
427 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and reads with quality scores <
428 20 were removed. Trimming was performed with TrimGalore 0.5.0
429 (www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with a stringency error of 2 bp.
430 Overall, we removed ~ 0.5% of the total sequencing reads after checking read quality and
431 trimming of the adaptors (Supplemental Table S1). The remaining reads were mapped onto the
432 honey bee reference genome assembly Amel_4.5 (54) using Bismark 0.16.1 (67) and Bowtie
433 2 2.3.5.1v (68). The honey bee genome assembly Amel_4.5 was the newest honey bee genome
434 version at the time the bioinformatic analyses were performed. Coverage varied between 27-
435 46 times across samples (Supplemental Table S1). Methylation calling was performed with
436 Bismark software. We used 10-times coverage as a threshold for adequate sequencing coverage
437 of each cytosine as in Herb et al. (2012), to make our data sets comparable. Methylation levels
438 were assessed by the C/T ratio of converted cytosines to unconverted bases (69). Significantly
439 methylated sites were identified using a binomial probability model that takes into account the
440 bisulfite conversion rate for each sample (Supplemental Table S1) as the probability of success,

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

441 followed by Bonferroni corrections at the 1% significance level using BWASP (34). We
442 removed methylated CpGs with > 500x coverage to avoid PCR-based bias in the analyses.
443 MethylKit (70) was used for differential methylation analysis. First, the honey bee
444 genome was partitioned into 200 bp sliding windows (step size = 100 bp). Only windows
445 containing at least four sufficiently covered CpGs (two in each strand) were analyzed. A
446 difference threshold of 10% in the methylation level between pairwise comparisons was
447 applied (Herb et al. 2012). A threshold >10 methylated cytosines (sum for all CpGs inside
448 given window) in at least one of the libraries was used to reduce methylome complexity. The
449 list of DMRs was then FDR-corrected and a q-value <0.01 was considered significant. A gene
450 containing at least one DMR was defined as a differentially methylated gene. Gene annotation
451 was performed with Homer 4.9.1 software (71). Analyses were performed in the *R* environment
452 (R Core Team 2018). Hierarchical clustering distances between methylome samples were
453 determined with the R package "pvclust" (73) .
454

455 **Differential methylation analysis of amplicon sequencing data**

456 Reads were checked for quality using FastQC 0.11.8
457 (www.bioinformatics.babraham.ac.uk/projects/fastqc), followed by trimming of adaptors and
458 removal of low quality reads (Phred score < 20) using Trimmomatic (74). Between 85-90% of
459 the reads were retained for each library. Bissulfite-converted DNA sequences of the 13
460 amplicon regions of interest and the Lambda control sequence were used as templates to
461 generate a Bowtie2 index prior to alignment. Data from each of the two duplicate libraries for
462 each sample were aligned with Bowtie2 2.3.5.1v (68) using paired-end default parameters and
463 then converted to BAM files with Samtools (75). BAM files were imported into Geneious
464 software 10.2.4 (76), and alignments for each amplicon were manually checked for each
465 sample. C-to-T variant frequencies were calculated using the 'Find variant' function for all CG

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

466 sites, with a minimum coverage of 50 and a minimum variant frequency of 2%. Additional
467 SNP variants that were still visible after bisulfite conversion, such as G-to-A polymorphisms,
468 were also recorded. The overall cytosine methylation frequency was determined for each of the
469 four treatment groups (QR brain, QL brain queenless, QR ovary, QL ovary) for all of the 13
470 amplicons by dividing the total amount of C (methylated cytosines) per the total amount of
471 C+T (total amount of methylated and unmethylated cytosines). The results are presented as
472 percentages, and when appropriate, regions were compared in relation to their colony of origin,
473 social context, and tissue type.
474

475 **Gene expression analysis**

476 Each sample consisted of four pairs of non-activated ovaries or one brain (n = 8 per source
477 colony and social context combination). Ovaries needed to be pooled to obtain a sufficient
478 amount of RNA. Brains and ovaries were macerated in TRIzol (Invitrogen). Total RNA was
479 extracted using the Direct-zol™ RNA™ Miniprep kit (Zymo Research) according to the
480 manufacturer's instructions. Samples were treated with Turbo DNase™ (Thermo-Fisher
481 Scientific), and RNA concentrations were determined using a Qubit 2.0 Fluorometer system
482 (Invitrogen). RNA samples were diluted with ultrapure water to a final concentration of 40
483 ng/µL (brain) and 15 ng/µL (ovary). We used 142.5 ng of ovary RNA and 600 ng of brain
484 RNA to synthesize cDNA using the SuperScript™ III Reverse Transcriptase Kit (Invitrogen)
485 with Oligo(dT) primer (Invitrogen). Ovary cDNA was diluted to 2 ng/µL due to its lower
486 concentration, while brain cDNA was diluted to 5 ng/µL in ultrapure water.

487 Relative expression was determined by RT-qPCR assays for total of 20 genes, 12 being
488 hypermethylated or hypomethylated in one source colony, four genes whose methylation
489 patterns differ between tissues and four *Dnmt* genes. Assays were set up with 2.5 µL
490 SsoAdvanced™ Universal SYBR® Green Supermix (Bio Rad), 1.25 pmol of each primer, 1

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

491 μL diluted cDNA in a total volume of 5 μL using a CFX384 Real-Time System (Bio-Rad). For
492 each sample we conducted three technical replicates and used their mean as the data point.
493 Cycle conditions were as follows: 95 °C for 10 min followed by 40 cycles of 95 °C for 10 s,
494 annealing temperature (Supplemental Table 8) for 10 s and 72 °C for 15 s. At the end of the
495 PCR cycles a melting curve analysis was run to confirm a single amplification peak. Relative
496 gene expression analyses were performed using the a formula that accounts for primer's
497 efficiency: $E^{CqMin-CqSample}$ – where "E" (Table S3) is the efficiency of primers, "CqMin" is the
498 lowest Cq value for a given gene and "CqSample" is the Cq of that sample (65, 77, 78). Then,
499 two reference genes (*Rp49* [also known as *Rpl32*] and *Ef1α*) were used to normalize the
500 expression levels of the target genes. These control genes have been previously validated for
501 honey bee quantitative PCR analyses (79) and were stable in our analysis according to the
502 BestKeeper software (80). Primer efficiencies (Supplemental Table 8) were calculated based
503 on an amplification curve of 10 points obtained through serial dilution of mixed cDNA
504 samples. The list of primers used is given in Supplemental Table S8. Specificity of the
505 respective amplification products was validated by Sanger sequencing (Macrogen, South
506 Korea).
507
508 **Statistical analysis**
509 Gene expression levels were analyzed as the dependent variable using a Generalized Linear
510 Mixed Models (GLMM) with 'colony' as random effect, and 'social context' and 'tissue' as
511 fixed effects. We used a log link function to all genes to approximate gene expression data to
512 a Gaussian distribution, which was checked by analyzing the residuals' distribution.
513 Alternative link functions and data transformations were applied as necessary (see
514 Supplemental Table S6). Given that the 'social environment' (presence/absence of the queen)
515 might influence gene expression in opposite directions in different tissues, as seen for Dnmt's

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

516 expression, we performed Tukey's post-hoc tests for all gene/tissue combinations. To
517 investigate whether colony of origin influenced gene expression, 'colony' and 'tissue' were
518 treated as fixed effects and 'social environment' as a random effect. Tukey's post-hoc tests
519 were performed to identify differences in gene expression between individuals from different
520 colonies (Supplemental Table S7). The Pearson's correlation coefficient between gene
521 expression and differential methylation was calculated with a two-tailed test of significance.
522 Statistical testing was performed in *R* (R Core Team, 2018) using the packages lme4, car and
523 lsmeans, or in the GraphPad Prism 7 statistics package. For all analyses, a p-value < 0.05 was
524 considered significant.
525
526 **Data access**
527 All raw sequencing data generated in this study have been submitted to the NCBI Sequence
528 Read Archive under accession number PRJNA714749.
529
530 **Acknowledgments**
531 This work was supported by grants from São Paulo Research Foundation (FAPESP
532 2016/15881-0 and 2017/09269-3 to CAM; and 2017/09128-0 to KH), the Brazilian National
533 Council for Scientific and Technological Development (CNPq 403646/20162 and
534 303401/2014-1 to KH) and Australian Research Council (DP180101696 to BO and A. Zayed).
535
536
537
538
539
540

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

541 **References**

542 1. C. D. Allis, T. Jenuwein, The molecular hallmarks of epigenetic control. *Nat. Rev.*
543 *Genet.* **17**, 487–500 (2016).
544 2. G. L. Gabor Miklos, R. Maleszka, Epigenomic communication systems in humans and
545 honey bees: from molecules to behavior. *Horm Behav* **59**, 399–406 (2011).
546 3. S. Takayama, *et al.*, Genome methylation in *D. melanogaster* is found at specific short
547 motifs and is independent of DNMT2 activity. *Genome Res* **24**, 821–830 (2014).
548 4. Y. Wang, *et al.*, Functional CpG methylation system in a social insect. *Science* **314**,
549 645–647 (2006).
550 5. C. Honeybee Genome Sequencing, Insights into social insects from the genome of the
551 honeybee *Apis mellifera*. *Nature* **443**, 931–949 (2006).
552 6. F. Lyko, *et al.*, The honey bee epigenomes: differential methylation of brain DNA in
553 queens and workers. *PLoS Biol.* **8**, e1000506 (2010).
554 7. R. Bonasio, *et al.*, Genome-wide and caste-specific DNA methylomes of the ants
555 *Camponotus floridanus* and *Harpegnathos saltator*. *Curr. Biol.* **22**, 1755–1764 (2012).
556 8. S. Shukla, *et al.*, CTCF-promoted RNA polymerase II pausing links DNA methylation
557 to splicing. *Nature* **479**, 74–79 (2011).
558 9. H. Li-Byarlay, *et al.*, RNA interference knockdown of *DNA methyl-transferase 3* affects
559 gene alternative splicing in the honey bee. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12750–
560 12755 (2013).
561 10. B. G. Hunt, K. M. Glastad, S. V. Yi, M. A. D. Goodisman, The function of intragenic
562 DNA methylation: Insights from insect epigenomes. *Integr. Comp. Biol.* **53**, 319–328
563 (2013).
564 11. H. Yan, *et al.*, DNA Methylation in social insects: how epigenetics can control behavior
565 and longevity. *Annu. Rev. Entomol.* **60**, 435–452 (2015).
566 12. H. Yan, *et al.*, Eusocial insects as emerging models for behavioural epigenetics. *Nat.*
567 *Rev. Genet.* **15**, 677–688 (2014).
568 13. S. Foret, *et al.*, DNA methylation dynamics, metabolic fluxes, gene splicing, and
569 alternative phenotypes in honey bees. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 4968–4973
570 (2012).
571 14. S. Alvarado, R. Rajakumar, E. Abouheif, M. Szyf, Epigenetic variation in the *Egfr* gene
572 generates quantitative variation in a complex trait in ants. *Nat. Commun.* **6**, 6513 (2015).
573 15. B. R. Herb, *et al.*, Reversible switching between epigenetic states in honeybee
574 behavioral subcastes. *Nat. Neurosci.* **15**, 1371–1373 (2012).
575 16. R. Kucharski, J. Maleszka, S. Foret, R. Maleszka, Nutritional control of reproductive
576 status in honeybees via DNA methylation. *Science* **319**, 1827–1830 (2008).
577 17. B. R. Herb, M. S. Shook, C. J. Fields, G. E. Robinson, Defense against territorial
578 intrusion is associated with DNA methylation changes in the honey bee brain. *BMC*
579 *Genomics* **19**, 216 (2018).
580 18. G. A. Lockett, P. Helliwell, R. Maleszka, Involvement of DNA methylation in memory
581 processing in the honey bee. *Neuroreport* **21**, 812–816 (2010).
582 19. K. M. Glastad, B. G. Hunt, S. V. Yi, M. A. D. Goodisman, Epigenetic inheritance and
583 genome regulation: Is DNA methylation linked to ploidy in haplodiploid insects? *Proc.*
584 *R. Soc. B Biol. Sci.* **281**, 20140411 (2014).
585 20. S. D. Biergans, C. Giovanni Galizia, J. Reinhard, C. Claudianos, *Dnmts* and *Tet* target
586 memory-associated genes after appetitive olfactory training in honey bees. *Sci. Rep.* **5**,
587 16223 (2015).
588 21. S. D. Biergans, C. Claudianos, J. Reinhard, C. G. Galizia, DNA methylation mediates
589 neural processing after odor learning in the honeybee. *Sci. Rep.* **7**, 43635 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

590 22. L. Holman, K. Trontti, H. Helantera, Queen pheromones modulate DNA
591 methyltransferase activity in bee and ant workers. *Biol Lett* **12**, 20151038 (2016).
592 23. S. Kay, D. Skowronski, B. G. Hunt, Developmental DNA methyltransferase expression
593 in the fire ant *Solenopsis invicta*. *Insect Sci.* **25**, 57–65 (2016).
594 24. C. A. M. Cardoso-Júnior, K. R. Guidugli-Lazzarini, K. Hartfelder, DNA methylation
595 affects the lifespan of honey bee (*Apis mellifera* L.) workers – Evidence for a regulatory
596 module that involves vitellogenin expression but is independent of juvenile hormone
597 function. *Insect Biochem. Mol. Biol.* **92**, 21–29 (2018).
598 25. C. A. M. Cardoso-Junior, M. Eyer, B. Dainat, K. Hartfelder, V. Dietemann, Social
599 context influences the expression of DNA methyltransferase genes in the honeybee. *Sci.*
600 *Rep.* **8**, 11076 (2018).
601 26. K. Flores, *et al.*, Genome-wide association between DNA methylation and alternative
602 splicing in an invertebrate. *BMC Genomics* **13**, 480 (2012).
603 27. P. Cingolani, *et al.*, Intronic non-CG DNA hydroxymethylation and alternative mRNA
604 splicing in honey bees. *BMC Genomics* **14**, 666 (2013).
605 28. K. M. Glastad, K. Gokhale, J. Liebig, M. A. D. Goodisman, The caste- and sex-specific
606 DNA methylome of the termite *Zootermopsis nevadensis*. *Sci. Rep.* **6**, 37110 (2016).
607 29. K. Kannan, M. Shook, Y. Li, G. E. Robinson, J. Ma, Comparative analysis of brain and
608 fat body gene splicing patterns in the honey bee, *Apis mellifera*. *G3 Genes, Genomes,*
609 *Genet.* **9**, 1055–1063 (2019).
610 30. E. J. Remnant, *et al.*, Parent-of-origin effects on genome-wide DNA methylation in the
611 Cape honey bee (*Apis mellifera capensis*) may be confounded by allele-specific
612 methylation. *BMC Genomics* **17** (2016).
613 31. R. Libbrecht, P. R. Oxley, L. Keller, D. J. C. Kronauer, Robust DNA methylation in the
614 clonal raider ant brain. *Curr. Biol.* **26**, 391–395 (2016).
615 32. X. Wu, *et al.*, Lineage and parent-of-origin effects in DNA methylation of honey bees
616 (*Apis mellifera*) revealed by reciprocal crosses and whole-genome bisulfite sequencing.
617 *Genome Biol. Evol.* **12**, 1482–1492 (2020).
618 33. B. P. Oldroyd, B. Yagound, The role of epigenetics in the evolution of caste in insect
619 societies. *Philos. Trans. R. Soc. B-Biological Sci.* **In press** (2021).
620 34. D. S. Standage, *et al.*, Genome, transcriptome and methylome sequencing of a
621 primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social
622 insect. *Mol. Ecol.* **25**, 1769–1784 (2016).
623 35. L. Wedd, R. Kucharski, R. Maleszka, Differentially methylated obligatory epialleles
624 modulate context-dependent LAM gene expression in the honeybee *Apis mellifera*.
625 *Epigenetics* **11**, 1–10 (2016).
626 36. Z. Lonsdale, *et al.*, Allele specific expression and methylation in the bumblebee,
627 *Bombus terrestris*. *PeerJ* **5**, e3798 (2017).
628 37. A. J. Bewick, *et al.*, *Dnmt1* is essential for egg production and embryo viability in the
629 large milkweed bug, *Oncopeltus fasciatus*. *Epigenetics and Chromatin* **12**, 6 (2019).
630 38. K. D. Harris, J. P. B. Lloyd, K. Domb, D. Zilberman, A. Zemach, DNA methylation is
631 maintained with high fidelity in the honey bee germline and exhibits global non-
632 functional fluctuations during somatic development. *Epigenetics and Chromatin* **12**, 62
633 (2019).
634 39. B. Yagound, N. M. A. Smith, G. Buchmann, B. P. Oldroyd, E. J. Remnant, Unique DNA
635 methylation profiles are associated with cis-variation in honey bees. *Genome Biol. Evol.*
636 **11**, 2517–2530 (2019).
637 40. B. Yagound, E. J. Remnant, G. Buchmann, B. P. Oldroyd, Intergenerational transfer of
638 DNA methylation marks in the honey bee. *Proc. Natl. Acad. Sci. U. S. A.*, Ahead of print
639 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

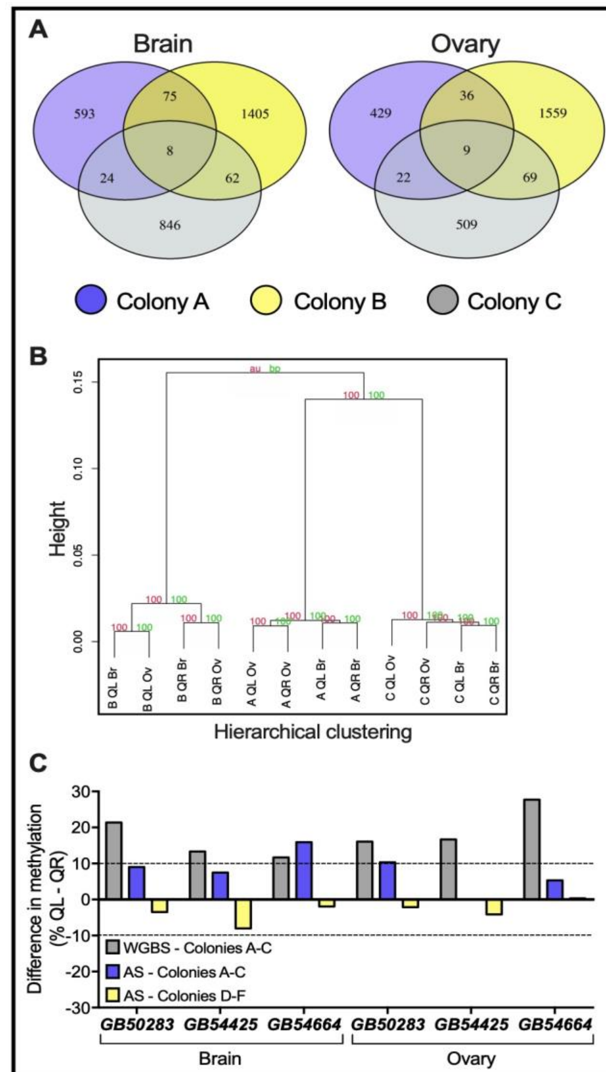
640 41. H. Marshall, A. R. C. Jones, Z. N. Lonsdale, E. B. Mallon, Bumblebee workers show
641 differences in allele-specific DNA methylation and allele-specific expression. *Genome*
642 *Biol. Evol.* **12**, 1471–1481 (2020).
643 42. X. Wang, *et al.*, Function and evolution of DNA methylation in *Nasonia vitripennis*.
644 *PLoS Genet.* **9**, e1003872 (2013).
645 43. S. V. Arsenault, B. G. Hunt, S. M. Rehan, The effect of maternal care on gene expression
646 and DNA methylation in a subsocial bee. *Nat. Commun.* **9**, 3468 (2018).
647 44. H. Marshall, Z. N. Lonsdale, E. B. Mallon, Methylation and gene expression differences
648 between reproductive and sterile bumblebee workers. *Evol. Lett.* **3**, 485–499 (2019).
649 45. P. R. Oxley, *et al.*, The genome of the clonal raider ant *Cerapachys Biroi*. *Curr. Biol.*
650 **24**, 451–458 (2014).
651 46. T. Pankiw, M. L. Winston, Z. Huang, G. E. Robinson, Queen mandibular gland
652 pheromone influences worker honey bee (*Apis mellifera* L.) foraging ontogeny and
653 juvenile hormone titers. *J. Insect Physiol.* **44**, 685–692 (1998).
654 47. M. L. Winston, *The biology of the honey bee.* (Harvard University Press, 1987).
655 48. S. E. Hoover, C. I. Keeling, M. L. Winston, K. N. Slessor, The effect of queen
656 pheromones on worker honey bee ovary development. *Naturwissenschaften* **90**, 477–
657 480 (2003).
658 49. C. M. Grozinger, N. M. Sharabash, C. W. Whitfield, G. E. Robinson, Pheromone-
659 mediated gene expression in the honey bee brain. *Proc. Natl. Acad. Sci. U. S. A.* **100**
660 **Suppl**, 14519–14525 (2003).
661 50. D. Niu, *et al.*, Transcriptome comparison between inactivated and activated ovaries of
662 the honey bee *Apis mellifera* L. *Insect Mol. Biol.* **23**, 668–681 (2014).
663 51. L. Holman, H. Helanterä, K. Trontti, A. S. Mikheyev, Comparative transcriptomics of
664 social insect queen pheromones. *Nat. Commun.* **10**, 1593 (2019).
665 52. E. J. Duncan, M. P. Leask, P. K. Dearden, Genome architecture facilitates phenotypic
666 plasticity in the honeybee (*Apis mellifera*). *Mol. Biol. Evol.* **37**, 1964–1978 (2020).
667 53. C. A. M. Cardoso-Júnior, I. Ronai, K. Hartfelder, B. P. Oldroyd, Queen pheromone
668 modulates the expression of epigenetic modifier genes in the brain of honey bee workers.
669 *Biol. Lett.* **16**, 20200440 (2020).
670 54. C. G. Elsik, *et al.*, Finding the missing honey bee genes: Lessons learned from a genome
671 upgrade. *BMC Genomics* **15**, 86 (2014).
672 55. R. A. Drewell, *et al.*, The dynamic DNA methylation cycle from egg to sperm in the
673 honey bee *Apis mellifera*. *Development* **141**, 2702–2711 (2014).
674 56. B. Li, *et al.*, Identification and caste-dependent expression patterns of DNA methylation
675 associated genes in *Bombus terrestris*. *Sci Rep* **8**, 2332 (2018).
676 57. Y. Wang, *et al.*, Regulation of behaviorally associated gene networks in worker honey
677 bee ovaries. *J. Exp. Biol.* **215**, 124–134 (2012).
678 58. R. Libbrecht, P. R. Oxley, D. J. C. Kronauer, Clonal raider ant brain transcriptomics
679 identifies candidate molecular mechanisms for reproductive division of labor. *BMC*
680 *Biol.* **16**, 89 (2018).
681 59. R. Kucharski, R. Maleszka, Exploring DNA methylation diversity in the honey bee brain
682 by ultra-deep amplicon sequencing. *Epigenomes* **4**, 10 (2020).
683 60. A. Zemach, I. E. McDaniel, P. Silva, D. Zilberman, Genome-wide evolutionary analysis
684 of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010).
685 61. N. K. E. Schulz, *et al.*, *Dnmt1* has an essential function despite the absence of CpG DNA
686 methylation in the red flour beetle *Tribolium castaneum*. *Sci. Rep.* **8**, 16462 (2018).
687 62. J. T. Washington, *et al.*, The essential role of *Dnmt1* in gametogenesis in the large
688 milkweed bug *Oncopeltus fasciatus*. *BioRxiv Prepr.* (2020).
689 63. A. U. Amukamara, *et al.*, More than DNA methylation: Does pleiotropy drive the

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

690 complex pattern of evolution of *Dnmt1*? *Front. Ecol. Evol.* **8**, 4 (2020).
691 64. J. Harbo, "Propagation and instrumental insemination." in *Rinderer T. E. (Ed.) Bee*
692 *Genetics and Breeding*, Rinderer T, (Academic Press, 1986), pp. 361–389.
693 65. C. A. M. Cardoso-Júnior, B. Oldroyd, I. Ronai, Vitellogenin expression in the ovaries
694 of adult honeybee workers provides insights into the evolution of reproductive and social
695 traits. *Insect Mol Biol* **Epub** (2021).
696 66. R. Kucharski, S. Foret, R. Maleszka, EGFR gene methylation is not involved in
697 Royalactin controlled phenotypic polymorphism in honey bees. *Sci. Rep.* **5**, 14070
698 (2015).
699 67. F. Krueger, S. R. Andrews, Bismark: A flexible aligner and methylation caller for
700 Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
701 68. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
702 **9**, 357–359 (2012).
703 69. K. M. Glastad, *et al.*, Variation in DNA methylation is not consistently reflected by
704 sociality in Hymenoptera. *Genome Biol. Evol.* **9**, 1687–1698 (2017).
705 70. A. Akalin, *et al.*, MethylKit: a comprehensive R package for the analysis of genome-
706 wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).
707 71. S. Heinz, *et al.*, Simple combinations of lineage-determining transcription factors prime
708 cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**,
709 576–589 (2010).
710 72. R. C. Team, R: A language and environment for statistical computing. R foundation for
711 statistical computing, Vienna, Austria. URL: http://www.R-project.org/ (2018).
712 73. R. Suzuki, H. Shimodaira, Pvclust: An R package for assessing the uncertainty in
713 hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).
714 74. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina
715 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
716 75. H. Li, *et al.*, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**,
717 2078–2079 (2009).
718 76. M. Kearse, *et al.*, Geneious basic: An integrated and extendable desktop software
719 platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–
720 1649 (2012).
721 77. J. Hellemans, G. Mortier, A. De Paepe, F. Speleman, J. Vandesompele, qBase relative
722 quantification framework and software for management and automated analysis of real-
723 time quantitative PCR data. *Genome Biol.* **8**, R19 (2007).
724 78. R. M. Brito, M. McHale, B. P. Oldroyd, Expression of genes related to reproduction and
725 pollen foraging in honey bees (*Apis mellifera*) narcotized with carbon dioxide. *Insect*
726 *Mol. Biol.* **19**, 451–461 (2010).
727 79. A. P. Lourenço, A. Mackert, A. S. Cristino, Z. L. Simoes, Validation of reference genes
728 for gene expression studies in the honey bee, *Apis mellifera*, by quantitative real-time
729 RT-PCR. *Apidologie* **39**, 372–385 (2008).
730 80. M. W. Pfaffl, A. Tichopad, C. Prgomet, T. P. Neuvians, Determination of stable
731 housekeeping genes, differentially regulated target genes and sample integrity:
732 BestKeeper - Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–
733 515 (2004).
734
735
736
737
738
739

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

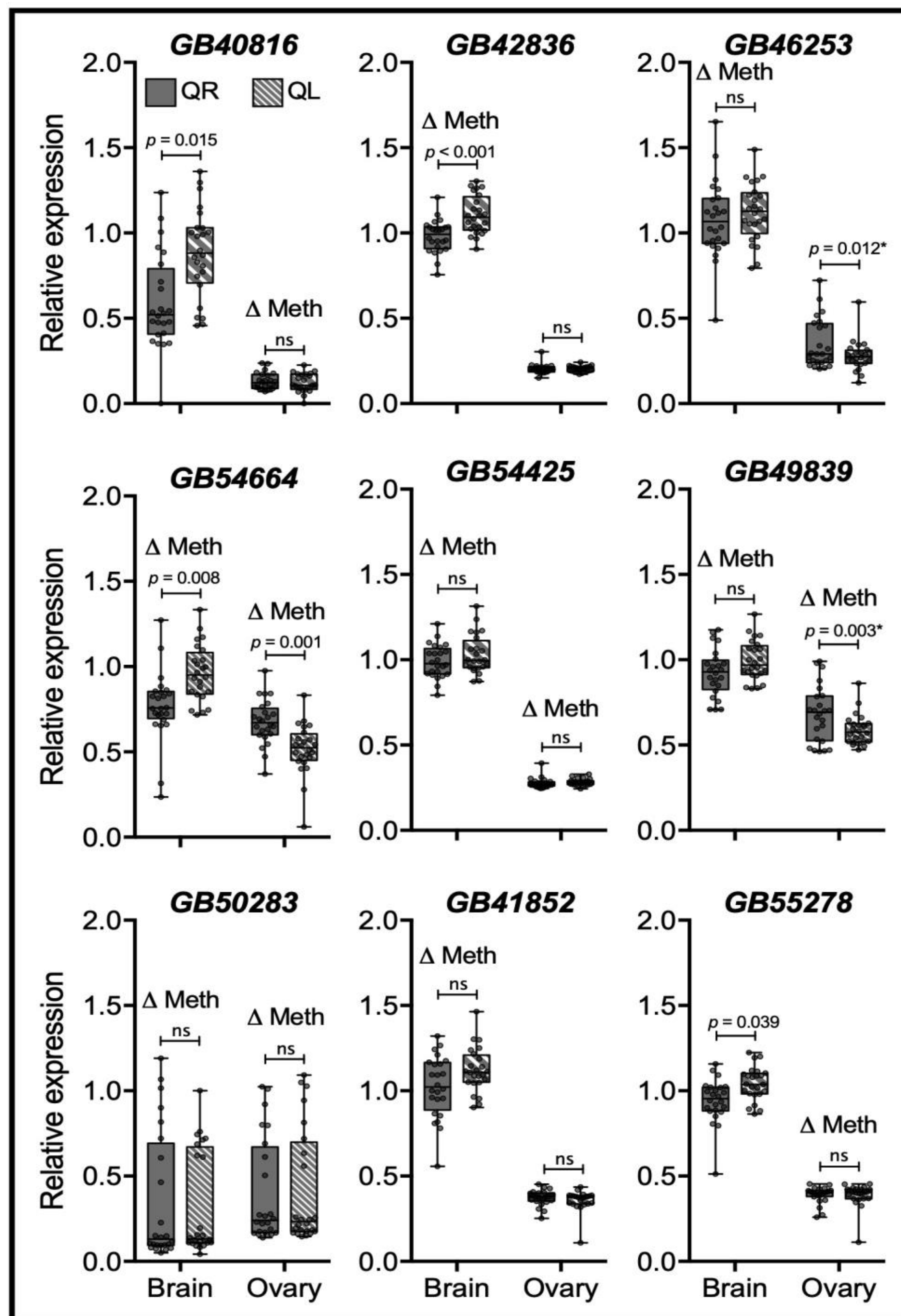
740 **Figures**



741

742 **Figure 1.** Social context does not drive significant alterations in the methylomes of honey bee
743 workers. (*A*) Number of DMRs comparing QR and QL methylomes across all three colonies
744 assessed for WGBS. Intersections show the number of DMRs shared by more than one colony.
745 (*B*) Hierarchical clustering showing the correlation-based distances of methylation patterns
746 across all WGBS samples; the numbers show high support values for both approximately
747 unbiased (au) and bootstrap probability (bp) statistics. Samples are identified by their colony
748 origin (A, B or C), social context (QR or QL) and tissue (Br – Brain, Ov – Ovary), respectively.
749 (*C*) Methylation frequency (%QL - %QR) of three top-ranked DMRs identified in the brain
750 and ovaries for the six colonies. The gray bars represent the frequency of methylated sites
751 between QL and QR bees obtained from whole-genome bisulfite sequencing (WGBS), and the
752 blue bars for the amplicon bisulfite sequencing (AS) for the original colonies A-C. The yellow
753 bars show the frequency of methylated sites (%QL - %QR) obtained from amplicon sequencing
754 of samples from three additional colony pairs (colonies D-F). The dashed lines represent the
755 10% threshold for considering a region as differentially methylated. Data from individual
756 cytosines and coverage can be found in Supplemental Fig. S2 and Supplemental Table S5,
757 respectively.

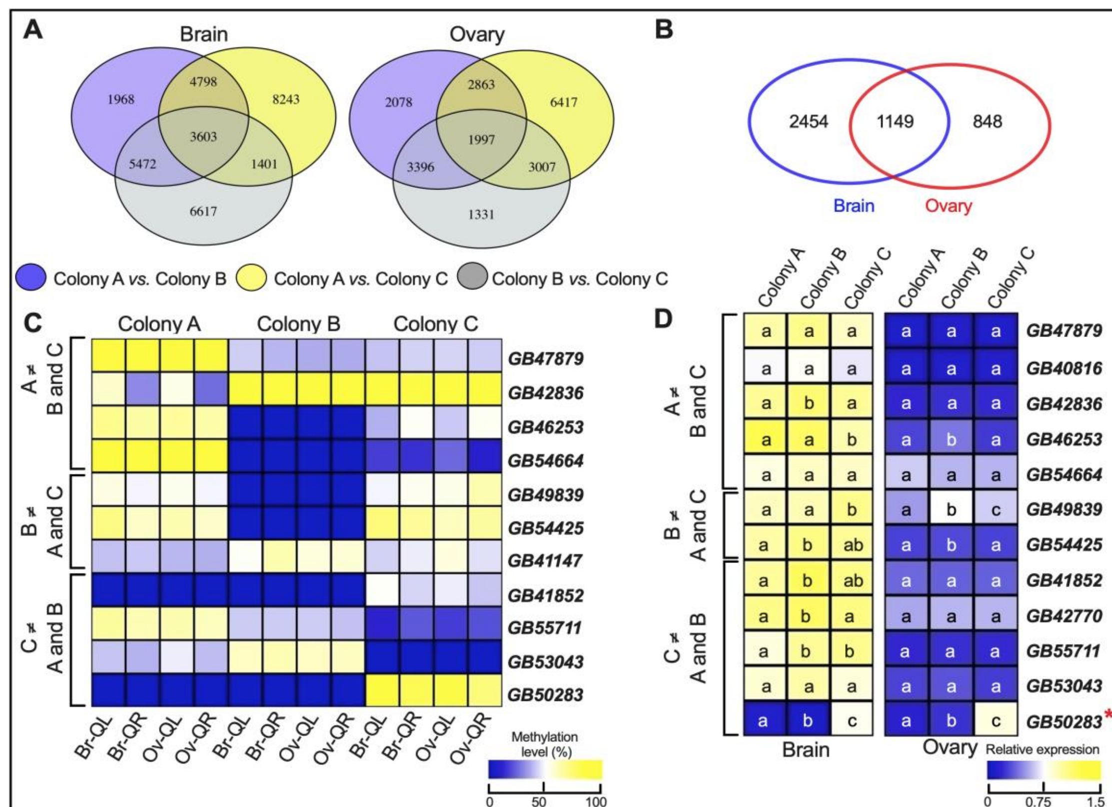
bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



758

759 **Figure 2.** Relative gene expression analysis of nine differentially methylated genes in the brain
760 and ovaries of workers from queenright (QR) and queenless (QL) colonies. "Δ Meth" shows
761 the tissue where DNA methylation was affected by social context (Supplemental Tables S3,
762 S4). Note that only *GB42836* (brain), *GB49839* (ovary) and *GB54664* (both tissues) were both
763 differently methylated and expressed in the QR *vs.* QL contrast. Each box shows the
764 interquartile range (25th-75th percentiles) and the median (line), while whiskers represent the
765 minimum and maximum values. Gray dots inside boxes represent individual samples (n = 3
766 colonies, 8 samples per colony, 24 in total). Statistical information: GLMM test of differences
767 between means with Tukey correction for multiple pairwise comparisons, an asterisk (*)
768 represents the genes with significant difference in the post-hoc test but non-significant for the
769 main 'social context' effect, 'ns' indicates $p > 0.05$, see Supplemental Table S6 for further
770 information.

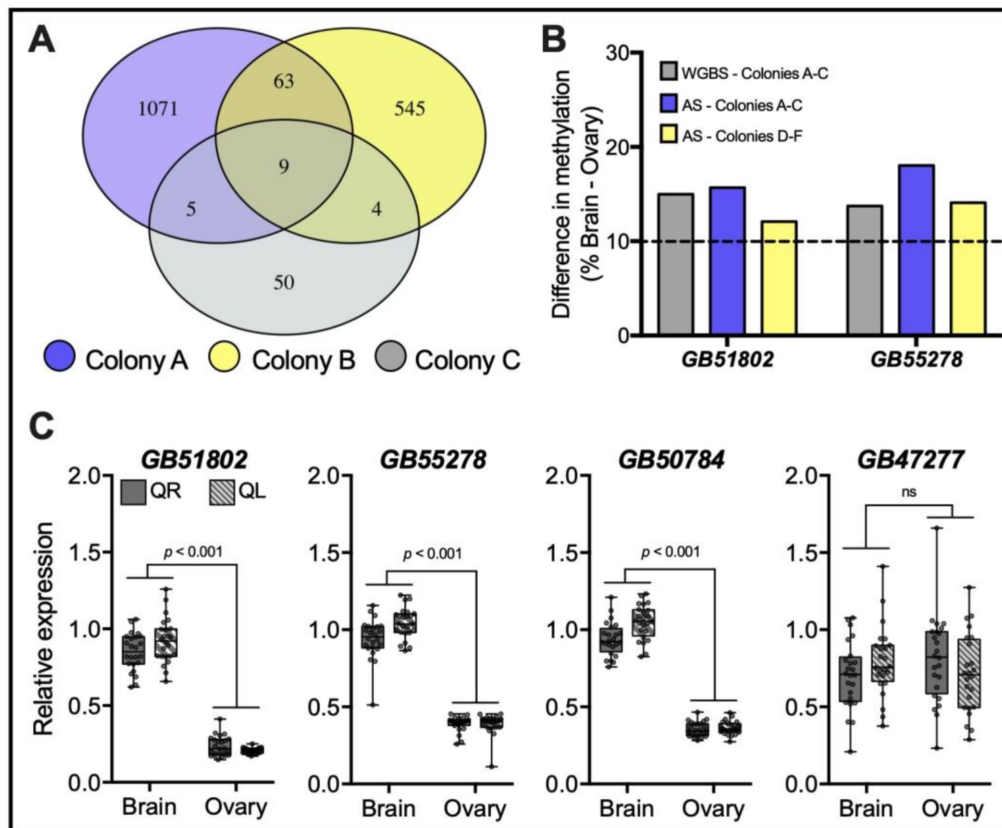
bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



771

772 **Figure 3.** Differences in the methylomes and gene expression levels between source colonies.
773 (*A*) Venn diagrams for the number of DMRs in the brain and ovaries comparing three different
774 colonies. Numbers inside the intersections represent the regions that show differential
775 methylation of at least 10% in one source colony compared to the other two source colonies.
776 For example, the intersect between blue and yellow region (e.g., 4798 DMRs for the brain or
777 2863 DMRs for the ovaries) represents the regions from colony A that differ in methylation
778 level by at least 10% from the same regions in colony B and C. The intersection of all three
779 comparisons, i.e. the central intersect, represents the regions with three degrees of DNA
780 methylation levels (for example, 0% methylation level for colony A, 20% for colony B and
781 70% for colony C). (*B*) Number of DMRs from central intersects displayed in Fig. 3A shared
782 by the brain and ovary tissues. (*C*) Heatmap showing colony-specific methylation validated by
783 high-throughput amplicon sequencing of 11 DMRs. For this analysis, we used the WGBS data
784 to select genes that had a specific pattern for one colony (hypermethylated or hypomethylated),
785 but differed from the patterns observed in the other two colonies. Methylation level for colonies
786 A-C are shown as follows: Br-QL – brain queenless; Br-QR – brain queenright; Ov-QL – ovary
787 queenless; Ov-QR – ovary queenright. Coverage of each region can be found in Supplemental
788 Table S5 and methylation level of individual CpG sites are displayed in Supplemental Figs.
789 S2, S6, S7. (*D*) Colony-specific gene expression of differentially methylated genes. Note that
790 only one gene, *GB50283* (red asterisk), had a methylation pattern (Fig. 3C) that seemed
791 correlated with its gene expression level. Note that the list of genes shown in Fig. 3D is the
792 same as in Fig. 3C, the exceptions being the gene *GB41147*, a differentially methylated gene
793 which relative expression was not examined, and the genes *GB40816* and *GB42770,* which
794 were differently methylated in the WGBS data but not validated by amplicon sequencing.
795 Different letters inside the heatmap boxes indicate statistical differences between colonies
796 (GLMM test, n = 16 samples per colony and tissue). Differential expression between tissues
797 was observed for all of the tested genes (Supplemental Table S7), with the exception of
798 *GB54664*.

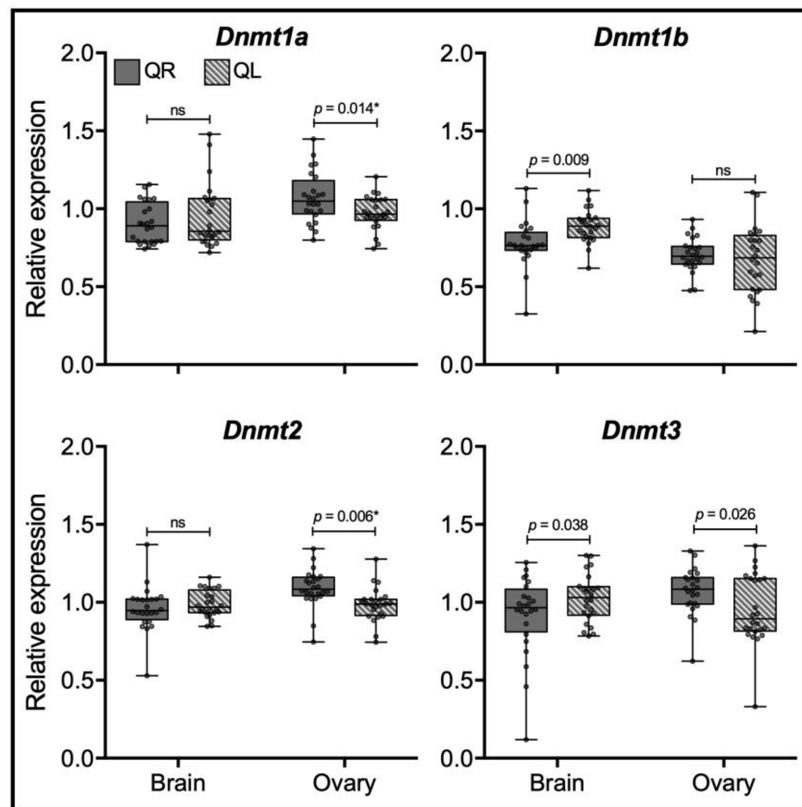
bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



799

800 **Figure 4.** Differences in the methylomes and gene expression levels between brain and ovary
801 tissue. (*A*) Venn diagram shows the number of DMRs in the contrast between brain vs. ovary
802 methylomes in all three colonies (A-C). These nine DMRs are associated with four genes. (*B*)
803 Difference in methylation for two of the four genes seen in the WGBS data (gray bars) and
804 validation by amplicon sequencing in colonies A-C (blue bars) and D-F (yellow bars). The
805 dashed line represents the 10% threshold for considering a region as differentially methylated.
806 Coverage of each region can be found in Supplemental Table S5 and methylation level of
807 individual CpG sites in Supplemental Fig. S7. (*C*) Transcript levels of the four differentially
808 methylated genes quantified by quantitative PCR in the brain and ovary of young honey bee
809 workers. Note that *GB55278* was also found to be differentially methylated between the ovaries
810 of QR and QL workers (Fig. 2 and Supplemental Table 4). Each box shows the interquartile
811 range (25th-75th percentiles) and the median (line), while whiskers represent the minimum and
812 maximum values. Sample size and statistical analysis are the same as in Fig. 2 and "ns"
813 indicates $p > 0.05$.
814
815
816
817
818
819
820
821
822

bioRxiv preprint doi: https://doi.org/10.1101/2021.03.12.435154; this version posted March 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



823

824 **Figure 5**. Relative expression of DNMTs encoding genes in honey bee workers kept in two
825 different social conditions, queen presence (QR) or queen absence (QL). Transcript levels were
826 assessed by quantitative PCR in the brain and ovary of young honey bee workers. Each box
827 shows the interquartile range (25th-75th percentiles) and the median (line), while whiskers
828 represent the minimum and maximum values. Sample size and statistical analysis are the same
829 as in Fig. 2; "*" represents the genes with a significant difference in the post-hoc test but non-
830 significant for the main 'social context' effect, "ns" indicates $p > 0.05$.