Parent of origin DNA methylation as a potential mechanism for genomic imprinting in bees.

Marshall, H.[1*], Nicholas, M.T.[1], van Zweden, J.S.[2], Wäckers, F.[3,4], Ross, L.[5], Wenseleers, T.[2], and Mallon, E.B.[1*]

[*]Corresponding Authors
[1]Department of Genetics and Genome Biology, The University of Leicester, UK.
[2]Laboratory of Socioecology and Social Evolution, Department of Biology, KU Leuven, Belgium.
[3]Biobest Belgium N.V., Westerlo, Belgium.
[4]The Lancaster Environmental Centre, University of Lancaster, UK.
[5]The Institute for Evolutionary Biology, University of Edinburgh, UK.

Friday 17th December, 2021

Contact: hollie_marshall@hotmail.co.uk, ebm3@leicester.ac.uk.

***Keywords*—** DNA methylation, reciprocal cross, genomic imprinting, social insect, bumblebee.

Abstract

Genomic imprinting is defined as parent-of-origin allele-specific expression. In order for genes to be expressed in this manner an 'imprinting' mark must be present to distinguish the parental alleles within the genome. In mammals imprinted genes are primarily associated with DNA methylation. Genes exhibiting parent-of-origin expression have recently been identified in two species of Hymenoptera with functional DNA methylation systems; Apis mellifera and Bombus terrestris. We carried out whole genome bisulfite sequencing of parents and offspring from reciprocal crosses of two B. terrestris subspecies in order to identify parent-of-origin DNA methylation. We were unable to survey a large enough proportion of the genome to draw a conclusion on the presence of parent-of-origin DNA methylation however we were able to characterise the sex- and caste-specific methylomes of B. terrestris for the first time. We find males differ significantly to the two female castes, with differentially methylated genes involved in many histone modification related processes. We also analysed previously generated honeybee whole genome bisulfite data to see if genes previously identified as showing parent-of-origin DNA methylation in the honeybee show consistent allele-specific methylation in independent data sets. We have identified a core set of 12 genes in female castes which may be used for future experimental manipulation to explore the functional role of parent-of-origin DNA methylation in the honeybee. Finally, we have also identified allele-specific DNA methylation in honeybee male thorax tissue which suggests a role for DNA methylation in ploidy compensation in this species.

Introduction

Genomic imprinting is defined as parent-of-origin allele-specific expression (Rodrigues and Zilberman, 2015). Deviating from Mendel's Law of Dominance, the expression of imprinted genes is determined by the parental origin of each allele irrespective of the underlying genotype. In order for genes to be expressed in this manner, an epigenetic 'imprinting mark' must be present to distinguish the parental alleles within the genome. In mammals and flowering plants, imprinted genes are primarily associated with DNA methylation in areas of the genome known as imprinting control regions (ICRs) (Barlow and Bartolomei, 2014). The presence of DNA methylation, as an imprinting mark, has been associated with both allelic silencing as well as allelic expression (Drewell *et al.*, 2012; Barlow, 2011).

DNA methylation can serve to repress expression of a individual allele by preventing a transcription factor from binding to a promoter region (Drewell *et al.*, 2012). It is also thought DNA methylation can signal the recruitment of histone modifications which result in heterochromatin formation, silencing the expression of all alleles within a region (Barlow, 2011; Xu *et al.*, 2021). DNA methylation can also enable allele-specific expression of imprinted genes by silencing an 'imprinting control element', such as a long non-coding RNA, which, when expressed, would cause transcriptional silencing of nearby genes (Barlow, 2011). It is also worth noting whilst most identified imprinted genes in mammals have some association with DNA methylation there are cases where imprinting has been observed independently of DNA methylation and instead associated solely with a particular histone modification (Inoue *et al.*, 2017).

The function of DNA methylation in insects is largely unknown and thought to be variable based on the range of overall levels between taxonomic orders (Provataris *et al.*, 2018). However, multiple insect species have now been show to display parent-of-origin gene expression, including the mealybug *Planococcus citri* (de la Filia *et al.*, 2021) and two Hymenopteran species, the bumblebee, *Bombus terrestris* (Marshall *et al.*, 2020b) and the honeybee *Apis mellifera* (Kocher *et al.*, 2015;

Galbraith *et al.*, 2016; Smith *et al.*, 2020; Galbraith *et al.*, 2021). In Hymenoptera, DNA methylation has been associated with caste differences in various species (Lyko *et al.*, 2010; Bonasio *et al.*, 2012; Amarasinghe *et al.*, 2014; Glastad *et al.*, 2016). However, a casual link has yet to be established (Oldroyd and Yagound, 2021b). The leading theory for the evolution of genomic imprinting, Haig's kinship theory (Haig, 2000), predicts social insects should display imprinted genes. It is also thought genomic imprinting in social insects could contribute to caste differentiation (Matsuura *et al.*, 2018) and the evolution of sociality (Matsuura, 2019). Although no association between the level of sociality of a species and the level of DNA methylation has been found (Weiner *et al.*, 2013; Glastad *et al.*, 2017).

Insect DNA methylation, like mammalian DNA methylation, is generally found in a CpG context (referring to a cytosine base immediately followed by a guanine base) (Glastad *et al.*, 2014). It is found at lower levels, with <1% - 14% of CpGs being methylated, compared to mammals where around 70% of CpG sites are methylated (Bewick *et al.*, 2016; Feng *et al.*, 2010). Additionally DNA methylation in insects is generally located in gene bodies and associated with more highly expressed genes, such as housekeeping genes (Provataris *et al.*, 2018; Elango *et al.*, 2009; Foret *et al.*, 2009). This suggests if DNA methylation does play a role in imprinted genes in insects, it may not function the same as imprinted DNA methylation marks in mammals.

Here, we set out to explore how robust parent-of-origin DNA methylation is within the species of bees which display genes showing parent-of-origin expression, to investigate the hypothesis that DNA methylation can act as an epigenetic mark for genomic imprinting in social insects (Oldroyd and Yagound, 2021a). Firstly, we attempt to identify parent-of-origin DNA methylation in the eusocial bumblebee *Bombus terrestris*. There is some evidence that parent-of-origin DNA methylation exists in honeybees (Wu *et al.*, 2020), however this was found to not directly correlate with genes previously identified as showing parent-of-origin expression. Given that genes showing parent-of-origin expression appear to have evolved rapidly within Hymenoptera (Marshall *et al.*, 2020b) it is worth investigating this potential role of DNA methylation in other species. *B. terrestris* has previously

been shown to exhibit parent-of-origin gene expression of both maternally derived and paternally derived alleles (Marshall *et al.*, 2020b), *B. terrestris* possess a fully functional methylation system (Amarasinghe *et al.*, 2014; Liu *et al.*, 2018; Marshall *et al.*, 2019) and it displays allele-specific methylation at multiple loci throughout the genome (Lonsdale *et al.*, 2017; Marshall *et al.*, 2020a). Some of these allele-specific DNA methylation events may represent parent-of-origin imprinting marks.

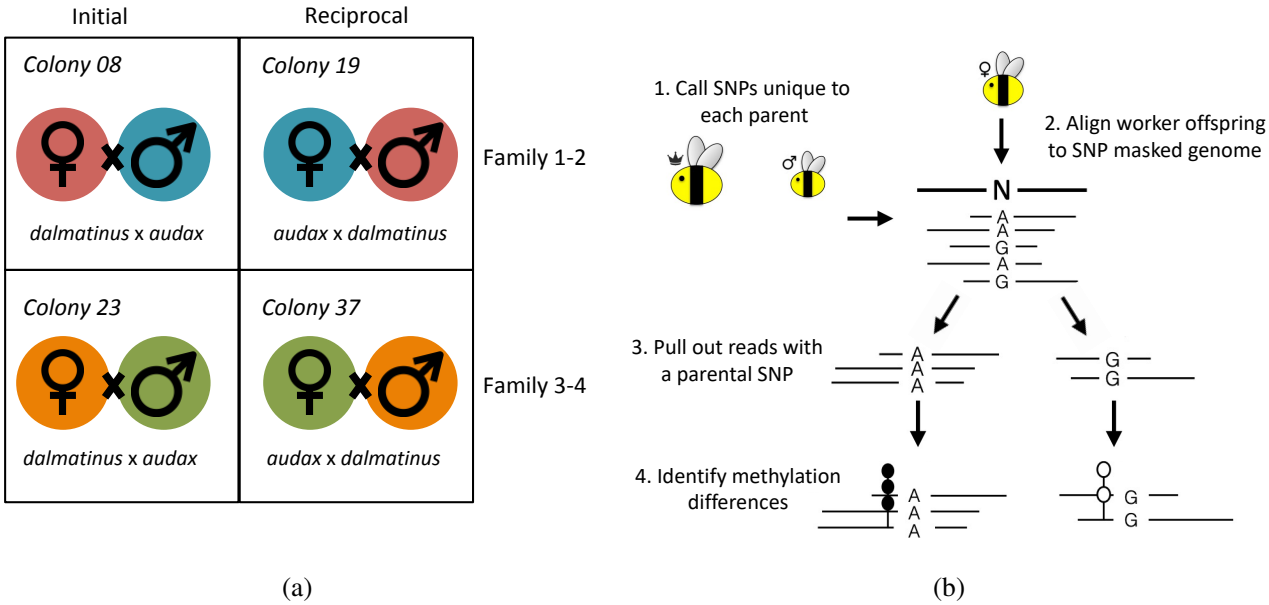
As part of this first component we have generated whole genome bisulfite sequencing (WGBS) libraries from head tissue of the parents and pooled worker offspring from two reciprocal crosses. Male *B. terrestris* are haploid meaning every worker daughter inherits the same copy of the paternal genome, whilst queens are diploid meaning there are two possible maternal alleles that can be present in the offspring. Whole genome re-sequencing (WGS) of the parents was used to identify SNPs unique to the father and homozygous SNPs unique to the queen mother to identify the parental origin of alleles within the offspring genome, allowing parent-of-origin DNA methylation to be identified. As the WGBS data generated here are also the first DNA methylation libraries for queens and males we have also explored the sex- and caste-specific methylome of *B. terrestris*.

In the second part of this study we explore how robust previously identified parent-of-origin DNA methylation in the honeybee is (Wu *et al.*, 2020). We have identified allele-specific DNA methylation in data from multiple recent honeybee studies to see if previously identified genes showing parent-of-origin DNA methylation appear to be present across independently generated data sets for both diploid female castes (queens and workers) and across different tissues. Finally, we also explore the potential presence of allele-specific DNA methylation in honeybee male somatic tissue, where some cells show varying levels of ploidy (Aron *et al.*, 2005), to see if female imprinted genes may hold an alternative role in haploid males.

Methods

Sample collection

Reciprocal crosses of *B. terrestris dalmatinus* (native to southern Europe) and *B. terrestris audax* (native to the UK) were carried out by Biobest, Leuven. Four successful colonies (one of each cross direction) from two genetically distinct 'families' (Fig. 1a) were housed at the University of Leuven and kept in 21°C with red light conditions, they were fed *ad libitum* with pollen and a sugar syrup. Callow workers were tagged with numbered disks in order to determine age. Worker reproductive status was confirmed by ovary dissection, ovaries were scored on a 0-4 scale as in Duchateau and Velthuis (1988), entire bodies were then stored at -80°C along with the original queen mothers and male fathers. Three reproductive workers, aged 16-17days, were selected from queenless conditions from each of the four colonies (supplementary 1.0.0, Fig.1a).



(a) (b)

Figure 1: (a) Graphic display of the family-wise reciprocal crosses carried out between *Bombus terrestris audax* and *Bombus terrestris dalmatinus*. Each colour refers to related individuals, i.e. the queen from colony 08 is the sister of the male used in colony 19. This design reduces genetic variability between the initial and reciprocal crosses as we do not have inbred lines of *B. terrestris*. (b) Overview schematic for identifying allelic methylation differences in the worker offspring. SNPs unique to either the mother or father are used to create N-masked reference genomes. The worker daughter sample is then aligned to the genome and reads are filtered to keep only those with an informative parental SNP. Methylation differences between the alleles can then be assessed and parent-of-origin DNA methylation can be inferred from comparing reciprocal crosses.

DNA extraction and sequencing

Whole genome bisulfite sequencing was generated for the parents and offspring of each colony. DNA was extracted from whole heads of the mother and father of each colony as well as from 12 reproductive workers (three per colony) using the Qiagen DNeasy® Blood & Tissue Kit following the manufacturers protocol. Reproductive workers were chosen to reduce the variation between samples as sterile and reproductive workers show different DNA methylation profiles (Marshall *et al.*, 2019). Each sample was treated with RNAse. DNA from the three reproductive worker samples per colony was pooled in equal quantities to produce one representative offspring sample per colony. DNA quantity and quality were determined by Nanodrop and Qubit® fluorometers as well as via

gel electrophoresis. Samples were sent to BGI Tech Solution Co., Ltd.(Hong Kong) for library preparation, bisulfite treatment and sequencing. Paired-end libraries (2 x 150bp) were sequenced across two lanes of an Illumina HiSeq 4000 platform with 40% phiX inclusion. A 1% lambda DNA spike was included in all libraries in order to assess bisulfite conversion efficiency, as the lambda genome is known to be unmethylated.

Whole genome re-sequencing of the parents was also carried out. DNA was extracted from half of the thorax of each mother and father per colony following a custom protocol (`https://github.com/agdelafilia/wet_lab/blob/master/gDNA_extraction_protocol.md`). DNA quantity and quality were determined by Nanodrop and Qubit® fluorometers as well as via gel electrophoresis. Samples were sent to Novogene Co., Ltd. for library preparation and sequencing. Paired-end libraries (2 x 150bp) were sequenced on an Illumina HiSeq 4000 platform.

Differential DNA methylation between castes and sexes

Whole genome bisulfite sequencing (WGBS) data of the parents and pooled worker offspring were checked using fastqc v.0.11.5 (Andrews, 2010) and poor quality bases were trimmed using cutadapt v.1.11 (Martin, 2011). Libraries were then aligned to the reference genome (Bter_1.0, Refseq accession no. GCF_000214255.1, (Sadd *et al.*, 2015)) using Bismark v.0.16.1 (Krueger and Andrews, 2011) and bowtie2 v.2.2.6 (Langmead and Salzberg, 2013) with standard parameters (supplementary 1.0.1). Bismark was also used to extract methylation calls and carry out deduplication. Coverage outliers (above the 99.9th percentile) were removed along with bases covered by less than 10 reads. The methylation status of each CpG was then determined via a binomial model, where the success probability is the non-conversion rate determined from the lambda spike. CpG sites were then filtered to remove any site that did not return as methylated in at least one sample.

Differential methylation was assessed at the CpG level in pair-wise comparisons (queen-male, queen-worker, male-worker) using the R package methylKit v.1.16.1 (Akalin *et al.*, 2012). A logistic regression model was applied to each comparison with Benjamini-Hochberg correction for multiple

testing (Benjamini and Hochberg, 1995). For a CpG to be differentially methylated a minimum difference of at least 10% methylation and a q-value of <0.01 were required. Genes were determined as differentially methylated genes if they contained an exon with at least two differentially methylated CpGs and an overall weighted methylation (Schultz *et al.*, 2012) difference across the exon of >15%. Two CpGs were chosen based on Xu *et al.* (2021), they find the methylation of two CpGs is enough to promote gene transcription in *Bombyx mori* via the recruitment of histone modifications.

Identification of parent-of-origin DNA methylation

Whole genome re-sequencing data of the parents were checked using fastqc v.0.11.5 (Andrews, 2010) and aligned to the reference genome (Bter_1.0, Refseq accession no. GCF_000214255.1, (Sadd *et al.*, 2015)) using bowtie2 v.2.2.6 (Langmead and Salzberg, 2013) in *–sensitive* mode (supplementary 1.0.2). Aligned reads were deduplicated and realigned around indels using GATK v.3.6 (McKenna *et al.*, 2010). SNPs were called using Freebayes v.0.9.21.7 (Garrison and Marth, 2012) which accounts for ploidy differences between males and females. SNPs were then filtered using VCFtools v.0.1.16 (Danecek *et al.*, 2011) with the following options: *–max-alleles 2 –minQ 20 –min-meanDP 10 –recode –recode-INFO-all*. A custom script was then used to filter SNPs to keep only homozygous alternative SNPs which are unique to either the mother or father of each colony. We also removed C-T and T-C SNPs as these are indistinguishable from bisulfite converted bases in WGBS. This left a mean of 365,372 SNPs per colony, allowing the parental alleles of the offspring to be identified (supplementary 1.0.3, Fig. 1b). The parental SNPs identified above were then used to create an N-masked genome for each colony (four total) using the BEDtools v.2.28.0 *maskfasta* command (Quinlan and Hall, 2010). WGBS libraries from the workers were then aligned to their respective colony's N-masked genome, using Bismark v.0.16.1 (Krueger and Andrews, 2011) as above. BEDtools v.2.28.0 (Quinlan and Hall, 2010) was then used to select reads from the WGBS alignments of the worker offspring which contained either a unique maternal or paternal SNP, i.e. reads from either the maternal allele or the paternal allele (supplementary 1.0.4). All other reads

which did not contain an informative SNP were discarded. Differential methylation between the maternal and paternal reads of all workers was then carried out using the R package methylKit v.1.16.1 (Akalin *et al.*, 2012) as above, with the exception of a minimum coverage of eight reads, as previously described in (Wang *et al.*, 2016).

Gene ontology enrichment

Gene ontology (GO) terms for *B. terrestris* were taken from a custom database made in Bebane *et al.* (2019). GO enrichment analysis was carried out using the hypergeometric test with Benjamini-Hochberg (Benjamini and Hochberg, 1995) multiple-testing correction, q <0.05. GO terms from differentially methylated genes between sexes and castes were tested against a GO term database made from the GO terms associated with all methylated genes. Genes were determined as methylated if they had a mean weighted methylation level greater than the bisulfite conversion error rate of 0.05 in either queens, males or workers. REVIGO (Supek *et al.*, 2011) was used to generate GO descriptions from the GO ids.

Honeybee comparative analysis

To confirm the consistency of parent-of-origin DNA methylation in independently generated honeybee WGBS data we used the R package DAMEfinder v.1.2.0 (Orjuela *et al.*, 2020) to identify allele-specific DNA methylation in samples from the following recent studies: Yagound *et al.* (2019, 2020); Cardoso-Júnior *et al.* (2021); Yi *et al.* (2021), totalling 58 unique samples after discarding poor quality libraries. We quality checked the raw data with fastqc v.0.11.5 (Andrews, 2010) and trimmed poor quality bases using cutadapt v.1.11 (Martin, 2011). Data were then aligned to the honeybee reference genome (Amel_HAv3.1, Refseq accession no. GCA_003254395.2, (Wallberg *et al.*, 2019)) using Bismark v.0.16.1 (Krueger and Andrews, 2011) and bowtie2 v.2.2.6 (Langmead and Salzberg, 2013) with standard parameters. We then used methtuple v.1.5.4 (`https://github.com/PeteHaitch/methtuple`) to extract the methylation calls of pairs of

CpGs. These data were then used in DAMEfinder v.1.2.0 (Orjuela *et al.*, 2020) to identify occurrences of allele-specific DNA methylation. For a region to be classed as showing allele-specific DNA methylation a minimum coverage of 10 was required and a minimum score of 0.8, which is considered representative of true allele-specific DNA methylation according to Orjuela *et al.* (2020). We then identified genes which contained allelically methylated regions using R and compared these gene lists to those which show parent-of-origin DNA methylation, as identified in Wu *et al.* (2020). Gene ontology enrichment was carried out as above using GO terms from the Hymenoptera Genome Database (Elsik *et al.*, 2016; Walsh *et al.*, 2021).

Results

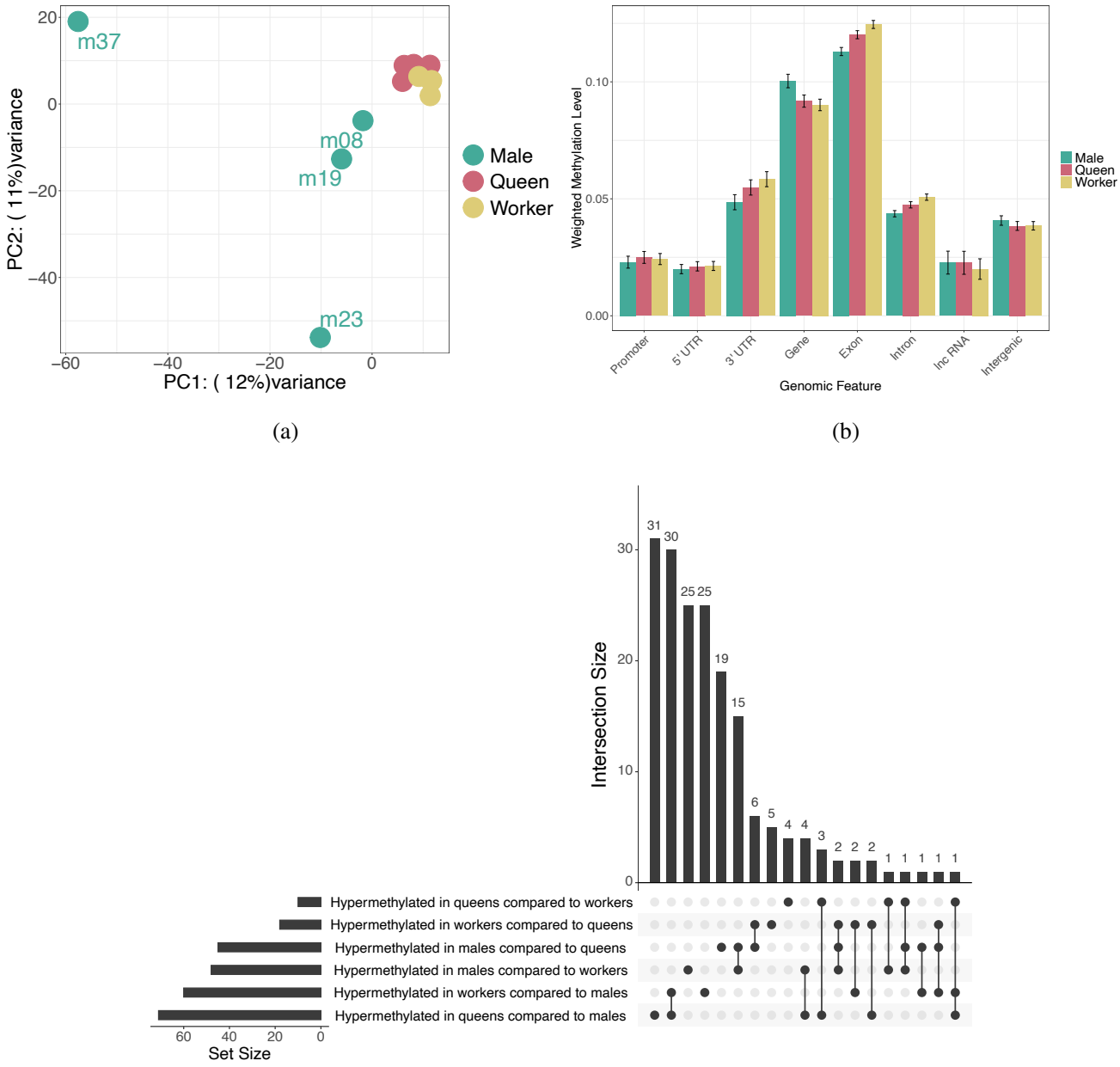
Genome-wide sex- and caste-specific DNA methylation

It is currently unknown to what extent DNA methylation varies between sexes and castes of *B. terrestris*. We have therefore taken this opportunity to also generally characterise the sex- and caste-specific methylomes of this species. We find low genome-wide levels similar to those previously reported in Bebane *et al.* (2019) and Marshall *et al.* (2019), on average 0.25% ± 0.05% across all samples, with little variation between sexes and castes (supplementary 1.0.1). Reproductive workers, queens and males do, however, show different CpG methylation profiles, with males clustering away from the two female castes (Fig.2a). We also see no clustering by sub-species for the males or queens, for example male 08 in Figure 2a represents *Bombus terrestris audax* and male 19 represents *Bombus terrestris dalmatinus*.

Genome-wide, we see overall similar levels of DNA methylation across various genomic features for both sexes and castes (Fig.2b). It has recently been shown that promoter DNA methylation exists in some insect species (Lewis *et al.*, 2020; Bain *et al.*, 2021). We have, therefore, annotated putative promoter regions in *B. terrestris*, defined at 500bp upstream of a gene with no overlap with other genomic features, we also added UTR regions and intergenic regions to further explore

the genome-wide methylation profile. We find the highest levels of DNA methylation for all sexes and castes are within exon regions, whilst promoter, and 5' UTR regions show a depletion in DNA methylation compared to intergenic regions (Fig.2b).

We also segregated genes into categories of differing levels of DNA methylation to explore the potential function of highly methylated genes across sexes and castes. There are a small number of genes classed as highly methylated (weighted methylation level >0.7) across each sex/caste (supplementary Fig.S2, supplementary 1.0.5). Most highly methylated genes in queens and workers are also found in another caste/sex. Whereas males show a larger number of unique genes which are highly methylated (supplementary Fig.S2, n = 62). We then carried out an gene ontology enrichment test for each list of highly methylated genes per sex/caste and compared these to lists of genes classed as methylated (i.e. a weighted methylation level across the gene greater than the lambda conversion rate) for each sex/caste. We find a variety of GO terms enriched across sexes and castes mostly involved in core cellular processes (supplementary 1.0.6). Male highly methylated genes did, however, have a few GO terms enriched for RNA splicing (GO:0008380, GO:0000377, GO:0000387) which were not present in the queen and worker enriched GO terms.



(a) (b)

(c)

Figure 2: (a) PCA plot based on the methylation level per CpG for all CpGs which had greater than 10X in all samples and were classed as methylated in at least one sample (n = 5,304). (b) Bar plot of the mean methylation level of each genomic feature for sexes and castes. Error bars represent 95% confidence intervals of the mean. Promoters are putative and represented by 500bp upstream of a gene without any other genomic feature overlap. (c) Upset plot showing common genes containing a hypermethylated exon per hypermethylated sex/caste per comparison. The set size indicates the total number of hypermethylated genes, the intersection size shows how many of those are common between sets, as indicated by the connections in the bottom panel. E.g. 31 genes are uniquely hypermethylated in queens compared to males and 30 genes are hypermethylated in both queen and workers compared to males.

Differential DNA methylation between sexes and castes

A differential DNA methylation analysis between sexes and castes found a total of 1,011 differentially methylated CpGs between males and reproductive workers, 824 differentially methylated CpGs between males and queens and 156 differentially methylated CpGs between queens and reproductive workers. Roughly equal numbers were hypermethylated in each sex/caste per comparison, except for males and queens where queens show slightly more hypermethylated sites (Chi-squared Goodness of Fit: χ^2 = 8.97, df = 1, p-value < 0.01, male n = 369, queen n = 455). The majority of all differentially methylated CpGs are located within genes and specifically within exons, we also find a slight depletion of differentially methylated CpGs in the first exon compared to the following exons (supplementary Fig.S1), this is in line with DNA methylation being slightly lower in the first exon in *B. terrestris* (Lewis *et al.*, 2020).

We next classed a gene as differentially methylated if a given exon contained at least two differentially methylated CpGs and had an overall weighted methylation difference of at least 15%. We find 155 genes are differentially methylated between males and workers, 165 between males and queens and 37 between queens and workers (supplementary 1.0.7). We carried out a GO enrichment analysis on all differentially methylated genes and on hypermethylated genes for each sex/caste per comparison (supplementary 1.0.8). Whilst most terms are involved in core cellular processes, we specifically find differentially methylated genes between queens and workers are enriched for chromatin remodelling-related terms (e.g. "*histone H3-K27 acetylation*" (GO:0043974) and "*chromatin organization involved in negative regulation of transcription*" (GO:0097549)) and reproductive terms (e.g. "*oogenesis*" (GO:0048477)). Differentially methylated genes between males and workers were also enriched for a large number of histone modification related terms (e.g. "*histone H3-K27 acetylation*" (GO:0043974), "*histone H3-K9 methylation*" (GO:0051567), "*regulation of histone H3-K9 trimethylation*" (GO:1900112)) as well as "*dosage compensation*" (GO:0007549) and some reproductive related terms (e.g. "*gamete generation*" (GO:0007276)).

Multiple histone related terms were also found for differentially methylated genes between males and queens, as well as the above we also found "*histone H4-K20 demethylation*" (GO:0035574), "*histone H4-K8 acetylation*" (GO:0043982), "*histone H4-K16 acetylation*" (GO:0043984) and "*histone H4-K5 acetylation*" (GO:0043981).

When looking specifically at hypermethylated genes per sex/caste compared to all differentially methylated genes per comparison we find only two enriched GO terms for hypermethylated genes in queens compared to workers: "*developmental process involved in reproduction*" (GO:0003006) and "*gamete generation*" (GO:0007276). In genes hypermethylated in males compared to queens and workers separately we find a large number of enriched GO terms related to neuron development amongst other cellular processes.

Most of the differentially methylated genes are common between males, queens and workers, with only 178 total unique genes changing methylation levels between sexes/castes (Fig.2c). Specifically, we find 31 genes are hypermethylated in queens and workers when compared to males and 18 genes are hypermethylated in males when compared to queens and workers. We carried out a GO enrichment on these genes using all differentially methylated genes from all comparisons as a background set. We find general cellular processes enriched in both gene lists with hypermethylated genes in the female castes also enriched for some telomere-related functions, e.g. "*telomere organization*" (GO:0032200) and "*regulation of telomere maintenance*" (GO:0032204).

Parent-of-origin and lineage-of-origin DNA methylation in *B. terrestris*

Using SNPs called from the parental genomes of each colony we were able to identify reads within the WGBS of the offspring which belong to each parental allele. We analysed each reciprocal cross separately to maximise the number of common CpG sites covered in the samples. This allowed us to identify the parental origin of 6,091 and 4,883 CpGs with at least eight times coverage in all worker samples per cross respectively. Given that in our sex- and caste-specific differential DNA methylation comparisons we have >2 million CpGs identified, we calculate that we have only been

able to survey around 0.25% of the bumblebee methylome for parent-of-origin and lineage-of-origin DNA methylation. We therefore present this component of the paper as a learning experience in the hope to improve future research in this field (see discussion).

Of the 6,091 and 4,883 CpG identified, three and two respectively were classed as being methylated in at least one sample by a binomial test and were tested for parent-of-origin and lineage-of-origin effects. None of these positions show evidence of parent-of-origin or lineage-of-origin DNA methylation.

Consistency of parent-of-origin DNA methylation in the honeybee

In order to explore the consistency of parent-of-origin DNA methylation in independently generated honeybee data sets we selected genes from Wu *et al.* (2020) which contained a minimum of two CpGs showing parent-of-origin DNA methylation. Using a list of unique genes from either sterile or reproductive samples from either of the two genetic blocks tested in Wu *et al.* (2020) we selected a final list of 166 unique genes which show parent-of-origin DNA methylation, 91 with paternal-of-origin DNA methylation and 75 with maternal-of-origin DNA methylation (supplementary 1.0.9).

We then called allele-specific DNA methylation in independently generated honeybee female data sets (n = 33) spanning various tissues of both workers and queens from Cardoso-Júnior *et al.* (2021); Yi *et al.* (2021) and Yagound *et al.* (2020). We find around 1,500 genes per sample show allele-specific DNA methylation with slightly lower numbers in worker thorax tissue (Fig.S3). Of the 166 identified genes which show parent-of-origin DNA methylation, 132 are found in at least one independent data set showing allele-specific DNA methylation. Specifically, each data set shows allele-specific DNA methylation for between 45-53% of the 166 genes with parent-of-origin DNA methylation (Fig.S3). From these we have identified a core set of 12 genes which are present in all 33 data sets, six of which show maternal-of-origin DNA methylation and six show paternal-of-origin DNA methylation as determined in Wu *et al.* (2020) (Table 1).

In order to further explore the function of these core 12 genes we carried out a gene ontology

enrichment analysis using all unique genes with allele-specific DNA methylation as a background set (n = 3,448). We find a variety of terms enriched, including many involved in nervous system development and the term "*social behaviour*" (GO:0035176) (supplementary 1.1.1).

Table 1: List of the 12 genes which show parent-of-origin DNA methylation in Wu *et al.* (2020) as well as allele-specific DNA methylation in all 33 independent honeybee female data sets. The mean ASM score represents the mean allele-specific DNA methylation score called by DAMEfinder taken across all 33 data sets, a score >0.8 is indicative of true allele-specific DNA methylation (Orjuela *et al.*, 2020). * These genes also show allele-specific DNA methylation in at least 50% of all male thorax samples tested.

BEEBASE	**LOC ID**	**Name**	**Parent-of-origin DNA methylation**	**Mean ASM Score**
GB41287	LOC412544	protein dopey-1 homolog	Maternal	1.12
GB41714	LOC727150	uncharacterized	Paternal	1.18
GB41884	LOC412746	calcineurin-binding protein cabin-1	Maternal	1.02
GB42246	LOC551731	rho GTPase-activating protein 190	Paternal	1.08
GB46478	LOC412224	tectonin beta-propeller repeat-containing protein	Maternal	1.12
GB47418	LOC409282	bromodomain-containing protein	Maternal	1.14
GB47466	LOC552221	F-BAR domain only protein 2	Maternal	1.18
GB48854*	LOC413947	sodium leak channel non-selective protein	Paternal	1.09
GB50958	LOC408286	2-oxoglutarate dehydrogenase, mitochondrial	Maternal	1.3
GB51276*	LOC551848	protocadherin-like wing polarity protein stan	Paternal	1.09
GB51956*	LOC412281	protein purity of essence	Paternal	1.15
GB53269	LOC412643	uncharacterized	Paternal	1.13

We also examined genes which show both parent-of-origin DNA methylation from Wu *et al.* (2020) and allele-specific DNA methylation in at least half of all data sets (n >= 16), which gave a total of 91 genes, 46 of which show maternal-of-origin DNA methylation and 45 show paternal-of-origin DNA methylation (supplementary 1.1.2). A GO enrichment analysis of these genes using all genes which show allele-specific DNA methylation as a background set revealed a variety of processes including multiple nervous system development terms, histone modification related terms (e.g. "*histone H4-K20 demethylation*" GO:0035574) and some terms involved in reproduction (e.g. "*oocyte nucleus migration involved in oocyte dorsal/ventral axis specification*" GO:0007312).

In addition to identifying this core set of genes which show potentially consistent parent-of-origin DNA methylation across multiple independent data sets, we also ran this pipeline for some male samples as previous research has shown a diploid genome exists in some tissues of *A. mellifera* males as a potential mechanism of dosage compensation (Mittwoch *et al.*, 1966; Aron *et al.*, 2005). As expected we find little to no allele-specific DNA methylation in sperm samples (supplementary Fig.S4). As there is only one allele present any arising allele-specific methylation events are likely false positives. However, we do find between 332-538 genes showing allele-specific DNA methylation in male thorax tissue, across 10 replicate samples from Yagound *et al.* (2020).

We checked these genes which show allele-specific DNA methylation to see if they were also called as showing parent-of-origin DNA methylation in Wu *et al.* (2020). Of the 166 genes which show parent-of-origin DNA methylation, 79 are found in at least one male sample showing allele-specific DNA methylation. Each data set shows allele-specific DNA methylation for between 18-24 (i.e. 10-14%) of the genes which show parent-of-origin DNA methylation identified in Wu *et al.* (2020) (Fig.S4). There is one gene which is present in all 10 data sets showing allele-specific DNA methylation and parent-of-origin DNA methylation, this is *protein crumbs* (LOC725591, GB46140), which is maternally methylated in females. GO terms associated with this gene are involved in multiple body plan developmental processes.

There are also 12 genes which show allele-specific DNA methylation in at least 50% of the male thorax data sets (n >=5, supplementary 1.1.2), seven of which show maternal methylation in females and five of which show paternal methylation in females. Gene ontology terms enriched for these 12 genes compared to all genes showing allele-specific DNA methylation in males (n = 2,010) included many histone demethylation terms (e.g. "*histone H3-K9 demethylation*" GO:0033169 and "*histone H3-K27 demethylation*" GO:0071557), "*social behaviour*" (GO:0035176) and various other developmental processes (supplementary 1.1.1). Of these 12 genes, three also show allele-specific DNA methylation in all female data sets (Table 1). Upon inspection of the GO terms related to these genes we find GB48854 is driving the enriched term involved in social behaviour and GB51276 and GB51956 are driving nervous system development enriched GO terms. Interestingly, whilst not enriched the GO term spermatid development is associated with GB51956.

Discussion

In this study we have explored the potential of DNA methylation as an imprinting mark in social bees. We conducted reciprocal crosses to explore parent-of-origin DNA methylation in the bumblebee *Bombus terrestris*. Whilst our crosses and data generation were successful, we were unable to confidently identify genome-wide parent-of-origin DNA methylation. We were, however, able to use these data to characterise the sex- and caste-specific DNA methylation profiles of *B. terrestris* for the first time. We find genome-wide that sexes and castes show similar DNA methylation profiles, however there are a number of genes which are differentially methylated between sexes and castes. Males specifically show a large number of differentially methylated genes involved in other epigenetic processes, such as histone modifications and chromatin dynamics. We also mined previously generated honeybee whole genome bisulfite sequencing data to explore the consistency of parent-of-origin DNA methylation, as identified in Wu *et al.* (2020), across independent data sets. We find a core set of 12 genes which exhibit parent-of-origin DNA methylation show allele-specific

DNA methylation in all 33 independently generated female data sets. We have also identified a potential role for allele-specific DNA methylation in some diploid tissues of male honeybees.

Recommendations for *B. terrestris* reciprocal cross design for parent-of-origin DNA methylation

We used whole genome re-sequencing data of the mother and father from two sets of reciprocal crosses in order to call SNPs to identify the parent-of-origin of the alleles in the offspring. We find a similar number of unique informative SNPs between our crosses compared to previous studies which have identified parent-of-origin expression in bumblebees (Marshall *et al.*, 2020b) and honeybees (Galbraith *et al.*, 2016) and parent-of-origin DNA methylation in honeybees (Wu *et al.*, 2020). This indicates the sub-species used in this study were genetically different enough to be able to identify the parent-of-origin of the alleles in the offspring. It should be noted for future study design, however, that the number of SNPs identified has a direct effect on the percentage of the genome that can be surveyed for parent-of-origin effects. For example, Smith *et al.* (2020) were only able to call around 7000 informative SNPs from honeybee reciprocal crosses which meant they could only survey 17% of all genes for parent-of-origin expression. Whilst we have obtained enough resolution through SNPs, we were still unable to identify parent-of-origin DNA methylation across the entire genome of *B. terrestris*. We explore the reasons for this and make the following recommendations for a replication of this work.

Firstly, we sequenced the worker offspring samples to a depth of 30X, this coverage yields enough data for standard differential DNA methylation analysis even after data loss due to low mapping rates of bisulfite converted data (generally less than 60% mapping efficiency (Tran *et al.*, 2014)) and removal of PCR duplicates. An additional step required to identify parent-of-origin DNA methylation involves allocating reads to the maternal or paternal chromosomes, this means only half of all remaining data is usable for each maternal/paternal-of-origin gene. Given that statistically a minimum coverage of 10 reads are required per CpG to avoid excess false positive calls (Ziller *et al.*,

2015), we would recommend sequencing to a depth of at least 50X in future work to maximise the final number of CpG sites which meet this minimum coverage requirement.

Secondly, we chose to pool DNA from worker head tissue per colony, resulting in only one sequencing library for each colony of each cross. It has recently been shown that variation in DNA methylation within Hymenopteran species can be largely driven by the underlying genotype, i.e. the genomic sequence (Yagound *et al.*, 2019; Marshall *et al.*, 2019; Yagound *et al.*, 2020). Pooling samples therefore adds additional variation per sequencing library which may confound any lineage-of-origin DNA methylation profiles. Avoiding pooling samples would therefore reduce variation per library.

Finally, whilst we were selective in the tissue we sequenced we would recommend sequencing even more specific tissue types as DNA methylation profiles are known to differ between tissues (Pai *et al.*, 2011). A relatively new technique called Enzymatic Methyl-Seq (Vaisvila *et al.*, 2021) may be used in place of traditional bisulfite sequencing which allows a much lower initial input DNA quantity and yields the same unmethylated-cytosine to thymine conversion as bisulfite sequencing. This may be used, for example, to examine parent-of-origin methylation profiles in individual brain samples.

Sex- and caste-specific methylomes of *B. terrestris*

We were able to use the data generated in this study to explore the sex- and caste-specific methylome of *B. terrestris*. Whilst genome-wide, males show similar DNA methylation profiles to the two female castes, in terms of DNA methylation localisation to exons and depletion in promoter regions, we also found a number of differentially methylated genes between males and females. Specifically, differentially methylated genes are enriched for many histone modification related processes. It has recently been found in the silk moth that the presence of DNA methylation promotes histone H3-K27 acetylation which changes the chromatin formation of a region allowing gene expression (Xu *et al.*, 2021). The relationship between DNA methylation and histone modifications in social

insects remains unknown. However, recent work by Choppin *et al.* (2021) shows a role for histone acetylation in the regulation of worker reproduction and gene expression in the ant *Temnothorax rugatulus*. An exploration of the functional relationship between DNA methylation and histone modifications is needed across a greater diversity of insect species in order to understand how these processes may interact to produce downstream gene expression and thus phenotype differences.

In addition to histone related genes we also find both female castes show hypermethylated genes which are involved in telomere functions. Telomere length is thought to correlate with lifespan (Cawthon *et al.*, 2003) and it has been found in the ant *Lasius niger* that the shorter lived males do indeed posses shorter telomeres than females (Jemielity *et al.*, 2007). The lifespan of *B. terrestris* queens is significantly longer than workers and males (Greeff and Schmid-Hempel, 2008; Smeets and Duchateau, 2003; Duchateau and Marin, 1995). One role for DNA methylation in *B. terrestris* may therefore be the regulation of caste differences through core cellular processes, such as telomere maintenance. Finally, we also find differentially methylated genes between queens and reproductive workers are involved in reproductive related processes. Previous work has suggested a role for DNA methylation in reproduction in *B. terrestris* (Amarasinghe *et al.*, 2014), as well as other social insects (Wang *et al.*, 2020; Bonasio *et al.*, 2012), although this does not appear to be consistent across Hymenoptera (Libbrecht *et al.*, 2016; Patalano *et al.*, 2015). Whilst the differentially methylated genes identified here suggest a role for DNA methylation in maintaining or generating caste differences, a direct causal link between DNA methylation and gene expression changes mediating phenotypes has yet to be found.

Consistency of parent-of-origin DNA methylation in the honeybee

Genes which show parent-of-origin expression have been identified in two social insect species to date, *B. terrestris* (Marshall *et al.*, 2020b) and *A. mellifera* (Wu *et al.*, 2020). Whilst a direct link between parent-of-origin DNA methylation and parent-of-origin expression has not been found in the honeybee (Wu *et al.*, 2020; Smith *et al.*, 2020), it is possible parent-of-origin DNA methylation may

mediate imprinted genes in a trans- or temporal-acting fashion (Xu *et al.*, 2021; Li-Byarlay *et al.*, 2020). For example, in mammals DNA methylation can act to silence an imprinting control element, which when normally expressed would silence nearby genes (Barlow, 2011). This highlights the complex nature of interactions between epigenetic mechanisms and gene expression and shows why direct correlation of expression and methylation levels of a given gene may not be indicative of the function of those specific epigenetic marks.

As part of this study set out to identify genes which may show consistent parent-of-origin DNA methylation across independently generated honeybee whole genome bisulfite sequencing data in order to provide a strong list of candidate genes which may be involved in genomic imprinting in *A. mellifera*. We find a core set of 12 genes which show allele-specific DNA methylation in 33 female data sets from both queens and workers across various tissues. One of these genes is specifically involved in regulating social behaviour. It is predicted by the kinship theory that imprinted genes in social insects should play a role in mediating worker reproductive behaviour (Haig, 2000; Queller, 2003). The identification of a specific gene which may regulate this process provides an ideal candidate for experimental epigenetic manipulation, through techniques such as CRISPR-Cas9 (Vojta *et al.*, 2016), enabling a functional assessment of the parent-of-origin specific DNA methylation marks present.

It is also worth noting that if parent-of-origin DNA methylation does act as an imprinting mark, in some capacity in *A. mellifera*, that we might expect there to be more than 12 genes involved in genomic imprinting. Firstly, some imprinted genes in mammals are known to be tissue specific (Prickett and Oakey, 2012) and here we only examined genes with allele-specific DNA methylation when it occurred across all tissues tested. Additionally, it may be expected that workers and queens display different imprinted genes, as genomic imprinting has been suggested to play a role in caste determination in some social insects (Matsuura, 2019). Again, this suggests the 12 genes identified may only represent a subset of all imprinted genes in *A. mellifera*. It has also been suggested that imprinted genes may respond in a plastic manner to environmental conditions, for example a silenced

allele may become activated in order to increase gene dosage in response to external stimuli (Radford *et al.*, 2011). Imprinted genes in *A. mellifera* workers are predicted to be involved in reproduction (Haig, 2000), and so a plastic response to queen presence of some imprinted genes may account for the small number in common across independent samples. Finally, we cannot rule out that more of the genes with parent-of-origin DNA methylation identified in Wu *et al.* (2020) are not consistent across *A. mellifera* females. As discussed in the 'Recommendations for *B. terrestris* reciprocal cross design for parent-of-origin DNA methylation' section above, it may be that we did not have sufficient coverage in some genome regions / samples from the data tested for these areas to show significant allele-specific DNA methylation. Although, it's worth noting the advantage of identifying allele-specific methylation through probabilistic models as opposed to using SNPs is that we can survey homozygous regions which would usually be discounted when differences in the underlying genotype are needed for allele identification (Orjuela *et al.*, 2020).

Finally, of the core 12 genes identified above we find three of those genes also show allele-specific DNA methylation in some male thorax tissue, including the gene involved in social behaviour. Different tissues are known to vary in levels of ploidy in some social insects (Aron *et al.*, 2005). DNA methylation has also previously been suggested as a possible mechanism of ploidy compensation in haplodiploid insects, with Glastad *et al.* (2014) finding diploid males of the fire ant, *Solenopsis invicta*, show more similar methylation patterns to diploid females compared to haploid males. The discovery of genes showing allele-specific DNA methylation in male thorax tissue is suggestive of a role for DNA methylation in tissue-specific ploidy compensation in *A. mellifera*.

Conclusion

This study provides the groundwork for future research exploring parent-of-origin DNA methylation as a potential imprinting mechanism in the bumblebee *Bombus terrestris*. We specifically highlight technical recommendations for adequate data generation from reciprocal crosses needed to identify parent-of-origin DNA methylation at a genome-wide scale. We have also characterised the sex-

and caste-specific methylome of *B. terrestris* identifying a potential role for DNA methylation in downstream epigenetic regulatory processes which may influence sex and caste phenotypic differences. Additionally, using genes previously identified as showing parent-of-origin DNA methylation in the honeybee we have identified a core set of these genes which appear to show conserved allele-specific DNA methylation across female castes and tissues. These genes can serve as candidates for experimental manipulation to explore the functional role of parent-of-origin DNA methylation in *A. mellifera*. Finally, we have identified allele-specific DNA methylation in honeybee male thorax tissue which suggests a role for DNA methylation in ploidy compensation in this species.

Acknowledgements

We thank Biobest N.V. for providing us with the forty bumblebee colonies of the desired reciprocal crosses and Dr Ben Hunt for analysis advice. Hollie Marshall had a PhD-scholarship from the Central England NERC Training Alliance (NERC, UK). This study was funded in the context of the NERC grant NE/N010019/1 and a Leverhulme Trust Research Project Grant (RPG-2020-363) awarded to E.B.M., the KU Leuven BOF Centre of Excellence Financing on 'Eco- and socio-evolutionary dynamics' (Project number PF/2010/07) and a grant from the Research Foundation-Flanders (FWO-Vlaanderen, grant G.0463.12) awarded to T.W. and the European Research Council grant: PGErepro awarded to L.R. This research used the ALICE2 High Performance Computing Facility at the University of Leicester. Thank you to Christian Thomas for shipping the frozen bees from Leicester to Edinburgh.

Author contributions

E.B.M., T.W. and H.M. conceived the study. The reciprocal crosses were carried out by Biobest (Westerlo, Belgium) under supervision of F.W. J.S.Z. carried out the dissections to confirm the reproductive status of workers. H.M. carried out the wet lab work. H.M. and M.T.N carried out the

analyses. L.R. contributed the whole genome re-sequencing. H.M. wrote the initial manuscript. All authors contributed to and reviewed the final manuscript.

Data Accessibility

Data has been deposited in GenBank under NCBI BioProject: PRJNA779586. All code is available at: `https://github.com/MooHoll/Parent_of_Origin_Methylation`.

References

Akalin, A., Kormaksson, M., Li, S., Garrett-bakelman, F. E., Figueroa, M. E., Melnick, A., and Mason, C. E. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13(R87).

Amarasinghe, H. E., Clayton, C. I., and Mallon, E. B. 2014. Methylation and worker reproduction in the bumble-bee (*Bombus terrestris*). *Proceedings of the Royal Society B: Biological Sciences*, 281(20132502).

Andrews, S. 2010. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. *http://www.bioinformatics.babraham.ac.uk/projects/fastqc/*, Accessed: 12-12-2020.

Aron, S., de Menten, L., Van Bockstaele, D. R., Blank, S. M., and Roisin, Y. 2005. When Hymenopteran Males Reinvented Diploidy. *Current Biology*, 15(9): 824–827.

Bain, S. A., Marshall, H., Filia, A. G., Laetsch, D. R., Husnik, F., and Ross, L. 2021. Sex-specific expression and DNA methylation in a species with extreme sexual dimorphism and paternal genome elimination. *Molecular Ecology*, page mec.15842.

Barlow, D. P. 2011. Genomic Imprinting: A Mammalian Epigenetic Discovery Model. *Annual Review of Genetics*, 45(1): 379–403.

Barlow, D. P. and Bartolomei, M. S. 2014. Genomic Imprinting in Mammals. *Cold Spring Harb Perspect Biol*, 6(a018282).

Bebane, P. S. A., Hunt, B. J., Pegoraro, M., Jones, A. R. C., Marshall, H., Rosato, E., and Mallon, E. B. 2019. The effects of the neonicotinoid imidacloprid on gene expression and DNA methylation in the buff-tailed bumblebee *Bombus terrestris*. *Proc. R. Soc. B*, 286: 20190718.

Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57(1): 289–300.

Bewick, A. J., Vogel, K. J., Moore, A. J., and Schmitz, R. J. 2016. Evolution of DNA methylation across insects. *Molecular Biology and Evolution*, 34(3): 654–665.

Bonasio, R., Li, Q., Lian, J., Mutti, N. S., Jin, L., Zhao, H., Zhang, P., Wen, P., Xiang, H., Ding, Y., Jin, Z., Shen, S. S., Wang, Z., Wang, W., Wang, J., Berger, S. L., Liebig, J. J., Zhang, G., and Reinberg, D. 2012. Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Current Biology*, 22(19): 1755–1764.

Cardoso-Júnior, C. A. M., Yagound, B., Ronai, I., Remnant, E. J., Hartfelder, K., and Oldroyd, B. P. 2021. DNA methylation is not a driver of gene expression reprogramming in young honey bee workers. *Molecular Ecology*, 30(19): 4804–4818.

Cawthon, R. M., Smith, K. R., O'Brien, E., Sivatchenko, A., and Kerber, R. A. 2003. Association between telomere length in blood and mortality in people aged 60 years or older. *The Lancet*, 361(9355): 393–395.

Choppin, M., Feldmeyer, B., and Foitzik, S. 2021. Histone acetylation regulates the expression of genes involved in worker reproduction in the ant Temnothorax rugatulus. *BMC Genomics*, 22(1): 871.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., and 1000 Genomes Project Analysis Group 2011. The variant call format and VCFtools. *Bioinformatics*, 27(15): 2156–2158.

de la Filia, A. G., Mongue, A. J., Dorrens, J., Lemon, H., Laetsch, D. R., and Ross, L. 2021. Males That Silence Their Father's Genes: Genomic Imprinting of a Complete Haploid Genome. *Molecular Biology and Evolution*, 38(6): 2566–2581.

Drewell, R. a., Lo, N., Oxley, P. R., and Oldroyd, B. P. 2012. Kin conflict in insect societies: A new epigenetic perspective. *Trends in Ecology and Evolution*, 27(7): 367–373.

Duchateau, M. J. and Marin, J. 1995. Sexual biology of haploid and diploid males in the bumble beeBombus terrestris. *Insectes Sociaux*, 42(3): 255–266.

Duchateau, M. J. and Velthuis, H. H. W. 1988. Development and reproduction strategies in *Bombus terrestris* colonies. *Behaviour*, 107: 186–207.

Elango, N., Hunt, B. G., Goodisman, M. a. D., and Yi, S. V. 2009. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, Apis mellifera. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27): 11206–11211.

Elsik, C. G., Tayal, A., Diesh, C. M., Unni, D. R., Emery, M. L., Nguyen, H. N., and Hagen, D. E. 2016. Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine. *Nucleic Acids Research*, 44(D1): D793–D800.

Feng, S., Cokus, S. J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M. G., Hetzel, J., Jain, J., Strauss, S. H., Halpern, M. E., Ukomadu, C., Sadler, K. C., Pradhan, S., Pellegrini, M., and Jacobsen, S. E. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences*, 107(19): 8689–8694.

Foret, S., Kucharski, R., Pittelkow, Y., Lockett, G. A., and Maleszka, R. 2009. Epigenetic regulation of the honey bee transcriptome: unravelling the nature of methylated genes. *BMC Genomics*, 10(1): 472.

Galbraith, D. A., Kocher, S. D., Glenn, T., Albert, I., Hunt, G. J., Strassmann, J. E., Queller, D. C., and Grozinger, C. M. 2016. Testing the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*). *Proceedings of the National Academy of Sciences*, 113(4): 1020–1025.

578 Galbraith, D. A., Ma, R., and Grozinger, C. M. 2021. Tissue-specific transcription patterns support
579 the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*). *Molecular Ecology*,
580 30(4): 1029–1041.

581 Garrison, E. and Marth, G. 2012. Haplotype-based variant detection from short-read sequencing.
582 *arXiv*. 1207.3907.

583 Glastad, K. M., Hunt, B. G., Yi, S. V., and Goodisman, M. a. D. 2014. Epigenetic inheritance and
584 genome regulation: is DNA methylation linked to ploidy in haplodiploid insects? *Proceedings.*
585 *Biological sciences / The Royal Society*, 281(1785): 20140411.

586 Glastad, K. M., Gokhale, K., Liebig, J., and Goodisman, M. A. D. 2016. The caste- and sex-specific
587 DNA methylome of the termite *Zootermopsis nevadensis*. *Scientific Reports*, 6(37110).

588 Glastad, K. M., Arsenault, S. V., Vertacnik, K. L., Geib, S. M., Kay, S., Danforth, B. N., Rehan,
589 S. M., Linnen, C. R., Kocher, S. D., and Hunt, B. G. 2017. Variation in DNA methylation is
590 not consistently reflected by sociality in hymenoptera. *Genome Biology and Evolution*, 9(6):
591 1687–1698.

592 Greeff, M. and Schmid-Hempel, P. 2008. Sperm reduces female longevity and increases melanization
593 of the spermatheca in the bumblebee Bombus terrestris L. *Insectes Sociaux*, 55(3): 313–319.

594 Haig, D. 2000. The Kinship Theory of Genomic Imprinting. *Annual Review of Ecology, Evolution,*
595 *and Systematics*, (31): 9–32.

596 Inoue, A., Jiang, L., Lu, F., Suzuki, T., and Zhang, Y. 2017. Maternal H3K27me3 controls DNA
597 methylation-independent imprinting. *Nature*, 547: 419–424.

598 Jemielity, S., Kimura, M., Parker, K. M., Parker, J. D., Cao, X., Aviv, A., and Keller, L. 2007. Short
599 telomeres in short-lived males: what are the molecular and evolutionary causes? *Aging Cell*, 6(2):
600 225–233.

601 Kocher, S. D., Tsuruda, J. M., Gibson, J. D., Emore, C. M., Arechavaleta-Velasco, M. E., Queller,
602 D. C., Strassmann, J. E., Grozinger, C. M., Gribskov, M. R., San Miguel, P., Westerman, R., and
603 Hunt, G. J. 2015. A Search for Parent-of-Origin Effects on Honey Bee Gene Expression. *G3*
604 *Genes|Genomes|Genetics*, 5(8): 1657–1662.

605 Krueger, F. and Andrews, S. R. 2011. Bismark: A flexible aligner and methylation caller for
606 Bisulfite-Seq applications. *Bioinformatics*, 27(11): 1571–1572.

607 Langmead, B. and Salzberg, S. L. 2013. Fast gapped-read alignment with Bowtie 2. *Nature methods*,
608 9(4): 357–359.

609 Lewis, S. H., Ross, L., Bain, S. A., Pahita, E., Smith, S. A., Cordaux, R., Miska, E. A., Lenhard, B.,
610 Jiggins, F. M., and Sarkies, P. 2020. Widespread conservation and lineage-specific diversification
611 of genome-wide DNA methylation patterns across arthropods. *PLOS Genetics*, 16(6): e1008864.

Li-Byarlay, H., Boncristiani, H., Howell, G., Herman, J., Clark, L., Strand, M. K., Tarpy, D., and Rueppell, O. 2020. Transcriptomic and Epigenomic Dynamics of Honey Bees in Response to Lethal Viral Infection. *Frontiers in Genetics*, 11: 566320.

Libbrecht, R., Oxley, P., Keller, L., and Kronauer, D. 2016. Robust DNA Methylation in the Clonal Raider Ant Brain. *Current Biology*, 26(3): 391–395.

Liu, L., Pei, C., Liu, S., Guo, X., Du, N., and Guo, W. 2018. Genetic and epigenetic changes during the invasion of a cosmopolitan species (*Phragmites australis*). *Ecology and Evolution*, (13): 6615–6624.

Lonsdale, Z., Lee, K., Kiriakidu, M., Amarasinghe, H., Nathanael, D., O'Connor, C. J., and Mallon, E. B. 2017. Allele specific expression and methylation in the bumblebee, *Bombus terrestris*. *PeerJ*, 5(e3798).

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 Biology*, 8(11).

Marshall, H., Lonsdale, Z. N., and Mallon, E. B. 2019. Methylation and gene expression differences between reproductive and sterile bumblebee workers. *Evolution Letters*, 3(5): 485–499.

Marshall, H., Jones, A. R. C., Lonsdale, Z. N., and Mallon, E. B. 2020a. Bumblebee Workers Show Differences in Allele-Specific DNA Methylation and Allele-Specific Expression. *Genome Biology and Evolution*, 12(8): 1471–1481.

Marshall, H., van Zweden, J. S., Van Geystelen, A., Benaets, K., Wäckers, F., Mallon, E. B., and Wenseleers, T. 2020b. Parent of origin gene expression in the bumblebee, *Bombus terrestris* , supports Haig's kinship theory for the evolution of genomic imprinting. *Evolution Letters*, 4(6): 479–490.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1): 10.

Matsuura, K. 2019. Genomic imprinting and evolution of insect societies. *Population Ecology*, 62(1): 38–52.

Matsuura, K., Mizumoto, N., Kobayashi, K., Nozaki, T., Fujita, T., Yashiro, T., Fuchikawa, T., Mitaka, Y., and Vargo, E. L. 2018. A Genomic Imprinting Model of Termite Caste Determination: Not Genetic but Epigenetic Inheritance Influences Offspring Caste Fate. *The American Naturalist*, 191(6).

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9): 1297–1303.

Mittwoch, U., Kalmus, H., and Webster, W. S. 1966. Deoxyribonucleic Acid Values in Dividing and Non-Dividing Cells of Male and Female Larvae of the Honey Bee. *Nature*, 210(5033): 264–266.

Oldroyd, B. P. and Yagound, B. 2021a. Parent-of-origin effects, allele-specific expression, genomic imprinting and paternal manipulation in social insects. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1826): 20200425.

Oldroyd, B. P. and Yagound, B. 2021b. The role of epigenetics, particularly DNA methylation, in the evolution of caste in insect societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1826): 20200115.

Orjuela, S., Machlab, D., Menigatti, M., Marra, G., and Robinson, M. D. 2020. DAMEfinder: a method to detect differential allele-specific methylation. *Epigenetics & Chromatin*, 13(1): 25.

Pai, A. A., Bell, J. T., Marioni, J. C., Pritchard, J. K., and Gilad, Y. 2011. A Genome-Wide Study of DNA Methylation Patterns and Gene Expression Levels in Multiple Human and Chimpanzee Tissues. *PLoS Genetics*, 7(2): e1001316.

Patalano, S., Vlasova, A., Wyatt, C., Ewels, P., Camara, F., Ferreira, P. G., Asher, C. L., Jurkowski, T. P., Segonds-Pichon, A., Bachman, M., González-Navarrete, I., Minoche, A. E., Krueger, F., Lowy, E., Marcet-Houben, M., Rodriguez-Ales, J. L., Nascimento, F. S., Balasubramanian, S., Gabaldon, T., Tarver, J. E., Andrews, S., Himmelbauer, H., Hughes, W. O. H., Guigó, R., Reik, W., and Sumner, S. 2015. Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. *Proceedings of the National Academy of Sciences*, 112(45): 13970–13975.

Prickett, A. R. and Oakey, R. J. 2012. A survey of tissue-specific genomic imprinting in mammals. *Molecular Genetics and Genomics*, 287(8): 621–630.

Provataris, P., Meusemann, K., Niehuis, O., Grath, S., and Misof, B. 2018. Signatures of DNA Methylation across Insects Suggest Reduced DNA Methylation Levels in Holometabola. *Genome biology and evolution*, 10(4): 1185–1197.

Queller, D. C. 2003. Theory of genomic imprinting conflict in social insects. *BMC Evolutionary Biology*, page 23.

Quinlan, A. R. and Hall, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6): 841–842.

Radford, E. J., Ferrón, S. R., and Ferguson-Smith, A. C. 2011. Genomic imprinting as an adaptative model of developmental plasticity. *FEBS Letters*, 585(13): 2059–2066.

Rodrigues, J. A. and Zilberman, D. 2015. Evolution and function of genomic imprinting in plants. *Genes and Development*, 29(24): 2517–2531.

Sadd, B. M., Barribeau, S. M., Bloch, G., de Graaf, D. C., Dearden, P., Elsik, C. G., Gadau, J., Grimmelikhuijzen, C. J., Hasselmann, M., Lozier, J. D., Robertson, H. M., Smagghe, G., Stolle, E., Van Vaerenbergh, M., Waterhouse, R. M., Bornberg-Bauer, E., Klasberg, S., Bennett, A. K.,

Câmara, F., Guigó, R., Hoff, K., Mariotti, M., Munoz-Torres, M., Murphy, T., Santesmasses, D., Amdam, G. V., Beckers, M., Beye, M., Biewer, M., Bitondi, M. M., Blaxter, M. L., Bourke, A. F., Brown, M. J., Buechel, S. D., Cameron, R., Cappelle, K., Carolan, J. C., Christiaens, O., Ciborowski, K. L., Clarke, D. F., Colgan, T. J., Collins, D. H., Cridge, A. G., Dalmay, T., Dreier, S., du Plessis, L., Duncan, E., Erler, S., Evans, J., Falcon, T., Flores, K., Freitas, F. C., Fuchikawa, T., Gempe, T., Hartfelder, K., Hauser, F., Helbing, S., Humann, F. C., Irvine, F., Jermiin, L. S., Johnson, C. E., Johnson, R. M., Jones, A. K., Kadowaki, T., Kidner, J. H., Koch, V., Köhler, A., Kraus, F. B., Lattorff, H. M. G., Leask, M., Lockett, G. A., Mallon, E. B., Antonio, D. S. M., Marxer, M., Meeus, I., Moritz, R. F., Nair, A., Näpflin, K., Nissen, I., Niu, J., Nunes, F. M., Oakeshott, J. G., Osborne, A., Otte, M., Pinheiro, D. G., Rossié, N., Rueppell, O., Santos, C. G., Schmid-Hempel, R., Schmitt, B. D., Schulte, C., Simões, Z. L., Soares, M. P., Swevers, L., Winnebeck, E. C., Wolschin, F., Yu, N., Zdobnov, E. M., Aqrawi, P. K., Blankenburg, K. P., Coyle, M., Francisco, L., Hernandez, A. G., Holder, M., Hudson, M. E., Jackson, L., Jayaseelan, J., Joshi, V., Kovar, C., Lee, S. L., Mata, R., Mathew, T., Newsham, I. F., Ngo, R., Okwuonu, G., Pham, C., Pu, L.-L., Saada, N., Santibanez, J., Simmons, D., Thornton, R., Venkat, A., Walden, K. K., Wu, Y.-Q., Debyser, G., Devreese, B., Asher, C., Blommaert, J., Chipman, A. D., Chittka, L., Fouks, B., Liu, J., O'Neill, M. P., Sumner, S., Puiu, D., Qu, J., Salzberg, S. L., Scherer, S. E., Muzny, D. M., Richards, S., Robinson, G. E., Gibbs, R. A., Schmid-Hempel, P., and Worley, K. C. 2015. The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biology*, 16(1): 76.

Schultz, M. D., Schmitz, R. J., and Ecker, J. R. 2012. 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. *Trends in Genetics*, 28(12): 583–585.

Smeets, P. and Duchateau, M. J. 2003. Longevity of *Bombus terrestris* workers (Hymenoptera: Apidae) in relation to pollen availability, in the absence of foraging. *Apidologie*, 34(4): 333–337.

Smith, N. M. A., Yagound, B., Remnant, E. J., Foster, C. S. P., Buchmann, G., Allsopp, M. H., Kent, C. F., Zayed, A., Rose, S. A., Lo, K., Ashe, A., Harpur, B. A., Beekman, M., and Oldroyd, B. P. 2020. Paternally-biased gene expression follows kin-selected predictions in female honey bee embryos. *Molecular Ecology*, 29(8): 1523–1533.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. 2011. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, 6(7).

Tran, H., Porter, J., Sun, M.-a., Xie, H., and Zhang, L. 2014. Objective and Comprehensive Evaluation of Bisulfite Short Read Mapping Tools. *Advances in Bioinformatics*, 2014: 1–11.

Vaisvila, R., Ponnaluri, V. C., Sun, Z., Langhorst, B. W., Saleh, L., Guan, S., Dai, N., Campbell, M. A., Sexton, B. S., Marks, K., Samaranayake, M., Samuelson, J. C., Church, H. E., Tamanaha, E., Corrêa, I. R., Pradhan, S., Dimalanta, E. T., Evans, T. C., Williams, L., and Davis, T. B. 2021. Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome Research*, 31(7): 1280–1289.

718 Vojta, A., Dobrinić, P., Tadić, V., Bočkor, L., Korać, P., Julg, B., Klasić, M., and Zoldoš, V. 2016.
719 Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research*,
720 44(12): 5615–5628.

721 Wallberg, A., Bunikis, I., Pettersson, O. V., Mosbech, M.-B., Childers, A. K., Evans, J. D., Mikheyev,
722 A. S., Robertson, H. M., Robinson, G. E., and Webster, M. T. 2019. A hybrid de novo genome
723 assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genomics*,
724 20(1): 275.

725 Walsh, A. T., Triant, D. A., Le Tourneau, J., Shamimuzzaman, M., and Elsik, C. 2021. Hymenoptera
726 Genome Database: new genomes and annotation datasets for improved go enrichment and
727 orthologue analyses. *Nucleic Acids Research*, page gkab1018.

728 Wang, H., Liu, Z., Wang, Y., Ma, L., Zhang, W., and Xu, B. 2020. Genome-Wide Differential DNA
729 Methylation in Reproductive, Morphological, and Visual System Differences Between Queen Bee
730 and Worker Bee (Apis mellifera). *Frontiers in Genetics*, 11: 770.

731 Wang, X., Werren, J. H., and Clark, A. G. 2016. Allele-Specific Transcriptome and Methylome
732 Analysis Reveals Stable Inheritance and Cis-Regulation of DNA Methylation in Nasonia. *PLOS*
733 *Biology*, 14(7): e1002500.

734 Weiner, S. A., Galbraith, D. A., Adams, D. C., Valenzuela, N., Noll, F. B., Grozinger, C. M., and Toth,
735 A. L. 2013. A survey of DNA methylation across social insect species, life stages, and castes reveals
736 abundant and caste-associated methylation in a primitively social wasp. *Naturwissenschaften*,
737 100(8): 795–799.

738 Wu, X., Galbraith, D. A., Chatterjee, P., Jeong, H., Grozinger, C. M., and Yi, S. V. 2020. Lineage
739 and Parent-of-Origin Effects in DNA Methylation of Honey Bees (*Apis mellifera*) Revealed by
740 Reciprocal Crosses and Whole-Genome Bisulfite Sequencing. *Genome Biology and Evolution*,
741 12(8): 1482–1492.

742 Xu, G., Lyu, H., Yi, Y., Peng, Y., Feng, Q., Song, Q., Gong, C., Peng, X., Palli, S. R., and Zheng, S.
743 2021. Intragenic DNA methylation regulates insect gene expression and reproduction through the
744 MBD/Tip60 complex. *iScience*, 24(2): 102040.

745 Yagound, B., Smith, N. M. A., Buchmann, G., Oldroyd, B. P., and Remnant, E. J. 2019. Unique
746 DNA Methylation Profiles Are Associated with cis-Variation in Honey Bees. *Genome Biology*
747 *and Evolution*, 11(9): 2517–2530.

748 Yagound, B., Remnant, E. J., Buchmann, G., and Oldroyd, B. P. 2020. Intergenerational transfer
749 of DNA methylation marks in the honey bee. *Proceedings of the National Academy of Sciences*,
750 117(51): 32519–32527.

751 Yi, Y., Liu, Y.-B., Barron, A. B., and Zeng, Z.-J. 2021. Effects of commercial queen rearing methods
752 on queen fecundity and genome methylation. *Apidologie*, 52(1): 282–291.

753 Ziller, M. J., Hansen, K. D., Meissner, A., and Aryee, M. J. 2015. Coverage recommendations for
754 methylation analysis by whole-genome bisulfite sequencing. *Nature Methods*, 12(3): 230–232.