

RESEARCH

Open Access



DNA methylation is maintained with high fidelity in the honey bee germline and exhibits global non-functional fluctuations during somatic development

Keith D. Harris¹ , James P. B. Lloyd^{2,3} , Katherine Domb¹ , Daniel Zilberman^{4*} and Assaf Zemach^{1*}

Abstract

Background: DNA methylation of active genes, also known as gene body methylation, is found in many animal and plant genomes. Despite this, the transcriptional and developmental role of such methylation remains poorly understood. Here, we explore the dynamic range of DNA methylation in honey bee, a model organism for gene body methylation.

Results: Our data show that CG methylation in gene bodies globally fluctuates during honey bee development. However, these changes cause no gene expression alterations. Intriguingly, despite the global alterations, tissue-specific CG methylation patterns of complete genes or exons are rare, implying robust maintenance of genic methylation during development. Additionally, we show that CG methylation maintenance fluctuates in somatic cells, while reaching maximum fidelity in sperm cells. Finally, unlike universally present CG methylation, we discovered non-CG methylation specifically in bee heads that resembles such methylation in mammalian brain tissue.

Conclusions: Based on these results, we propose that gene body CG methylation can oscillate during development if it is kept to a level adequate to preserve function. Additionally, our data suggest that heightened non-CG methylation is a conserved regulator of animal nervous systems.

Keywords: Epigenetics, DNA methylation, Gene body methylation, Non-CG methylation, DNMT3, Germline

Background

Cytosine methylation is an ancient DNA modification that regulates the functioning of eukaryotic genomes [1, 2]. Methylation can be epigenetically inherited, most classically through the action of the Dnmt1 methyltransferase that maintains methylation within symmetrical CG dinucleotides in plant and animal genomes [3]. Inheritance of DNA methylation patterns across generations is robust and well-established in plants, but variable and controversial in mammals [4–6]. Plants and animals also

have Dnmt3-family methyltransferases that establish new methylation patterns and contribute to the maintenance of existing ones [1]. Methylation in animal genomes has long been thought to exist exclusively within the CG context, but recent work has revealed the presence of Dnmt3-catalyzed non-CG methylation in specific mammalian cell types and tissues, particularly in the brain [7, 8].

DNA methylation represses transposable elements in plants, vertebrates, fungi, and likely other species [9–11]. Methylation of promoters and other gene regulatory elements also generally causes transcriptional silencing, and this type of methylation is modulated to regulate gene expression and development in plants and vertebrates [2, 12]. More enigmatic targets of methylation are the transcribed portions of genes, a phenomenon known as

*Correspondence: daniel.zilberman@jic.ac.uk; assafze@tauex.tau.ac.il

¹ School of Plant Sciences and Food Security, Tel-Aviv University, 69978 Tel-Aviv, Israel

⁴ Department of Cell and Developmental Biology, John Innes Center, Norwich, UK

Full list of author information is available at the end of the article



gene body methylation [13]. In animals and flowering plants, gene body methylation is enriched in the exons of highly conserved genes that are ubiquitously and moderately transcribed, e.g. housekeeping genes that are constitutively expressed in all cell types and under diverse conditions [14–28]. These associations suggest a similar biological function, and possibly a common evolutionary origin, for gene body methylation in animals and plants. Extensive gene body methylation in the green algae *Chlorella variabilis* and *Klebsormidium nitens* [11, 29] also suggests a common evolutionary origin. Other findings, including the lack of gene body methylation in early-branching land plants and the linkage of genic methylation with a plant-specific DNA methyltransferase, support a convergent evolutionary model, in which gene body methylation evolved separately in animals and plants [30–33].

As a robust epigenetic feature that targets thousands of coding sequences in many eukaryotic genomes [9, 10], gene body methylation is expected to have an important function [34], especially because cytosine methylation is known to be mutagenic and should therefore be disfavored in coding sequences [35, 36]. Since its discovery, extensive efforts have been made to investigate the molecular and developmental roles of gene body methylation. Gene body methylation was found to fluctuate during animal and plant development [37–40], and several studies have linked artificial or developmental changes in gene body methylation to altered expression of particular genes [41–43]. Gene body methylation was also reported to regulate the splicing of individual genes [43–46]. However, DNA methyltransferase mutants in which methylation within gene bodies is nearly eliminated show no obvious global expression changes of body-methylated genes [28, 31, 32, 47], suggesting that gene body methylation in general either functions downstream of transcription [32, 48] or that its effects on transcription are subtle and difficult to detect by standard RNA sequencing techniques [34]. For example, the association between gene body methylation and constitutive gene expression could imply a role in homeostatic regulation of transcription, such as transcriptional stabilization or repression of unstable aberrant transcripts [49, 50]. Indeed, Dnmt3-mediated methylation of actively-transcribed genes was recently shown to inhibit aberrant intragenic transcriptional initiation in mouse cells [51], and gene body methylation has been linked with suppression of aberrant antisense transcripts in *Arabidopsis* [52].

Developmental or experimental changes in gene body methylation within vertebrate and plant genomes occur in the context of altered methylation in transposons and gene regulatory sequences [53, 54], which complicates assignment of functional outcomes specifically to

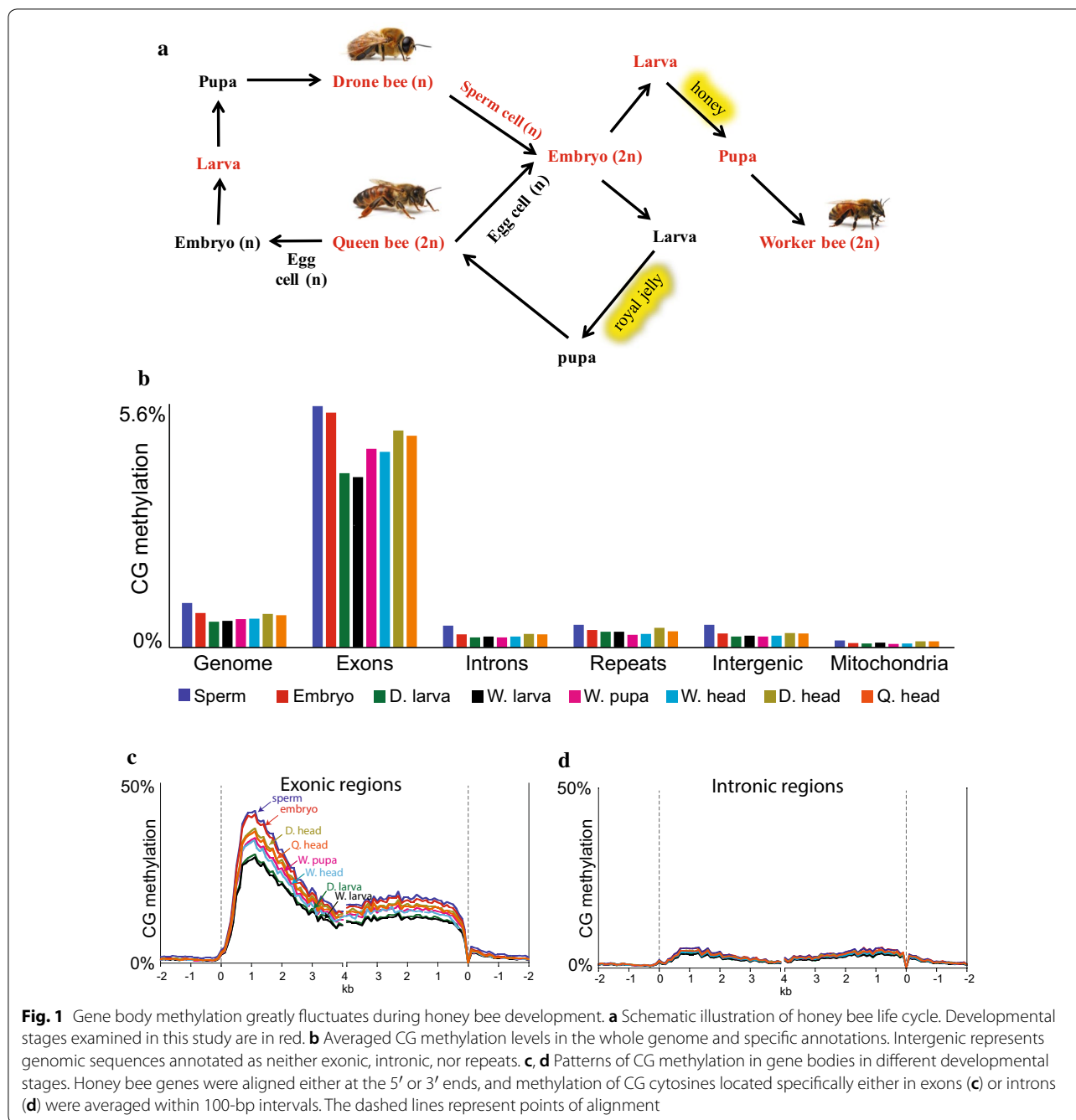
changes in gene body methylation. For this reason, honey bees and other invertebrates, in which methylation is specifically targeted to gene bodies, are highly useful models that provide a direct link between genic methylation and phenotype [9, 14]. Knockdown of Dnmt3 in honey bees shifted the development of worker larvae into adult queens [55], implying that an artificial reduction of gene body methylation can alter development. However, it is still unclear whether gene body methylation actively participates in executing developmental pathways in bees and other insects. Dynamic changes in the level and distribution of methylation during development would be consistent with an active role in developmental regulation. Several studies have reported that gene body methylation patterns in honey bee and other insects differ between developmental stages [43, 56, 57]. However, the robustness and biological reproducibility of the reported changes have been recently questioned [58].

To investigate DNA methylation dynamics during insect development, we profiled and analyzed the methylomes and transcriptomes derived from eight honey bee developmental stages (Fig. 1a). Using these biologically replicated data sets, we found that although the average methylation level of most genes alters substantially during honey bee development, very few sequences are specifically methylated in a particular developmental stage. Furthermore, DNA methylation changes are highly correlated between developmental stages. Fidelity of CG methylation is highest in sperm cells and embryos, the latter also exhibiting highest expression of Dnmt1. We did not find a significant association between methylation changes and developmentally regulated gene expression or alternative splicing. These results indicate that although the global levels of gene body methylation fluctuate during honey bee development, the overall methylation patterns are robustly maintained, suggesting that the main function of gene body methylation is homeostatic. The high methylation fidelity in sperm and the reported stability of methylation in wasp hybrids [59] suggest that gene body methylation in bees and related insects is trans-generationally inherited through robust germline maintenance, whereas methylation efficiency is lower and more variable during somatic development. Finally, we identified Dnmt3-associated non-CG methylation in the gene bodies of adult honey bee heads, suggesting that this may be an ancient and conserved feature of animal neurological development [7].

Results

Gene body methylation fluctuates during honey bee development

To investigate the dynamics of DNA methylation during honey bee development, we profiled the genome-wide methylomes of eight developmental stages: sperm,



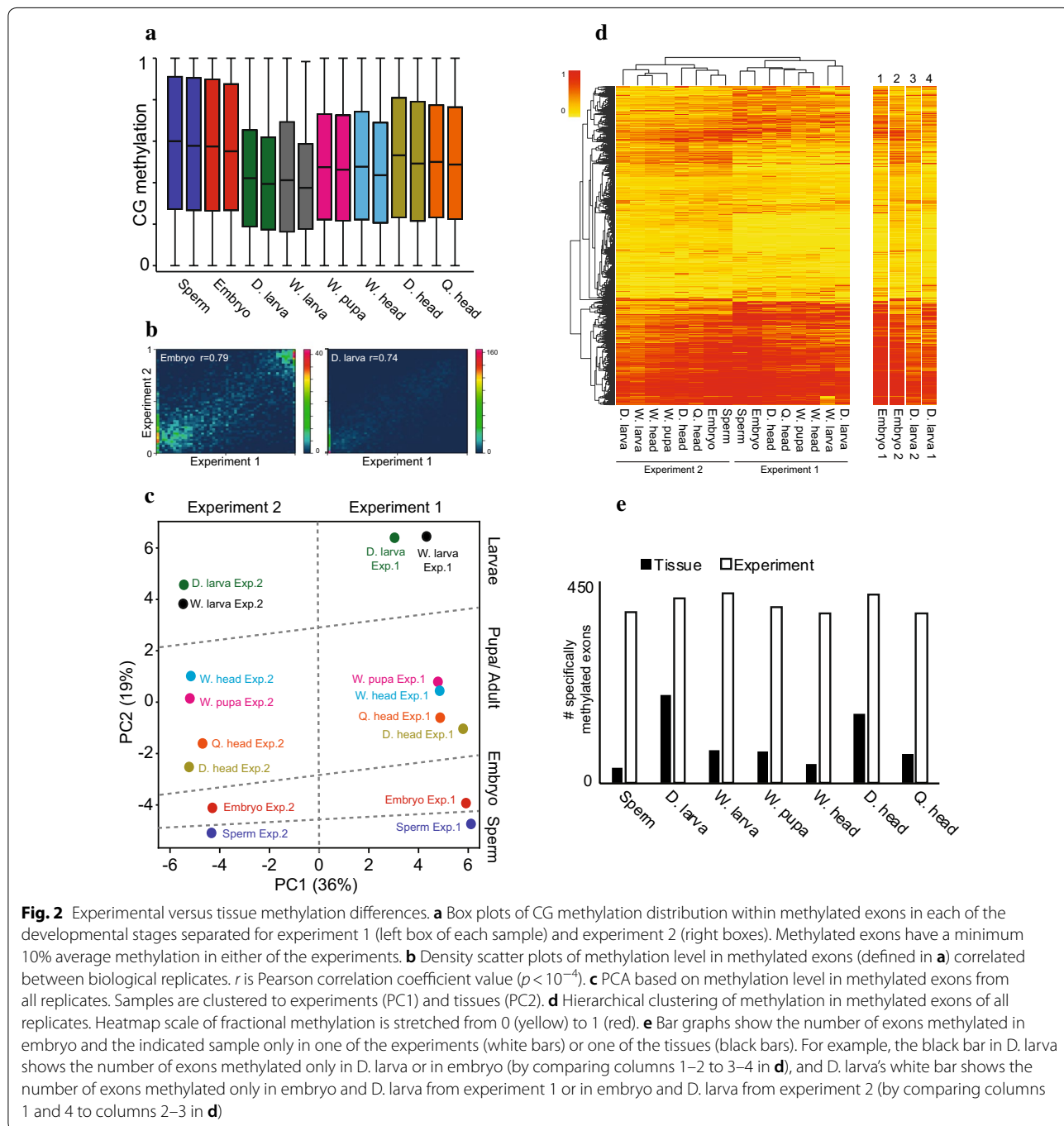
worker embryo, worker larva, drone larva, worker pupa, worker head, drone head, and queen head (Fig. 1a). We conducted two independent experiments (experiments 1 and 2, Additional file 1: Figure S1A), collecting material from each of the developmental stages about 3 weeks apart. Quantification of the averaged CG methylation level of the two biological replicates, either separately (Additional file 1: Figure S1A) or combined (Fig. 1b), in the whole nuclear genome, as well as within specific genetic elements, including exons, introns, and repeats,

revealed two main characteristics of DNA methylation in honey bee. First, methylation is targeted specifically to exonic sequences in all developmental stages (Fig. 1b–d and Additional file 1: Figure S1). Average CG methylation for the whole genome is around 1%, within all exons is about 5%, and within the set of methylated exons (~25% of all exons) reaches approximately 40% (Fig. 1b and Additional file 1: Figure S1A). Second, the genomic average of gene body methylation within exons fluctuates during honey bee development. Exon methylation is high

in the sperm and embryo, drops by about 25% in drone and worker larvae, and then increases to an intermediate level in pupa and adult heads (Fig. 1b and Additional file 1: Figure S1A). Meta-analyses of averaged exonic or intronic methylation across honey bee genes confirmed the above global methylation dynamics, further showing that exonic methylation changes across the entire gene sequence (Fig. 1c, d and Additional file 1: Figure S1B).

Developmental methylation differences are consistent between diverse biological replicates

Meta-analyses of exon methylation showed consistent methylation levels between biological replicates (Fig. 2a). However, methylation of individual exons shows only moderate correlation between biological replicates, with Pearson’s correlation coefficients ranging between 0.65 and 0.79 (Fig. 2b and Additional file 1: Figure S2). Principal component analysis



(PCA) separated methylation datasets by experiment in the first principal component (PC1), grouping datasets of experiment 1 away from those of experiment 2 (Fig. 2c). Similar separation by experiment was produced by hierarchical clustering (Fig. 2d). PCA analysis separated datasets by tissue along the second principal component (PC2), with tissues ordered roughly by the amount of overall methylation (compare Fig. 2a, c). Hierarchical clustering also grouped datasets by tissue methylation level, consistently separating samples with high methylation (sperm and embryo) from those with low methylation (larva; Fig. 2d). Comparison of methylation patterns in embryo and another tissue across experiments revealed many more exons specifically methylated in each experiment than in any tissue (Fig. 2e). These results demonstrate a greater divergence of methylation patterns between experiments than between tissues and illustrate the importance of biological replication for assessing tissue-specific methylation patterns in insects, as highlighted by an earlier study [58]. Biological replicates were collected at two separate time points (several weeks apart); thus, the experiment effect (Fig. 2c, d) could reflect a biological difference, e.g., seasonality. Comparison of methylation patterns between any tissue in experiment 1 and another in experiment 2 would identify many differentially methylated exons and create a false sense of extensive tissue-specific methylation. For the remainder of the study, a site is scored as differentially methylated between tissues only when comparisons within experiments 1 and 2 are in agreement and statistically significant.

Genic CG methylation changes globally but not locally during honey bee development

Differential methylation between tissues could result from global and moderate fluctuations within many genic sequences, from strong fluctuations in a limited set of sequences, or from a combination of these mechanisms.

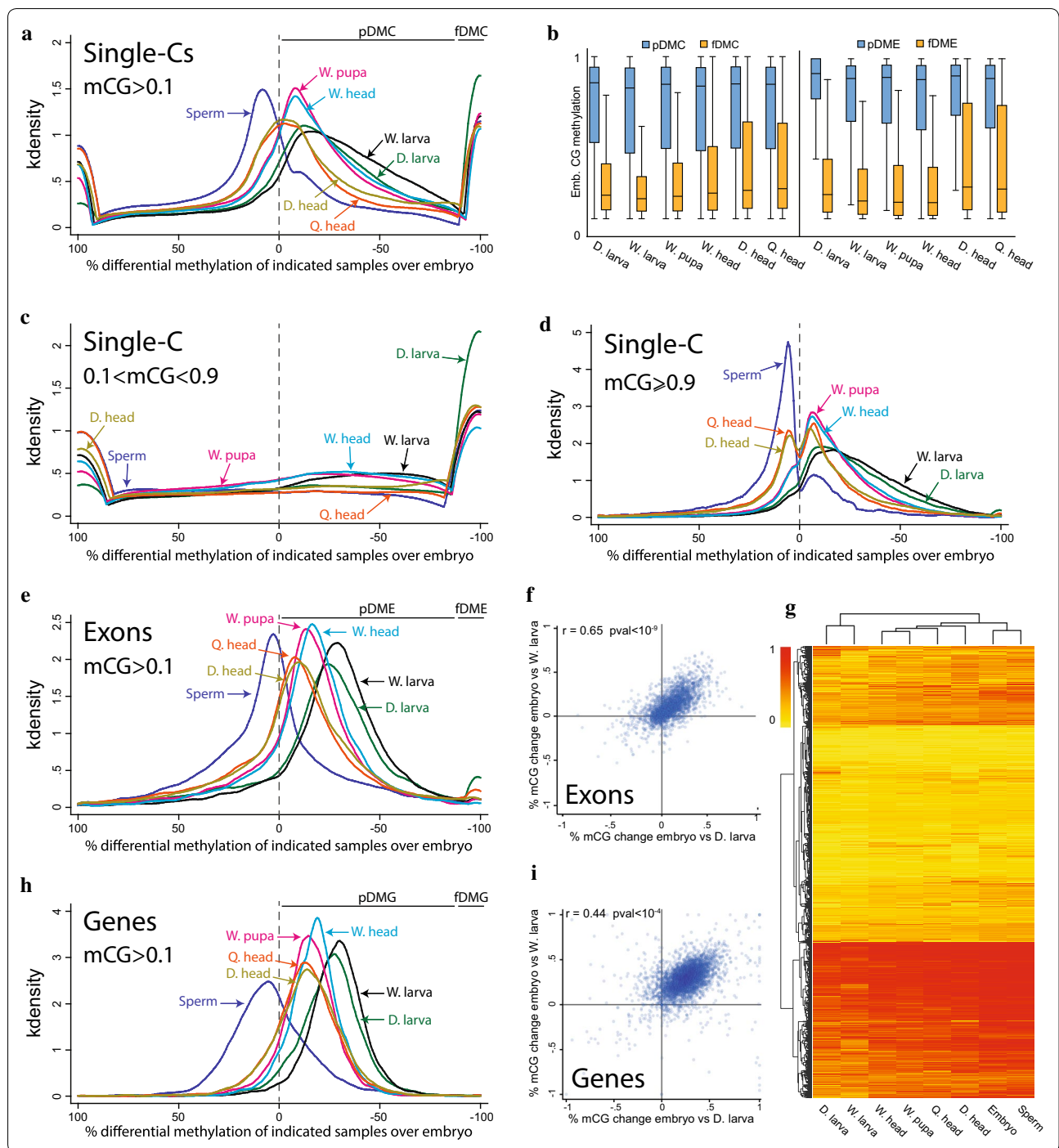
To distinguish between these possibilities, we calculated the percentage change in methylation between the embryo and each of the other biological samples in three levels of resolution: individual cytosines, exons, and entire genes.

Density plots of percent-methylation-change show that most individual methylated CG sites are hypermethylated in the sperm over the embryo, whereas in all postembryonic samples, most methylated CGs are hypomethylated compared to embryonic CGs (Fig. 3a). Furthermore, these plots show that differentially methylated cytosines (DMCs) can be divided into two groups: partial DMCs (<90% change) and full DMCs ($\geq 90\%$ change; Fig. 3a). pDMCs outnumber fDMCs in all developmental pairwise comparisons (1.5- to 4.8-fold) (Additional file 1: Figure S3A). 22–44% of fDMCs between embryo and other tissues are shared between biological replicates (Additional file 1: Figure S3B), suggesting that at least some sites do lose all methylation during development. Nevertheless, we found that the level of CG methylation in embryos is much higher in pDMCs cytosines (median level of 95%) than in fDMCs cytosines (median level $\sim 20\%$; Fig. 3b). Consistently, almost all fDMCs are methylated below 90% in the embryo, whereas pDMCs have methylation $\geq 90\%$ (Fig. 3c, d). These results indicate that highly methylated embryonic cytosines are usually partially hypomethylated and very rarely lose all their methylation during development.

As gene body methylation in honey bee is localized specifically in exons (Fig. 1b–d), we next examined DNA methylation dynamics across individual exons during bee development. Most exons are partially hypomethylated in postembryonic samples, with a median hypomethylation ranging from 8% in adult heads to 29% in larval samples (Fig. 3e and Additional file 1: Figure S3C). Exon methylation changes between embryos and post-embryonic samples are correlated (Fig. 3f and Additional file 1: Figure S3C), so that a given exon tends to have an altered methylation level in the same direction and to a similar

(See figure on next page.)

Fig. 3 Gene body methylation is robustly maintained during honey bee development. **a** Kernel density plots of percent-methylation-change between averaged methylation in embryo versus indicated samples calculated for single cytosines (Single-Cs). Negative and positive numbers in the x-axis indicate greater and lower methylation in embryo and the other sample, respectively. Partial and fully differentially methylated cytosines (pDMC/fDMC) were considered if the percent-methylation-change was lower and higher than 90%, respectively. **b** Box plots of averaged embryonic methylation in single cytosines (left) and exons (right) derived from partially and fully differentially methylated sites between embryo and indicated samples. **c, d** Kernel density plots of percent-methylation-change between averaged methylation of single-Cs in embryo versus indicated samples, of cytosines derived from lowly methylated in embryo (**c**) or highly methylated in embryo (**d**). **e** Kernel density plots of percent-methylation-change between averaged methylation in embryo versus indicated samples calculated for single exons. Partial and fully differentially methylated exons (pDME/fDME) were annotated as in **a**. **f** Scatter plot of percent-methylation-change between embryo/D. larva and embryo/W. larva in exons. r is Pearson correlation coefficient value. **g** Hierarchical clustering heatmap showing average methylation in individual exons in the different biological samples. Scale bar indicates methylation level. **h** Kernel density plots of percent-methylation-change between averaged methylation in embryo versus indicated samples calculated for single genes. Partial and fully differentially methylated genes (pDMG/fDMG) were annotated as in **a**. **i** Scatter plot of percent-methylation-change between embryo/D. larva and embryo/W. larva in genes



extent in all post-embryonic samples. This result is also illustrated in hierarchical clustering of methylation in exons, which segregated samples according to their overall exonic methylation level (Fig. 3g). On average, only 2.9% of the differentially methylated exons (181 exons) had completely lost their methylation in postembryonic

samples, in contrast to an average of 6240 exons that were partially hypomethylated (Additional file 1: Figure S3A). Similar to individual cytosines, we found that fully differentially methylated exons are much less methylated in the embryo (median level of 21–27%) than exons with partial differential methylation (median level of 83–85%;

Fig. 3b), suggesting that lowly methylated CG sites and exons are more easily fully demethylated than highly methylated sites and exons.

Finally, we found that very few genes (3–9 genes) completely lose their methylation during development (Additional file 1: Figure S3A), whereas most genes show methylation reduction between 13 and 30% (Fig. 3h and Additional file 1: Figure S3C). Similar to exons, genic methylation changes between embryos and post-embryonic samples are correlated (Fig. 3i and Additional file 1: Figure S3D), so that a given gene tends to change methylation in the same direction and to a similar extent in all post-embryonic samples.

In summary, our DNA methylation analyses show that global gene body methylation fluctuates during honey bee development, but specific differential methylation patterns at the level of complete genes or exons are rare. The patterns of gene body methylation in honey bee are, therefore, very robust and are largely maintained in distinct tissues throughout the bee life cycle.

Gene body methylation dynamics are not generally associated with differential gene expression during bee development

Gene body methylation in honey bee, as well as in other animals and plants, was previously shown to be anticorrelated with gene responsiveness and to be enriched at ubiquitously expressed genes [14–17, 21–28]. We confirmed this finding using transcriptional profiles from all developmental stages examined in this study (Fig. 4a). Consistently, most differentially expressed genes (DEGs) between developmental stages are not methylated (Fig. 4b and Additional file 1: Figure S4A). In addition, unmethylated genes present a broader range of differential expression than methylated genes (Fig. 4b and Additional file 1: Figure S4A). The number of overlapping DEGs and differentially methylated genes (DMGs) is smaller (in most cases much smaller) than expected by random chance in all pairwise comparisons of the examined developmental stages (column 9 in Table 1). Therefore, both methylated and differentially methylated genes are underrepresented among developmentally regulated genes. Taken together with the above-described behavior of DNA methylation during bee development, these results suggest that gene body methylation does not function primarily to regulate developmentally-specialized transcriptional programs.

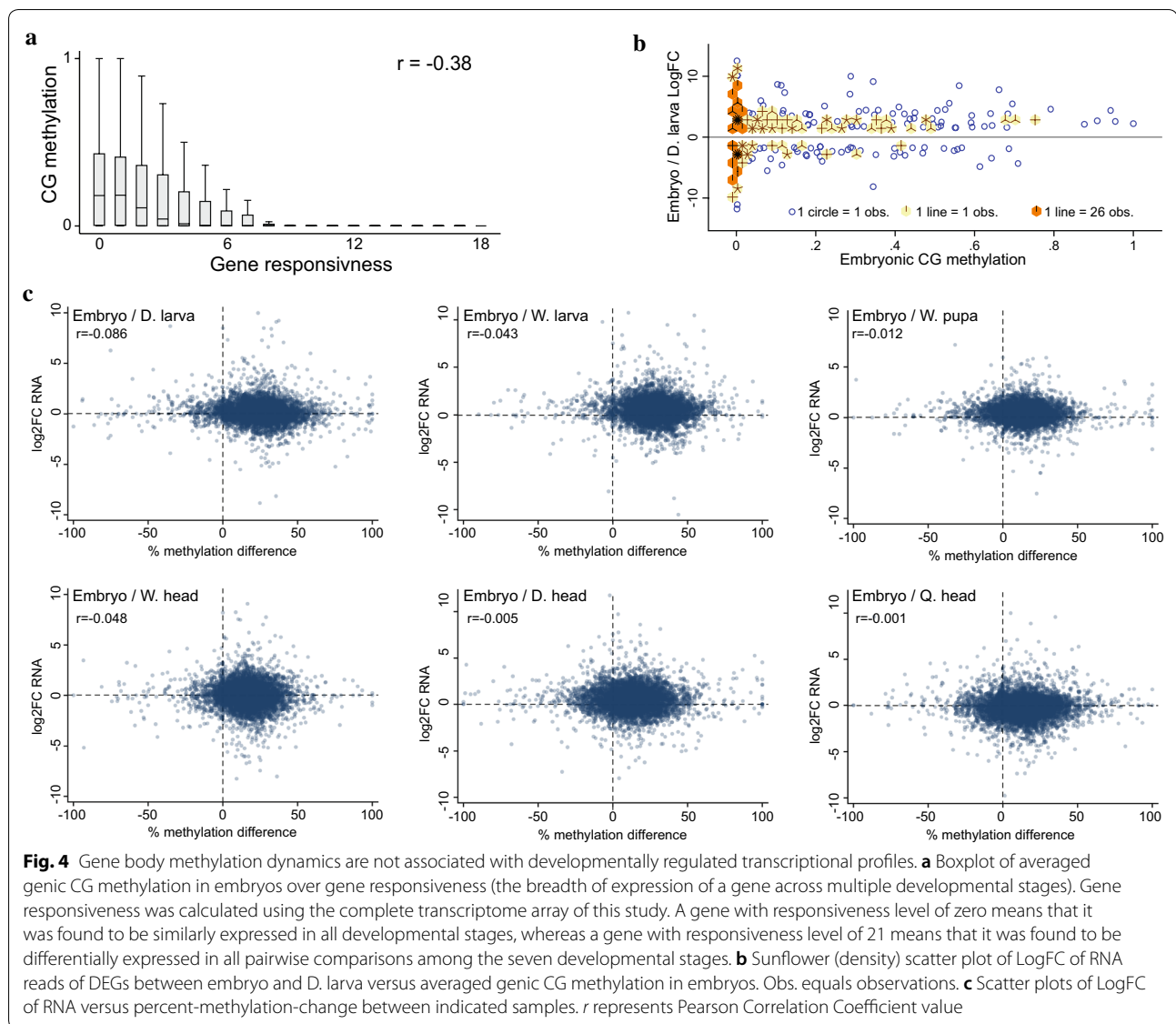
However, the global fluctuations in methylation during honey bee development might nonetheless influence developmentally-regulated transcription of methylated genes. To test this hypothesis, we focused our analyses specifically on differentially expressed methylated genes (DEG-me/column 7 in Table 1). We found that for the majority of the pairwise comparisons, including the

ones with the highest methylation dynamics (i.e., embryo vs. larvae), the number of overlapping genes between DEGs and DMGs is just as expected by random chance if only methylated genes are considered (column 10 in Table 1). Thus, DEGs are not generally enriched among DMGs. Four pairwise comparisons did show a significantly enriched overlap between DMGs and DEGs, and all of these contained at least one adult bee head (Table 1, marked with #). To further test the relationship between DEGs and DMGs, we correlated the levels of methylation and RNA changes between all pairs of developmental stages. This analysis found no association between methylation and expression dynamics (Fig. 4c), including between the DEG/DMG enriched overlap samples (Additional file 1: Figure S4B). Overall, our results indicate that gene body methylation dynamics in honey bee are not generally associated with changes in gene expression. Neurological tissues may be an exception, with some evidence that differential methylation is linked to differential expression (Table 1).

Gene body methylation dynamics are not associated with differential RNA splicing during bee development

To investigate the association between intragenic methylation and RNA splicing in honey bee, we examined the distributions of various alternative splicing (AS) events—exon skipping, intron retention, and alternative 5'/3' splicing—within methylated gene sequences across developmental stages. These distribution patterns show that no AS types correlate with the most highly methylated genic regions (Fig. 5a). Whereas methylation is particularly enriched within the second kb of gene sequences, all the examined AS profiles are relatively depleted from this region and enhanced either upstream or downstream (Fig. 5a). One interpretation of this result is that methylation suppresses AS, consistent with previous findings showing that genic methylation enhances the inclusion of alternatively spliced regions [44].

Next, we investigated whether developmentally associated gene body methylation dynamics are connected to differential levels of alternative splicing. For this purpose, we measured the number of differentially alternatively spliced skipped exons (DSEs) among all methylated exons, and compared this with the number of differentially methylated exons (DMEs) in paired developmental stages (Fig. 5b). DSEs overlapped DMEs significantly less than expected by random chance, with typically half of the expected number of DSEs corresponding to DMEs (Fig. 5b). This result indicates that the methylation level of most DSEs is stable during honey bee development. We next compared the levels of methylation of alternatively spliced exons and their level of inclusion in mature mRNA. To



this end, we correlated percent-methylation-change of exons between two developmental stages against the delta of percent-spliced-in (PSI) of these exons. This analysis did not reveal a correlation of appreciable magnitude between methylation and developmentally regulated PSI changes (Fig. 5c). Overall, our results do not indicate that global fluctuations in gene body methylation promote alternative splicing during honey bee development.

DNA methylation is distinctly regulated in the soma and germline

To explain the molecular mechanism of gene body methylation dynamics during honey bee development, we examined the expression of honey bee DNA methyltransferases. The honey bee genome encodes two

Dnmt1 CG methylation maintenance enzymes and one Dnmt3 de novo methyltransferase [60]. Our RNA-seq data showed that all honey bee Dnmts are expressed and differentially regulated during development (Fig. 6a). Both Dnmt1s are most highly expressed in the embryo (Fig. 6a). In contrast, Dnmt3 RNA level is the lowest in the embryo and consistently increases during development, being most enriched in adult heads (Fig. 6a). The highest expression level of both Dnmt1s in the embryo corresponds with the highest CG methylation level during this stage, whereas the increase in CG methylation in post-larva stages corresponds with an increase in Dnmt3 expression (Figs. 1a, 6a). Accordingly, the drop in CG methylation in the larva could be explained by passive DNA demethylation due to a decrease in Dnmt1 expression, whereas the recovery in

Table 1 Gene body methylation dynamics are not associated with developmentally regulated transcriptional profiles

Sample1	Sample2	All genes	meGenes	DMG	DEG-all	DEG-me	Overlap	Rep. factor	Rep. factor
W. head	W. pupa	13285	4437	1058	1941	204	63	0.4 ****	1.3*#
D. head	W. pupa	13285	4496	1438	1343	85	27	0.2 ****	1
Q. head	W. pupa	13285	4480	1522	2028	430	152	0.7 ****	1
Q. head	W. head	13285	4462	1535	1122	162	61	0.5 ****	1.1
W. larva	D. larva	13285	4318	1586	297	43	15	0.4 ***	0.9
Q. head	D. head	13285	4508	1589	878	140	62	0.6 ****	1.3*#
D. head	W. head	13285	4483	1624	486	42	15	0.3 ****	1
D. head	Embryo	13285	4589	1857	1796	261	113	0.5 ****	1.1
W. head	D. larva	13285	4417	1904	2129	192	89	0.3 ****	1.1
Q. head	D. larva	13285	4449	1911	2050	325	153	0.5 ****	1.1
W. pupa	D. larva	13285	4449	2045	1296	106	47	0.2 ****	1
Q. head	Embryo	13285	4584	2055	2331	417	210	0.6 ***	1.1*#
W. pupa	W. larva	13285	4392	2166	1585	174	86	0.3 ****	1
W. head	W. larva	13285	4359	2223	2462	339	169	0.4 ****	1
D. head	D. larva	13285	4480	2252	1450	100	47	0.2 ****	0.9
W. pupa	Embryo	13285	4576	2524	1170	157	92	0.4 ***	1.1
Q. head	W. larva	13285	4412	2528	2386	553	356	0.8 ****	1.1***#
D. head	W. larva	13285	4442	2538	1629	129	58	0.2 ****	0.8*
W. head	Embryo	13285	4561	2744	2285	308	199	0.4***	1.1
W. larva	Embryo	13285	4542	3200	1939	440	296	0.6***	1
D. larva	Embryo	13285	4577	3402	1612	216	165	0.4***	1

Table includes the names of the pairwise samples (columns 1–2), the number of all genes (column 3), number of methylated genes (>0.1 in either of the selected samples; column 4), number of differentially methylated genes (averaged genic mCG > 0.1 in either sample, Fisher Exact Test pval < 0.05 in both of biological replicates; column 5), number of differentially expressed genes among either all genes or within methylated ones (columns 6 and 7, respectively), number of overlapped genes between DMG and DEG (column 8), and the ratio between the number of actual overlapped genes with that of the expected one when considering all genes (column 9) or only methylated ones (column 10)

* Denote the strength of *p* values. Brackets above the table point to columns used for calculating relevant representation factors

Marks observations with a significant overlap between DMGs and DEGs among methylated genes

DNA methylation in the pupa and adult samples could be a result of increased Dnmt3 activity.

Consistent with the high expression of Dnmt1 enzymes in the embryo, this developmental stage shows more robust maintenance of CG methylation than any of the post-embryonic stages. Embryonic methylation at single CG cytosines shows a binary distribution: most cytosines are either not methylated or fully methylated (Fig. 6b). In comparison, cytosine methylation in postembryonic samples has a more gradual distribution with few completely methylated sites (Fig. 6b). The difference in the distribution and median level of CG methylation in the embryo versus post-embryonic samples is clear when only methylated cytosines are considered (Fig. 6c). Interestingly, these analyses also show an even stronger binary distribution of CG methylation in sperm cells, compared to the embryo (Fig. 6b), with the highest median methylation level (90%) among all samples (Fig. 6c). CG methylation in sperm is also the highest in complete exons containing single, double, and triple CG sites (which represent 84% of methylated exons) (Fig. 6d–f). Sperm and embryo have more fully methylated exons (>90%)

than partially methylated ones (>0.1 and ≤0.9), whereas adult somatic samples have the opposite ratio (Fig. 6g–i). Finally, by analyzing methylation level in single BS-seq molecules, we found sperm to have the highest frequency of fully methylated reads and lowest frequency of partially and unmethylated reads (Fig. 6j–l). Overall, these results suggest that methylation in sperm is more efficient than in somatic tissues, both globally among all methylated exons in the genome as well as locally within intra-exonic neighboring CG sites.

Further examination of the distribution of CG sites along genes, differentiated by their methylation level in sperm, revealed that highly methylated CGs (mCG ≥ 0.8) are enriched in upstream genic sequences, whereas lowly methylated CGs (0.05 < mCG < 0.8) are depleted from these regions (Fig. 6m). The distribution pattern of highly methylated CGs in sperm matches that of overall CG methylation in bee tissues (Fig. 1b). Thus, the efficiency of sperm methylation and the probability of inheriting a methylated site may shape bee methylation patterns over generations.

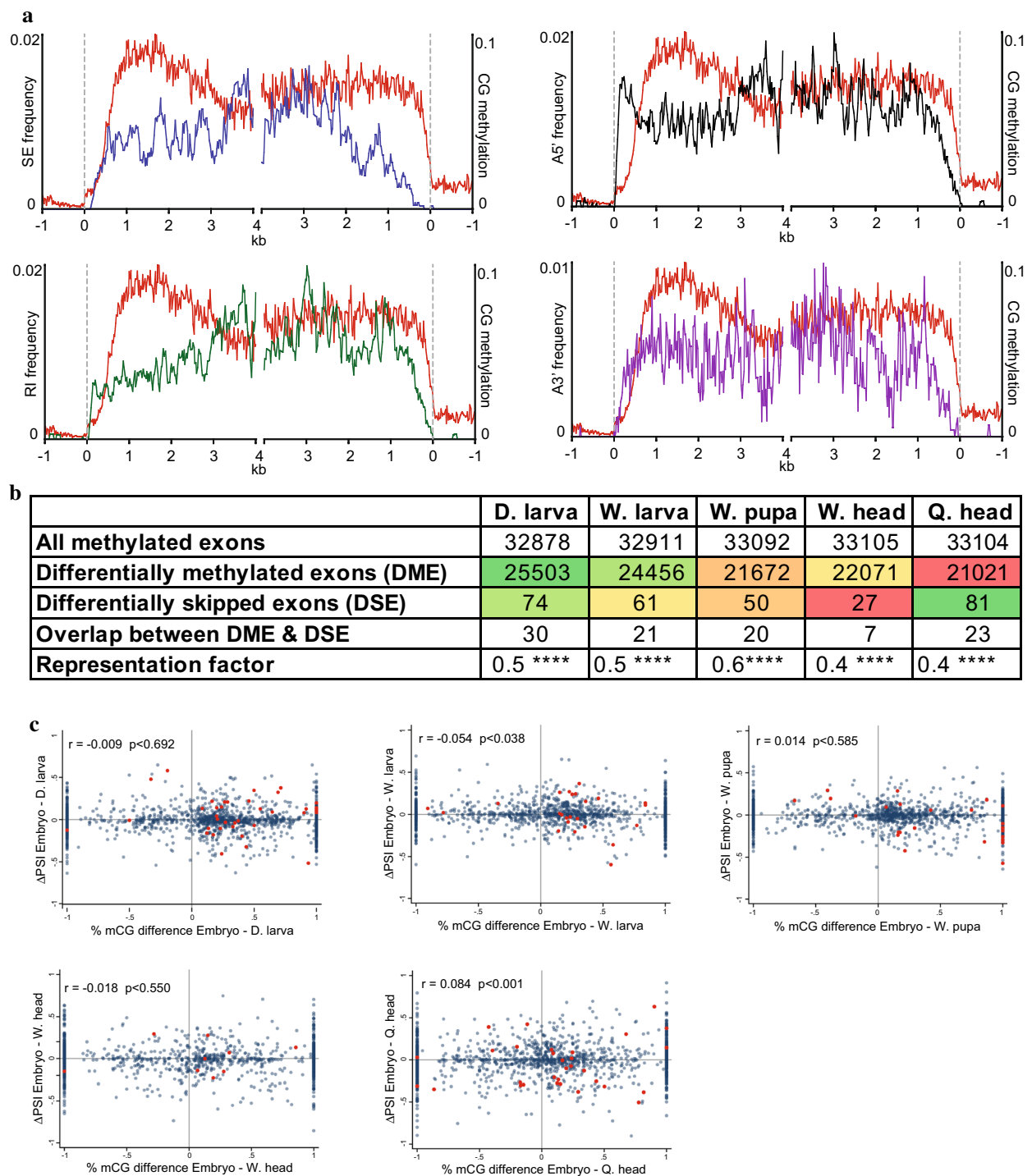
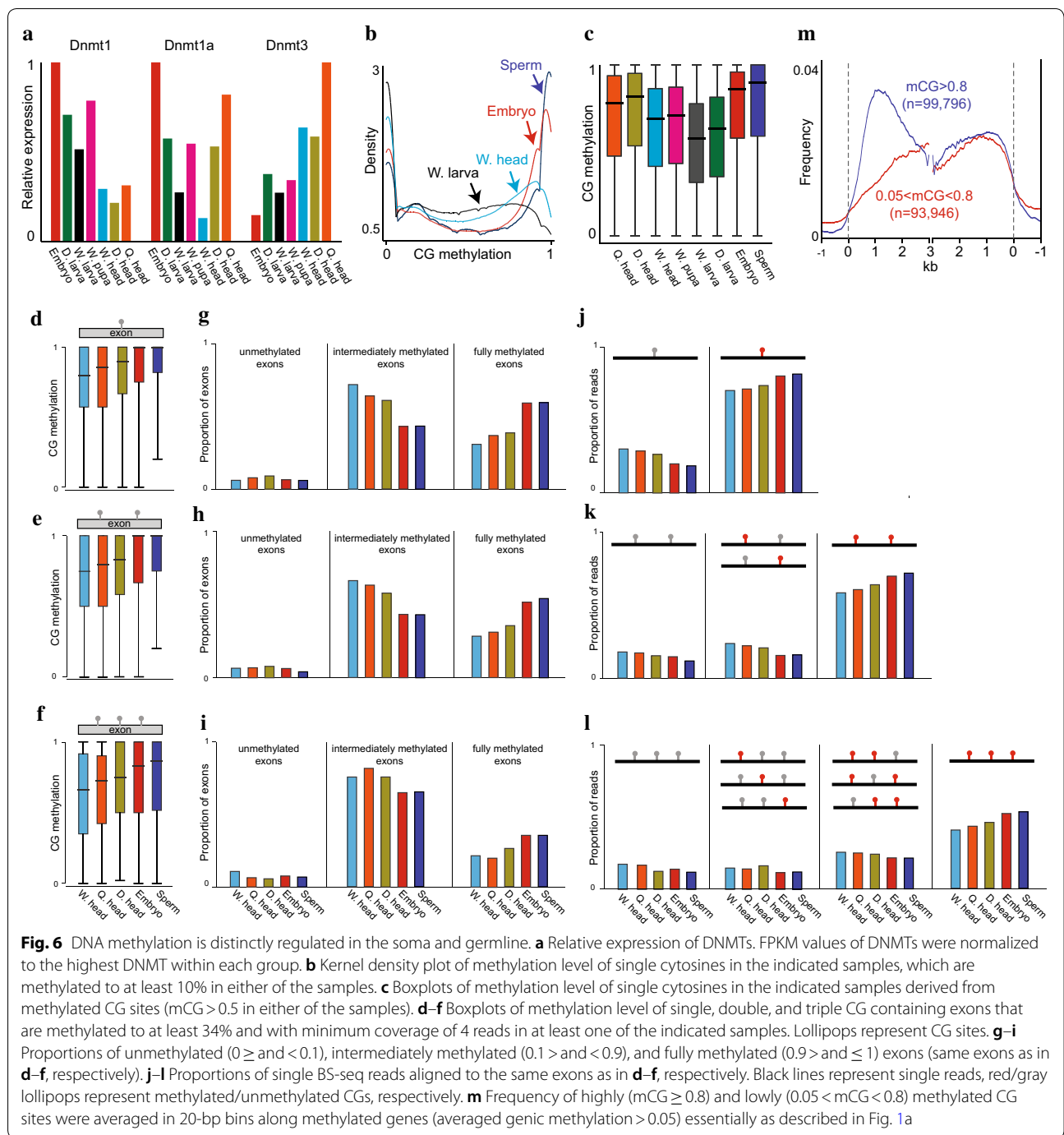


Fig. 5 Gene body methylation dynamics are not associated with developmentally regulated RNA splicing patterns. **a** Skipped exons (blue), retained introns (green) and alternative 5'/3' spliced sites (black/purple) frequencies were averaged in 20-bp bins along methylated genes (averaged genic methylation > 0.05) essentially as described in Fig. 1a. Averaged CG methylation profile is illustrated with red lines. **b** This table includes the number of all methylated exons (mCG > 0.05 in either embryo or the other biological sample (indicated on the top line of the columns), number of differentially methylated exons (DMEs; Fisher Exact Test $p < 0.05$ in both of biological replicates), number of differentially AS-skipped exons (DSEs), number of overlap exons between DMEs and DSEs, and representation factors, which is the ratio between the number of observed overlapped exons with that of the expected one. A representation factor < 1 means that it is under-representated. *Denote p value strength. **c** Scatter plots of delta PSI (Δ PSI) of exons between embryo and indicated samples versus percent-methylation-change of exons in the same pairwise samples. Blue dots include values of all methylated exons, and red dots include values of only differentially skipped and methylated exons (i.e., DSEs and DMEs)



Non-CG methylation is a conserved regulator of animal nervous systems

Whereas the absolute expression levels of Dnmt1 and Dnmt1a are comparable, the transcript level of Dnmt3 is on average 18 and 11 times higher than that of Dnmt1 and Dnmt1a, respectively (Fig. 7a). This differs from mammals, in which Dnmt1 expression is usually higher than that of Dnmt3 [61]. Although most methylation

in animals is concentrated in CG sites, recent reports showed that non-CG (CH, H=A, C or T) methylation also exists and is mostly targeted by Dnmt3 [7]. In mammals, CH Dnmt3-dependent methylation is particularly enriched in brain cells [7]. These findings prompted us to check whether the increase of Dnmt3 expression in honey bee heads is similarly associated with CH methylation. By examining methylation in genic sequences, we

(See figure on next page.)

Fig. 7 CW (CA or CT) methylation is enriched in adult honey bee heads. **a** Averaged FPKM values of DNMTs in the different biological samples. **b** Heads and W. larva CH methylation located specifically in exons was averaged in 100 bp bins along honey bee genes essentially as described in Fig. 1a. CH methylation of all developmental stages can be found in Additional file 1: Figure S5B. **c** Heads and W. larva CH methylation located specifically in introns was averaged in 100 bp bins along honey bee genes essentially as described in Fig. 1a. **d** Q. heads CG, CA, and CT methylation located specifically in exons were averaged in 100 bp bins along honey bee methylated (averaged genic CG methylation > 0.05) and unmethylated (averaged genic CG methylation < 0.005) genes essentially as described in Fig. 1a

discovered an increase in CH methylation in adult heads (Fig. 7b and Additional file 1: Figure S5A). CH methylation was mostly enriched in queen heads, as compared to drone and worker heads (Fig. 7b and Additional file 1: Figure S5A), in accordance with the highest Dnmt3 expression level found in this sample (Figs. 6a, 7a). The increased CH signal is not due to nucleotide polymorphism between the sequenced samples and the reference genome, i.e., CG sites in reads that align to CH sites in the genome, as the incidents of such events were lower than 5% in any of the samples and the lowest in queen heads (1.68%), which have the highest CH methylation signal (Additional file 1: Figure S5B). CH methylation was particularly enriched on cytosines adjacent to adenine (CA), and to a lesser extent to thymine (CT), similar to mammalian CH methylation context preferences [7]. CA methylated sites were not found to be further enriched within a particular motif, such as TACAC in glia and neuron cells (Additional file 1: Figure S5C) [7]. Finally, CW (W = A or T) methylation in the gene bodies of adult heads is specifically enriched in genes that also have CG methylation (Fig. 7d) and is localized in the same genic region (upstream exons) where CG methylation is highest (Fig. 7b, d), also similar to CH methylation in mammalian brain tissue [7]. Overall, while CH methylation in bee heads is low (~0.2%), its signal is continuous (~200 bp) and above the noise level, associated with Dnmt3 expression, and enriched in a particular context (CA), tissue (heads), and genes (CG methylated), indicating that it is a real biological signal and implying that increased CH methylation during brain development is an ancient epigenetic phenomenon shared by insects and mammals.

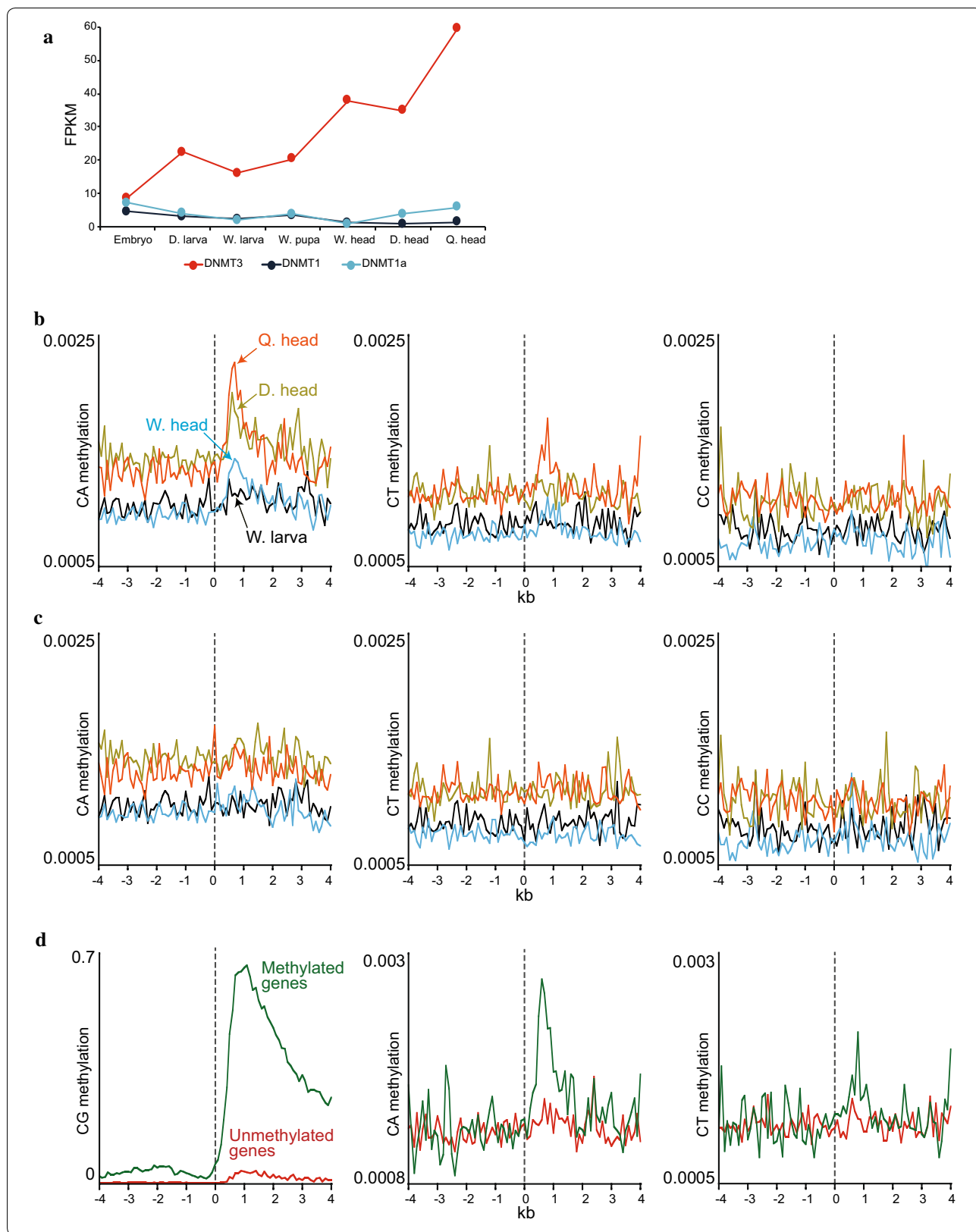
Discussion

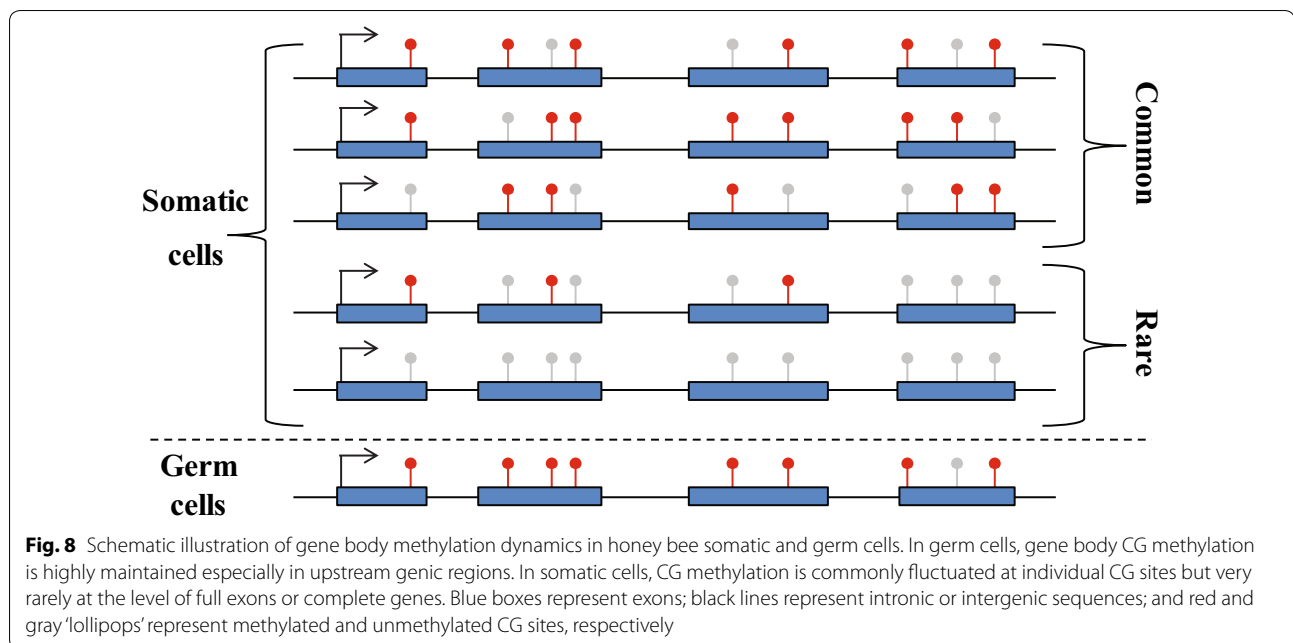
We show that gene body methylation globally fluctuates during honey bee development, yet the patterns of methylation remain essentially unchanged from sperm to adult (Figs. 1, 2, and 8). Developmentally robust patterns of gene body methylation have been proposed in ants [58] and in the invertebrate chordate *Ciona intestinalis* [16], and are also the norm in plants [62–64], suggesting that developmental stability of gene body methylation may be ancient and conserved.

We find the highest levels of methylation in bee sperm, whereas methylation fidelity in somatic tissues is considerably lower (Fig. 6). More efficient maintenance of methylation in male and female gametes has also been reported in flowering plants [65, 66]. Plants have robust transgenerational inheritance of DNA methylation [4, 67, 68], which necessitates efficient preservation of methylation patterns in reproductive tissues. The high efficiency of DNA methylation in sperm suggests that DNA methylation is also transgenerationally inherited in bees. Consistent with this hypothesis, chromosomes in hybrid wasps faithfully maintain their ancestral methylation patterns [59]. Thus, transgenerational epigenetic inheritance of DNA methylation may be a common feature of invertebrates and plants.

There could be a number of molecular mechanisms regulating methylation fidelity during development. Our data suggest that alterations in the expression of DNMT1 and DNMT3 (Fig. 6a) could play a role in enhancing, or passively reducing, the efficiency of maintenance (DNMT1) and de novo (DNMT3) DNA methylation activities during honey bee development. Alternatively or additionally, methylation fidelity could be regulated biochemically via DNMT cofactors, e.g. UHRF1 [69], or through the involvement of active DNA demethylation pathways [70], which are yet to be explored in honey bee.

Except for adult heads, we did not find any significant associations between methylation dynamics and any transcriptional or posttranscriptional alterations among various developmental stages (Figs. 4 and 5, and Table 1). Developmentally regulated genes also have a strong tendency not to be methylated, as has been observed in many animal and plant species [14–28]. We, therefore, propose that for the most part, gene body methylation is not involved in regulating transcriptional changes during honey bee development. Instead, gene body methylation is likely to have a homeostatic function in maintaining normal gene expression, for example by suppressing intergenic transcript initiation [51]. Alternatively, genic methylation might be involved in genomic activities unrelated to gene expression, such as genome structure or integrity [32]. A possible explanation for the global gene body methylation fluctuations during development is that the highest levels of methylation are not required for proper function. This hypothesis is based on the finding





that evolutionary selection acts on entire regions of gene body methylation, rather than individual cytosines [71]. In somatic cells with limited division potential, imperfect maintenance of methylation may, therefore, preserve functionality if methylation of an entire exon or gene is maintained above a threshold level.

Our finding that differential methylation and gene expression are correlated in comparisons involving adult heads (Table 1) may relate to non-CG methylation. Bee CH methylation is specifically targeted to CW sites, associated with Dnmt3 expression, localized in CG methylated regions, and particularly enriched in adult heads (Fig. 7). These characteristics are similar to CH methylation specificities described in mammals, and CH methylation was shown to regulate transcription and differentiation of human neurons [7]. Similar regulation may occur in honey bee, and would be consistent with the finding that knockdown of Dnmt3 caused worker larvae to develop into queens [55]. Therefore, tissue-specific fluctuations in the overall levels of gene body methylation, coupled with variable non-CG methylation, may regulate the expression of certain genes and be important for normal development. However, our findings strongly argue that this type of regulation is unlikely to be the core function of gene body methylation.

Conclusions

We demonstrate that honey bee DNA methylation is distinctly regulated in the soma and germline. In the germline, methylation is efficiently maintained, likely to preserve methylation patterns across generations. In

contrast, methylation can fluctuate in somatic cells, as long as overall methylation of exons and genes remains at an adequate level. Genome-wide methylation dynamics do not associate with gene regulation, and most developmentally regulated gene expression alterations occur in unmethylated genes. Accordingly, we conclude that gene body methylation is required for the homeostasis of gene expression and that global fluctuation of genic methylation is a non-functional feature of cells with limited division potential. Finally, our data suggest that heightened non-CG methylation is a conserved regulator of animal nervous systems, which may influence gene expression and development.

Methods

Biological samples

European honey bees were grown in Langstroth hives at the Harry H. Laidlaw Jr. Honey Bee Research Facility at University of California Davis. Honey bee samples were collected during summer in two rounds separated about 3 weeks apart (for two biological replicates). Worker embryos (eggs) were transferred from honeycomb cells into microcentrifuge tubes (50 per/tube). Drone and worker larvae were obtained at about 3rd and 4th instar developmental stages. Worker pupae were collected at about pink eyes developmental stage. Worker heads were collected from bees newly emerged from their own growing cells. Drone heads were collected from the same colony as drone larvae and worker embryos, larvae, pupae, and adult heads were collected from. Semen was collected from mature drones essentially as described

in [72]. All samples were flash frozen in liquid nitrogen immediately upon collection. Except for queen heads, for which we used a single head per experiment, the rest of the biological sampling consisted of a mix of individuals, three larvae/pupae/heads, 50 embryos, and 20ul of semen.

Bisulfite sequencing (BS-Seq)

About 500 ng of genomic DNA was isolated from various honey bee developmental stages, fragmented by sonication, end repaired, and ligated to custom synthesized methylated adapters (Eurofins MWG Operon) according to the manufacturer's (Illumina) instructions for gDNA library construction. Adaptor-ligated libraries were subjected to two successive treatments of sodium bisulfite conversion using the EpiTect Bisulfite kit (QIAGEN), as outlined in the manufacturer's instructions. The bisulfite-converted libraries were then amplified by PCR using ExTaq DNA polymerase (Takara Bio) for 12–14 cycles. The enriched libraries were purified using the solid-phase reversible immobilization method using AM-Pure beads (Beckman Coulter) prior to quantification with a Bioanalyzer (Agilent). Sequencing on the Illumina GAII and HiSeq 2000 platform was performed at the Vincent J. Coates Genomic Sequencing Laboratory at the University of California, Berkeley, USA (UC Berkeley) to generate single-end 76 and 100 base reads.

RNA-Seq

Total RNA samples were isolated using the RNeasy mini kit (QIAGEN #74106) including on-column DNase treatment. mRNA was purified from 10 to 50 µg of total RNA by two cycles of poly-A enrichment using the Oligotex kit (QIAGEN #72022) followed by an rRNA removal step using the RiboMinus Eukaryote Kit for RNA-Seq (Invitrogen #A1083702). Precipitated mRNA samples were eluted with 9 ml of RNase-free water and fragmented with 1 ml of fragmentation buffer (Ambion, #AM8740) at 70 °C. Reactions were stopped after 5 min by adding 1-ml stop buffer, and RNA was purified by ethanol precipitation. cDNA was synthesized from 100 to 300 ng of mRNA using SuperScript III reverse transcriptase (Invitrogen #18080-051). Double-stranded DNA was synthesized according to the instructions using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen). DNA was cleaned with a QIAquick PCR spin column (QIAGEN, #28106), sequencing adapters were ligated according to the Illumina protocol, and the library was amplified by 18 cycles of PCR using Phusion DNA polymerase (NEB, #F-530). Bands around 300 bp were gel purified and libraries were sequenced on the Illumina HiSeq 2000 platform at the Vincent J. Coates Genomic Sequencing Laboratory at UC Berkeley to generate single-end 100 base reads.

Data analysis

Differential methylation

Identifying differential methylation between various samples was performed by Fisher Exact Test using the number of Cs (methylated cytosines) and Ts (unmethylated cytosines) of any pair of samples. A site was determined to be differentially methylated if its *p* value was lower than 0.001 in both biological replicates.

Percent-methylation-change

This number was calculated by dividing the difference in methylation level between two samples by the level of methylation in the sample with the higher methylation level. For example, percent-methylation-change between embryo and W. larva was calculated as follows:

$$\frac{\text{Embryo mCG} - \text{W.larva mCG}}{\text{Embryo mCG}}$$

if Embryo mCG > W.larva mCG

$$- \frac{\text{W.larva mCG} - \text{Embryo mCG}}{\text{W.larva mCG}}$$

if Embryo mCG < W.larva mCG

Kernel density plots

Kernel density plots compare percent-methylation-change within either single-Cs, exons, or genes. For comparisons, we used genomic sites with at least 10 and 20 informative sequenced cytosines for single-C and complete exons or genes, respectively. Additionally, we used genomic sites with fractional methylation of at least 0.1 in at least one of the samples being compared.

Hierarchical clustering

The hierarchical clustering was performed and visualized in heatmaps using clustermap function from the Python seaborn library, with Euclidean distance matrix and average linkage method.

Gene expression

Raw Illumina RNA-Seq 100 base reads were first mapped to the most recent honey bee genome assembly (Amel_4.5; [73]) using Tophat [74], with the following changes to the default Tophat v2.0.1 parameters –I 100,000 (maximum intron length) and –no-novel-juncs (limiting the alignment to v3.2 honey gene annotation). Gene expression abundance and changes were calculated using the Cufflinks and Cuffdiff softwares [74] based on the honeybee 4.5 genome, as well as the v3.2 gene annotation file and default parameters with the addition of –min-reps-for-js-test. Differential expression was considered as statistically significant when the *q*-value (FDR

correction) was lower than 0.05 and the FPKM fold change between two samples was higher than two.

Gene responsiveness

We calculated gene responsiveness per annotated gene by counting the number of times it was differentially expressed within the seven developmental stages, i.e., 21 pairwise comparisons in total. Accordingly, a gene that was not differentially expressed in any of the pairwise comparisons, i.e., was similarly expressed in all developmental stages, was given a gene responsiveness score of zero. The maximum level of gene responsiveness in our data is 21, which was given to genes that were found to be differentially expressed in all 21 pairwise comparisons.

Significance of overlap between differentially methylated and alternatively expressed genes

RNA splicing: Alternative splicing events, and subsequently differential alternative splicing events, were identified by first running Cufflinks (version 2.2.1) on each sample and then Cuffmerge (version 2.2.1) to find novel transcript isoforms with $-\text{min-intron-length } 70$ $-\text{max-intron-length } 100,000$ settings [75]. To quantify transcript isoform level expression, the transcriptome annotation file from Cuffmerge was transformed into a transcriptome fasta file (using the gffread tool in Cufflinks version 2.2.1) and Kallisto (version 0.42.3) was run with fastq files as input (fragment length parameter $-l$ set to 200) [76]. Transcripts with an expression of less than 2 transcripts per million (TPM) were then removed. The SUPPA program [77] was used to analyze alternative splicing. Using the Cuffmerge GTF, SUPPA identifies all alternative events in the transcriptome, and then uses transcript abundances calculated by Kallisto to find the percentage spliced in (PSI) of each alternative event in each sample. Average PSI values were then calculated for each event between the two biological replicates and a t-test with multiple correction testing was performed to find differential alternative splicing. Differential events were defined by having a delta PSI of ≥ 0.1 and a q -value of < 0.05 .

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13072-019-0307-4>.

Additional file 1: Figure S1. Experimental reproducibility of gene body methylation dynamics during honey bee development. (A) Averaged CG methylation levels in the specified annotations, calculated separately for each of the biological replicates. Methylated exons (Me. exons) are exons with a minimum of 10% methylation in either of the samples. The y-axis is broken into two linear scales 0–10% and 20–50%. (B) Patterns of CG methylation in gene bodies within exonic sequences (i.e. excluding introns) separated to experiments. CG methylation profiles were generated similar to Fig. 1c. (C, D) Genomic snapshots of CG methylation of each of the replicates in a large-scale genomic region (D) and zoom in on single

exons (E). **Figure S2.** Correlation of exon methylation between biological replicates. Density scatter plots of methylation level in methylated exons (defined in S1A) correlated between each of the biological replicates. Note the high signal at maximum methylation in sperm and embryo samples. r is Pearson correlation coefficient value ($p < 10^{-4}$ for all correlations).

Figure S3. Gene body methylation is robustly maintained during honey bee development. (A) Bar graphs of the total number of partially and fully differentially methylated cytosines (left panel), exons (middle panel), and genes (right panel) between embryo and indicated developmental stages. (B) Sunflower plots (density scatter plot) of percent-methylation-change of individual CG cytosines between two biological replicates. The top right corner enclosed by dashed lines, holds the fDMC sites ($mCG \geq 0.9$). Note the single hexagon on the top right corner of the graphs containing multiple dark lines, which sum to 22–44% out of total cytosines in the graphs. (C) Kernel density plots of percent-methylation-change between averaged methylation in embryo versus indicated samples calculated for exons or genes separated to experiment 1 and experiment 2. Genes and exons were selected for the analysis if their average CG methylation was at least 10% in either of the biological stages in the relative experiment. (D) Scatter plots of percent-methylation-change in exons (upper plots) or genes (lower plots) between two comparison of embryo vs. other indicated samples. r is Pearson correlation coefficient value ($p < 10^{-4}$ for all correlations). **Figure S4.** Gene body methylation dynamics are not associated with developmentally regulated transcriptional profiles. (A) Sunflower plot of LogFC of RNA reads between embryo and indicated samples versus averaged genic CG methylation in embryos. obs. equals observations. (B) Scatter plots of LogFC of RNA versus percent-methylation-change between indicated samples. Red dots are of only genes that were found to be both differentially methylated (Fisher exact test $p < 0.05$ in both experiments) and alternatively expressed (T test $p < 0.05$ and $FC > 2$). r represents Pearson Correlation Coefficient values. **Figure S5.** CW methylation is enriched in adult honey bee heads. (A) CH methylation of all sequences (left) or specifically located in exons (middle) or introns (right) were averaged in 100 bp bins along honey bee genes essentially as described in Fig. 1a. (B) Quantification of CH error rate in each of the samples which is derived from methylated CG sites in reads overlapping CH sites in the reference genome. (C) Representative methylated CA motifs from queen bee heads. The relative frequencies of nucleotides around the methylated cytosine are represented through the Python seqlogo script. The height of each letter represents its information content.

Acknowledgements

We wish to thank Elizabeth Frost, Susan Cobey, and the Harry H. Laidlaw Jr. Honey Bee Research Facility at University of California Davis, for their help in collecting honey bee samples.

Authors' contributions

AZ and DZ designed the study. AZ and DZ supervised and provided funding for the study. AZ, KDH, JPBL, and KD generated and/or analyzed the genomic data. AZ and DZ interpreted results and wrote the manuscript. All authors read and approved the final manuscript.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No [679551]) to AZ and KDH. This work was also supported by Israel Science Foundation [Grant 757/12] to AZ and KD and by a Beckman Young Investigator award to DZ. JPBL wishes to acknowledge the Center for RNA Systems Biology at UC Berkeley (NIH P50 GM102706), and the Brenner Research Group at UC Berkeley for supporting him, as well as grants (NIH R01 GM071655, NIH U01 HG004271, and NIH R01 HG004872) for supporting the development of splicing analysis tools or funding servers.

Availability of data and materials

The raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo>) under accession numbers of GSE116629.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ School of Plant Sciences and Food Security, Tel-Aviv University, 69978 Tel-Aviv, Israel. ² Center for RNA Systems Biology, University of California, Berkeley, Berkeley, CA 94720, USA. ³ Present Address: ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Perth, WA 6009, Australia. ⁴ Department of Cell and Developmental Biology, John Innes Center, Norwich, UK.

Received: 26 April 2019 Accepted: 25 September 2019

Published online: 10 October 2019

References

- Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010;11:204–20. <https://doi.org/10.1038/nrg2719>.
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet.* 2012;13:484–92. <https://doi.org/10.1038/nrg3230>.
- Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol.* 2015;16:519–32. <https://doi.org/10.1038/nrm4043>.
- Quadrana L, Colot V. Plant transgenerational epigenetics. *Annu Rev Genet.* 2016;50:467–91. <https://doi.org/10.1146/annurev-genet-120215-035254>.
- Gilmore JH. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell.* 2008;29:1883–9.
- Heard E, Martienssen RA. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell.* 2014;157:95–109.
- He Y, Ecker JR. Non-CG methylation in the human genome. *Annu Rev Genomics Hum Genet.* 2015;16:55–77. <https://doi.org/10.1146/annurev-genom-090413-025437>.
- Luo C, Keown CL, Kurihara L, Zhou J, He Y, Li J, et al. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science.* 2017;357:600–4.
- Bewick AJ, Vogel KJ, Moore AJ, Schmitz RJ. Evolution of DNA methylation across insects. *Mol Biol Evol.* 2017;34:654–65.
- Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, et al. Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* 2016;17:194.
- Zemach A, Zilberman D. Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Curr Biol.* 2010;20:R780–5. <https://doi.org/10.1016/j.cub.2010.07.007>.
- Kim MY, Zilberman D. DNA methylation as a system of plant genomic immunity. *Trends Plant Sci.* 2014;19:320–6. <https://doi.org/10.1016/j.tplants.2014.01.014>.
- Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet.* 2008;9:465–76.
- Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science.* 2010;328:916–9. <https://doi.org/10.1126/science.1186366>.
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell.* 2006;126:1189–201.
- Suzuki MM, Yoshinari A, Obara M, Takuno S, Shigenobu S, Sasakura Y, et al. Identical sets of methylated and nonmethylated genes in *Ciona intestinalis* sperm and muscle cells. *Epigenetics Chromatin.* 2013;6:38.
- Keller TE, Han P, Yi SV. Evolutionary transition of promoter and gene body DNA methylation across invertebrate–vertebrate boundary. *Mol Biol Evol.* 2016;33:1019–28.
- Downen RH, Pelizzola M, Schmitz RJ, Lister R, Downen JM, Nery JR, et al. Widespread dynamic DNA methylation in response to biotic stress. *Proc Natl Acad Sci.* 2012;109:E2183–91.
- Wang X, Wheeler D, Avery A, Rago A, Choi J-H, Colbourne JK, et al. Function and evolution of DNA methylation in *Nasonia vitripennis*. *PLoS Genet.* 2013;9:e1003872. <https://doi.org/10.1371/journal.pgen.1003872>.
- Dixon GB, Bay LK, Matz MV. Evolutionary consequences of DNA methylation in a basal metazoan. *Mol Biol Evol.* 2016;33:2285–93.
- Aceituno FF, Moseyko N, Rhee SY, Gutiérrez RA. The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC Genomics.* 2008;9:1–14.
- Takuno S, Gaut BS. Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly. *Mol Biol Evol.* 2012;29:219–27.
- Sarda S, Zeng J, Hunt BG, Yi SV. The evolution of invertebrate gene body methylation. *Mol Biol Evol.* 2012;29:1907–16.
- Dimond JL, Roberts SB. Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change. *Mol Ecol.* 2016;25:1895–904.
- Gavery MR, Roberts SB. Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. *PeerJ.* 2013;1:e215.
- Gavery MR, Roberts SB. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). *BMC Genomics.* 2010;11:483.
- Elango N, Hunt BG, Goodisman MA, Yi SV. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc Natl Acad Sci USA.* 2009;106:11206–11.
- Li Z, Dai H, Martos SN, Xu B, Gao Y, Li T, et al. Distinct roles of DNMT1-dependent and DNMT1-independent methylation patterns in the genome of mouse embryonic stem cells. *Genome Biol.* 2015;16:115.
- de Mendoza A, Bonnet A, Vargas-Landin DB, Ji N, Li H, Yang F, et al. Recurrent acquisition of cytosine methyltransferases into eukaryotic retrotransposons. *Nat Commun.* 2018;9:1341.
- Takuno S, Ran J-H, Gaut BS. Evolutionary patterns of genic DNA methylation vary across land plants. *Nat Plants.* 2016;2:15222.
- Bewick AJ, Ji L, Niederhuth CE, Willing E-M, Hofmeister BT, Shi X, et al. On the origin and evolutionary consequences of gene body DNA methylation. *Proc Natl Acad Sci USA.* 2016;113:9111–6.
- Bewick AJ, Sanchez Z, McKinney EC, Moore AJ, Moore PJ, Schmitz RJ. Dnmt1 is essential for egg production and embryo viability in the large milkweed bug, *Oncopeltus fasciatus*. *Epigenetics Chromatin.* 2019;12:6. <https://doi.org/10.1186/s13072-018-0246-5>.
- Wendte JM, Zhang Y, Ji L, Shi X, Hazarika RR, Shahryary Y, et al. Epimutations are associated with CHROMOMETHYLASE 3-induced de novo DNA methylation. *Elife.* 2019;8:e47891.
- Zilberman D. An evolutionary case for functional gene body methylation in plants and animals. *Genome Biol.* 2017;18:87. <https://doi.org/10.1186/s13059-017-1230-2>.
- Drewell RA, Bush EC, Remnant EJ, Wong GT, Beeler SM, Stringham JL, et al. The dynamic DNA methylation cycle from egg to sperm in the honey bee *Apis mellifera*. *Development.* 2014;141:2702–11. <https://doi.org/10.1242/dev.110163>.
- Yi SV, Goodisman MAD. Computational approaches for understanding the evolution of DNA methylation in animals. *Epigenetics.* 2009;4:551–6.
- Herb BR, Wolschin F, Hansen KD, Aryee MJ, Langmead B, Irizarry R, et al. Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nat Neurosci.* 2012;15:1371–3.
- Rodrigues JA, Ruan R, Nishimura T, Sharma MK, Sharma R, Ronald PC, et al. Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. *Proc Natl Acad Sci USA.* 2013;110:7934–9.
- McGaughey DM, Abaan HO, Miller RM, Kropp PA, Brody LC. Genomics of CpG methylation in developing and developed zebrafish. *G3.* 2014;4:861–9.
- Schroeder DJ, Jayashankar K, Douglas KC, Thirkill TL, York D, Dickinson PJ, et al. Early developmental and evolutionary origins of gene body DNA methylation patterns in mammalian placentas. *PLoS Genet.* 2015;11:e1005442.

41. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell*. 2014;26:577–90. <https://doi.org/10.1016/j.ccr.2014.07.028>.
42. Du M, Luo M, Zhang R, Finnegan EJ, Koltunov AMG. Imprinting in rice: the role of DNA and histone methylation in modulating parent-of-origin specific expression and determining transcript start sites. *Plant J*. 2014;79:232–42.
43. Maleszka R. Epigenetic code and insect behavioural plasticity. *Curr Opin Insect Sci*. 2016;15:45–52. <https://doi.org/10.1016/j.cois.2016.03.003>.
44. Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. *Trends Genet*. 2015;31:274–80.
45. Li-Byarlay H, Li Y, Stroud H, Feng S, Newman TC, Kaneda M, et al. RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. *Proc Natl Acad Sci*. 2013;110:12750–5.
46. Bonasio R, Li Q, Lian J, Mutti NS, Jin L, Zhao H, et al. Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Curr Biol*. 2012;22:1755–64.
47. Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, et al. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*. 2013;153:193–205. <https://doi.org/10.1016/j.cell.2013.02.033>.
48. Teixeira FK, Colot V. Gene body DNA methylation in plants: a means to an end or an end to a means? *EMBO J*. 2009;28:997–8. <https://doi.org/10.1038/emboj.2009.87>.
49. Coleman-Derr D, Zilberman D. DNA methylation, H2A.Z, and the regulation of constitutive expression. *Cold Spring Harb Symp Quant Biol*. 2012;LXXVIII:147–54.
50. Huh I, Zeng J, Park T, Yi SV. DNA methylation and transcriptional noise. *Epigenetics Chromatin*. 2013;6:9.
51. Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature*. 2017;543:72–7. <https://doi.org/10.1038/nature21373>.
52. Choi J, Lyons DB, Kim Y, Moore JD, Zilberman D. DNA methylation and histone H1 cooperatively repress transposable elements and aberrant intragenic transcripts. *bioRxiv*. 2019;527523. <https://doi.org/10.1101/527523v1>.
53. Stroud H, Greenberg M, Feng S. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell*. 2013;152:352–64.
54. Liao J, Karnik R, Gu H, Ziller MJ, Tsankov AM, Akopian V, et al. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol*. 2015;47:469–78.
55. Kucharski R, Maleszka J, Foret S, Maleszka R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science*. 2008;319:1827–30.
56. Zwier MV, Verhulst EC, Zwahlen RD, Beukeboom LW, van de Zande L. DNA methylation plays a crucial role during early *Nasonia* development. *Insect Mol Biol*. 2012;21:129–38.
57. Glastad KM, Gokhale K, Lieb J, Goodisman MAD. The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. *Sci Rep*. 2016;6:1–14. <https://doi.org/10.1038/srep37110>.
58. Libbrecht R, Oxley PR, Keller L, Kronauer DJC. Robust DNA methylation in the clonal raider ant brain. *Curr Biol*. 2016;26:391–5. <https://doi.org/10.1016/j.cub.2015.12.040>.
59. Wang X, Werren JH, Clark AG. Allele-specific transcriptome and methylome analysis reveals stable inheritance and cis-regulation of DNA methylation in *Nasonia*. *PLoS Biol*. 2016;14:e1002500. <https://doi.org/10.1371/journal.pbio.1002500>.
60. Wang Y, Jorda M, Jones PL, Maleszka R, Ling X, Robertson HM, et al. Functional CpG methylation system in a social insect. *Science*. 2006;314:645–7.
61. Uysal F, Ozturk S, Akkoyunlu G. DNMT1, DNMT3A and DNMT3B proteins are differently expressed in mouse oocytes and early embryos. *J Mol Histol*. 2017;48:417–26. <https://doi.org/10.1007/s10735-017-9739-y>.
62. Kawakatsu T, Stuart T, Valdes M, Breakfield N, Schmitz RJ, Nery JR, et al. Unique cell-type-specific patterns of DNA methylation in the root meristem. *Nat Plants*. 2016;2:16058.
63. Bouyer D, Kramdi A, Kassam M, Heese M, Schnittger A, Roudier F, et al. DNA methylation dynamics during early plant life. *Genome Biol*. 2017;18:179.
64. Kawakatsu T, Nery JR, Castanon R, Ecker JR. Dynamic DNA methylation reconfiguration during seed development and germination. *Genome Biol*. 2017;18:171.
65. Hsieh P, He S, Buttress T, Gao H, Couchman M, Fischer RL, et al. Arabidopsis male sexual lineage exhibits more robust maintenance of CG methylation than somatic tissues. *Proc Natl Acad Sci*. 2016;113:15132–7. <https://doi.org/10.1073/pnas.1619074114>.
66. Park K, Kim MY, Vickers M, Park J-S, Hyun Y, Okamoto T, et al. DNA demethylation is initiated in the central cells of Arabidopsis and rice. *Proc Natl Acad Sci USA*. 2016;113:15138–43.
67. Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, et al. Transgenerational epigenetic instability is a source of novel methylation variants. *Science*. 2011;334:369–73.
68. Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, et al. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature*. 2011;480:245–9.
69. Bostick M, Jong KK, Estève PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science*. 2007;317:1760–4.
70. Piccolo FM, Fisher AG. Getting rid of DNA methylation. *Trends Cell Biol*. 2014;24:136–43.
71. Takuno S, Gaut BS. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc Natl Acad Sci USA*. 2013;110:1797–802.
72. Robinson KO, Ferguson HJ, Cobey S, Vaessin H, Smith BH. Sperm-mediated transformation of the honey bee, *Apis mellifera*. *Insect Mol Biol*. 2000;9:625–34.
73. Elsik CG, Worley KC, Bennett AK, Beye M, Camara F, Childers CP, et al. Finding the missing honey bee genes: lessons learned from a genome upgrade. *BMC Genomics*. 2014;15:86.
74. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7:562–78.
75. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28:511–5.
76. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34:525–7.
77. Alamancos GP, Pagès A, Trincado JL, Bellora N, Eyra E. Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA*. 2015;21:1521–31.
78. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*. 2011;6:e21800. <https://doi.org/10.1371/journal.pone.0021800>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.