Supplementary Figures for:

Comparative methylomics reveals gene-body H3K36me3 in *Drosophila* predicts DNA methylation and CpG landscapes in other invertebrates

Lisa Nanty[1,4], Guillermo Carbajosa[1,4], Graham A. Heap[1,4], Francis Ratnieks[2], David A. van Heel[1], Thomas A. Down[3], & Vardhman K. Rakyan[1]

[1] The Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, 4 Newark Street, London E1 2AT, UK.

[2] Laboratory of Apiculture and Social Insects, Department of Biological and Environmental Science, University of Sussex, Falmer, Brighton BN1 9QG, UK

[3] The Gurdon Institute and Department of Genetics, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK.

[4] These authors contributed equally to this work.

Correspondence should be addressed to VKR (v.rakyan@qmul.ac.uk)

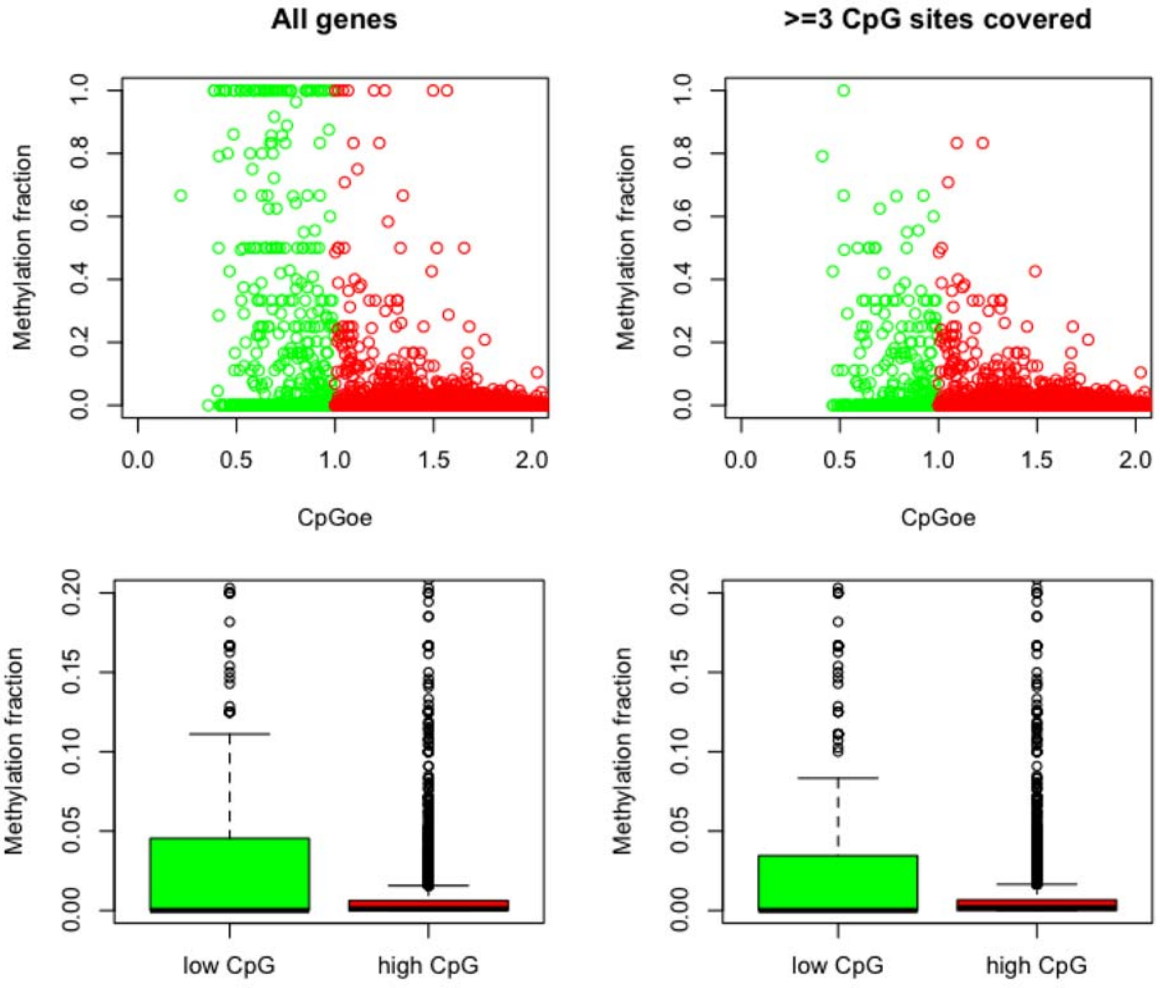
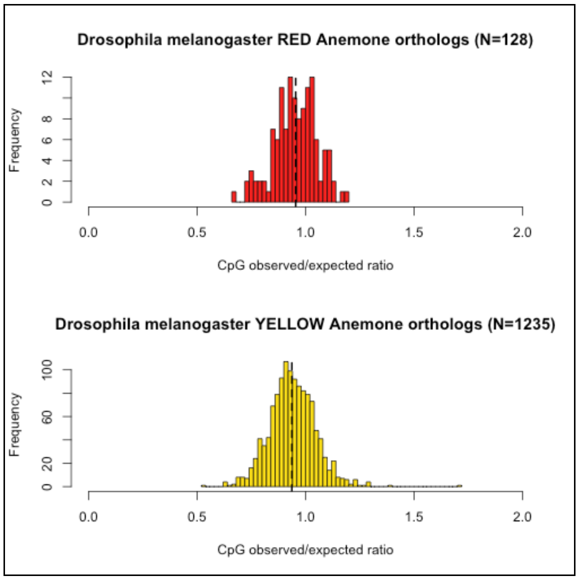
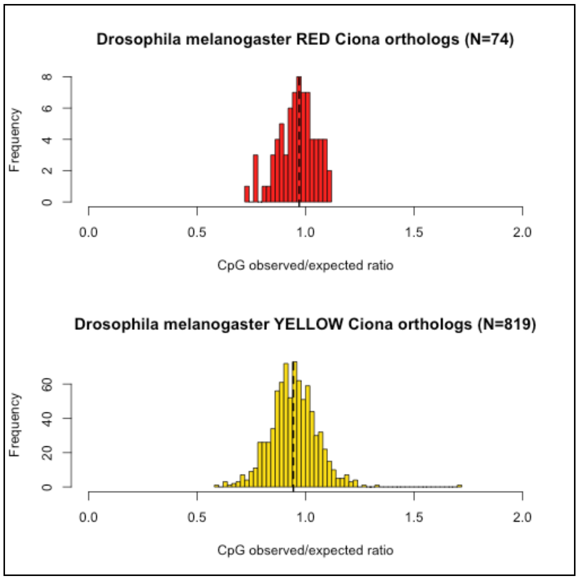


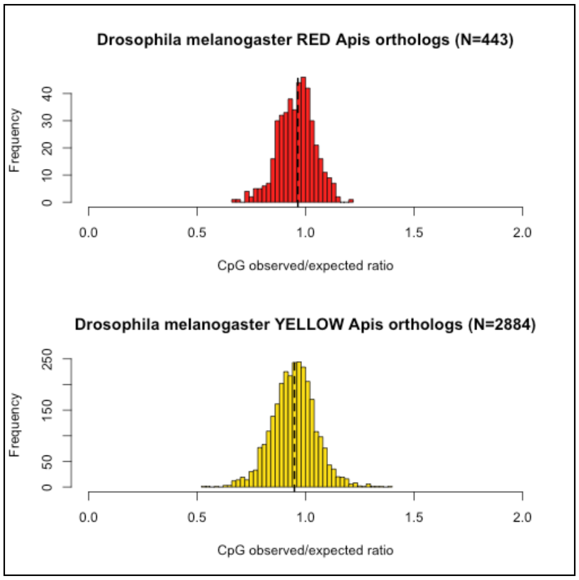
Figure S1. We generated RRB-seq data for A. mellifera mature spermatozoa (from a pool of 6 drones). Average gene-body CpGo/e and DNA methylation levels were calculated for each gene (Materials and Methods). The “All genes” plots on the left contain 933 low-CpG genes (green) and 4,537 high-CpG genes (red). The “≥3 CpG sites covered” plots on the right contain 487 low-CpG genes (green) and 4,183 high-CpG genes (red). The plots in the second row are the same data as the top row but shown as box-plots. When using all genes, or only those for which we have data for at least 3 CpG sites, low CpG genes are significantly more methylated than high-CpG genes in sperm (for all cases Mann Whitney test p-value < 0.001).


Drosophila melanogaster RED Anemone orthologs (N=128)

Drosophila melanogaster YELLOW Anemone orthologs (N=1235)


Drosophila melanogaster RED Ciona orthologs (N=74)

Drosophila melanogaster YELLOW Ciona orthologs (N=819)


Drosophila melanogaster RED Apis orthologs (N=443)

Drosophila melanogaster YELLOW Apis orthologs (N=2884)

Figure S2 (previous page). CpG content profiles of Drosophila YELLOW and RED chromatin-associated genes whose orthologs were used in the analysis presented in Figure 5 of the main text. To test that the difference between gene-body CpG density profiles for YELLOW and RED gene orthologs in *N. vectensis*, *C. intestinalis*, and *A. mellifera* is significantly greater than the difference between gene-body CpG density profiles for the orthologous YELLOW and RED genes in *Drosophila*, we performed the following test: Firstly, we calculated the difference in CpG density between each RED gene ortholog in *N. vectensis* and a randomly selected YELLOW gene ortholog in *N. vectensis*. Then we calculated the absolute value of the difference in CpG density between the Drosophila orthologs of those two *N. vectensis* genes. If the difference was bigger for the *N. vectensis* genes we gave it a score of 1, otherwise the score given was 0. Performing a binomial test with all the scores, using the ‘binom.test’ function in R, and using $P = 0.5$ as the null hypothesis (i.e. the difference is the same in *N. vectensis* and *Drosophila*) we obtained $P < 0.001$, that is, *N. vectensis* genes are significantly more different in CpG density compared with the orthologs in Drosophila. We repeated the test separately for *C. intestinalis*, and *A. mellifera,* obtaining a $P < 0.001$ in both cases.

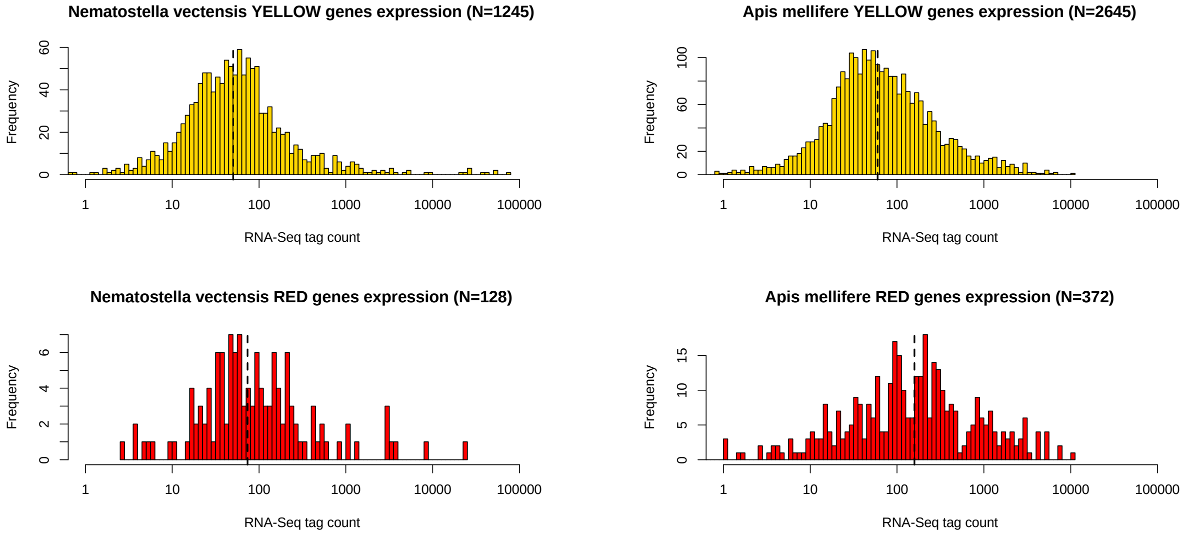


Figure S3. Gene expression differences between YELLOW and RED gene orthologs used in the analysis presented in Figure 5 of the main text. We obtained mRNA-seq RPKM values from Zemach et al., 2010 and compared gene expression data for YELLOW vs. RED gene orthologs. RED genes seem to be expressed at slightly higher levels compared with YELLOW genes in both *N. vectensis* and *A. Mellifera*. Wilcoxon rank sum test with continuity correction for both yields $P < 1 \times 10^{-4}$ in both cases (insufficient expression data were available for YELLOW and RED gene orthologs in *C. intestinalis* and *B. mori*).