Replicated DNA methylation differences between stick insect ecotypes

Clarissa de Carvalho[1], Jon Slate[2], Romain Villoutreix[3], Víctor Soria-Carrasco[4], Rüdiger Riesch[5], Jeff Feder[6], Zachariah Gompert[7], and Patrik Nosil[8]

[1]The University of Sheffield
[2]University of Sheffield
[3]Centre d'Ecologie Fonctionnelle et Evolutive
[4]John Innes Centre
[5]Royal Holloway University of London
[6]U. Notre Dame
[7]Utah State University
[8]CEFE

June 16, 2023

Abstract

Epigenetic mechanisms, such as DNA methylation, can influence gene regulation and affect phenotypic variation, raising the possibility that they contribute to ecological adaptation. To being to address this issue requires high-resolution sequencing studies of natural populations to pinpoint epigenetic regions of potential ecological and evolutionary significance. However, such studies are still relatively uncommon, especially in insects, and are mainly restricted to a few model organisms. Here, we characterize patterns of DNA methylation for natural populations of Timema cristinae adapted to two host plant species (i.e., ecotypes). By integrating results from sequencing of whole transcriptomes, genomes, and methylomes, we investigate whether environmental, host, and genetic differences of these stick insects are associated with methylation levels of cytosine nucleotides in CpG context. We report an overall genome-wide methylation level for T. cristinae of ~14%, being enriched in gene bodies and impoverished in repetitive elements. Genome-wide DNA methylation variation was strongly positively correlated with genetic distance (relatedness), but also exhibited significant host-plant effects. Using methylome-environment association analysis, we pinpointed specific genomic regions that are differentially methylated between ecotypes, with these regions being enriched for genes with functions in membrane processes. The observed association between methylation variation with genetic relatedness and the ecologically-important variable of host plant suggest a potential role for epigenetic modification in T. cristinae adaptation. To substantiate such adaptive significance, future studies could test if methylation has a heritable component and the extent to which it responds to experimental manipulation in field and laboratory studies.

Hosted file

`methylation_MolEcol_final.docx` available at https://authorea.com/users/629638/articles/649791-replicated-dna-methylation-differences-between-stick-insect-ecotypes

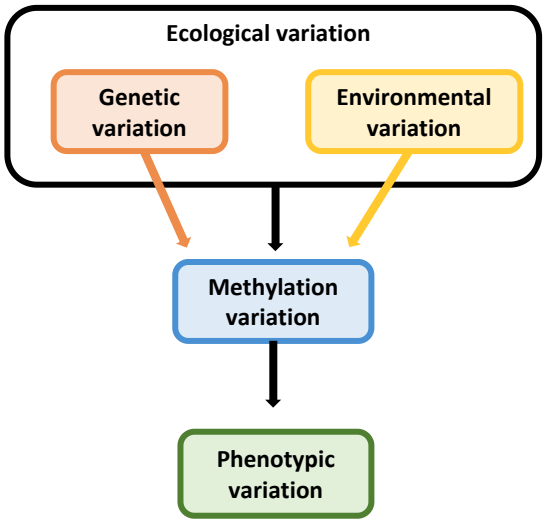

Fig. 1. Factors affecting methylation variation and its consequenses for phenotypic variation. The genetic background can influence methylation patterns. Additionally, environmental factors can affect methylation variation independently of or via an interaction with the genetic background (G x E). Knowing how methylation varies with ecological variation is another factor required to understand if and how methylation might contribute to variation in traits affecting fitness.

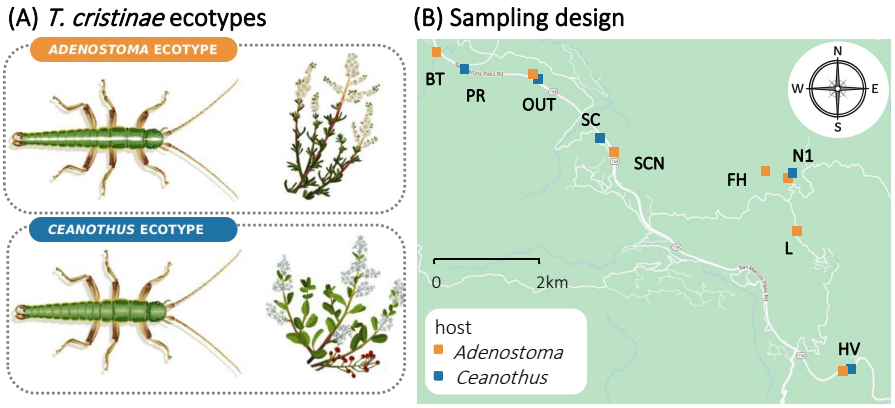

Fig. 2. The T. cristinae study system. (**A**) *T. cristinae* and their host plants: *Adenostoma fasciculatum* and *Ceanothus spinosus*. The ecotypes not only differ by the frequency of the dorsal white stripe, but also by differences in host-preference, body size, mate choice and cuticular hydrocarbons. (**B**) Map of the populations used in this study, selected based on host-plant species and their abundance, as well as elevation, climatic and geographic distance between populations (*Supplementary Materials*; Table S2). The general study area is situated in the Santa Ynez Moutains, in California, USA.

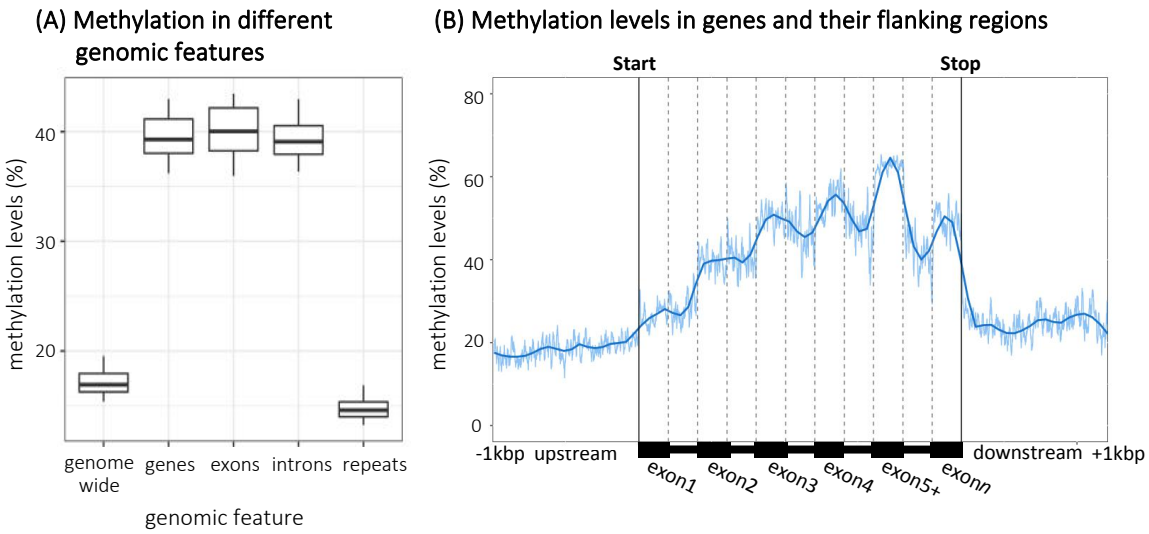

Fig. 3. Comparison of methylation levels for different components of the *T. cristinae* genome. (**A**) Methylation levels are enriched in both exons and introns (i.e., the gene body) compared to genome-wide levels, and genetic repeats tend to be impoverished in methylation. (**B**) DNA methylation levels in genes and their flanking regions. The graph represents 1kbp in the 5' downstream flanking region, multiple exons and introns in the depicted genetic region, and 1kbp in the 3' downstream region. The graph shows mean methylation levels estimated at CpG sites found in at least 12 samples (n=14,656 genes). The x-axis represents nucleotide position from the beginning or from the end of the genomic feature. To compare exons and introns of different genes, we used the mean methylation in the first 100bp at 5' and the last 100bp 3' of each exon and each intron, followingGlastad et al. (2016) and Hunt et al. (2013).

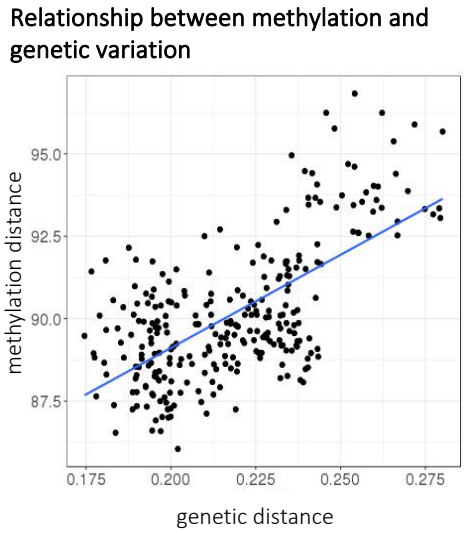

Fig. 4. Genome-wide methylation differences are correlated with genetic distance. Pairwise methylation distances were estimated using Euclidean distances between individuals, and genetic distance using the Kimura two-parameter model using GBS alignments. Regression was evaluated for significance using a Mantel test, with more complex multivariate analyses using Bayesian regression reported in the main text.

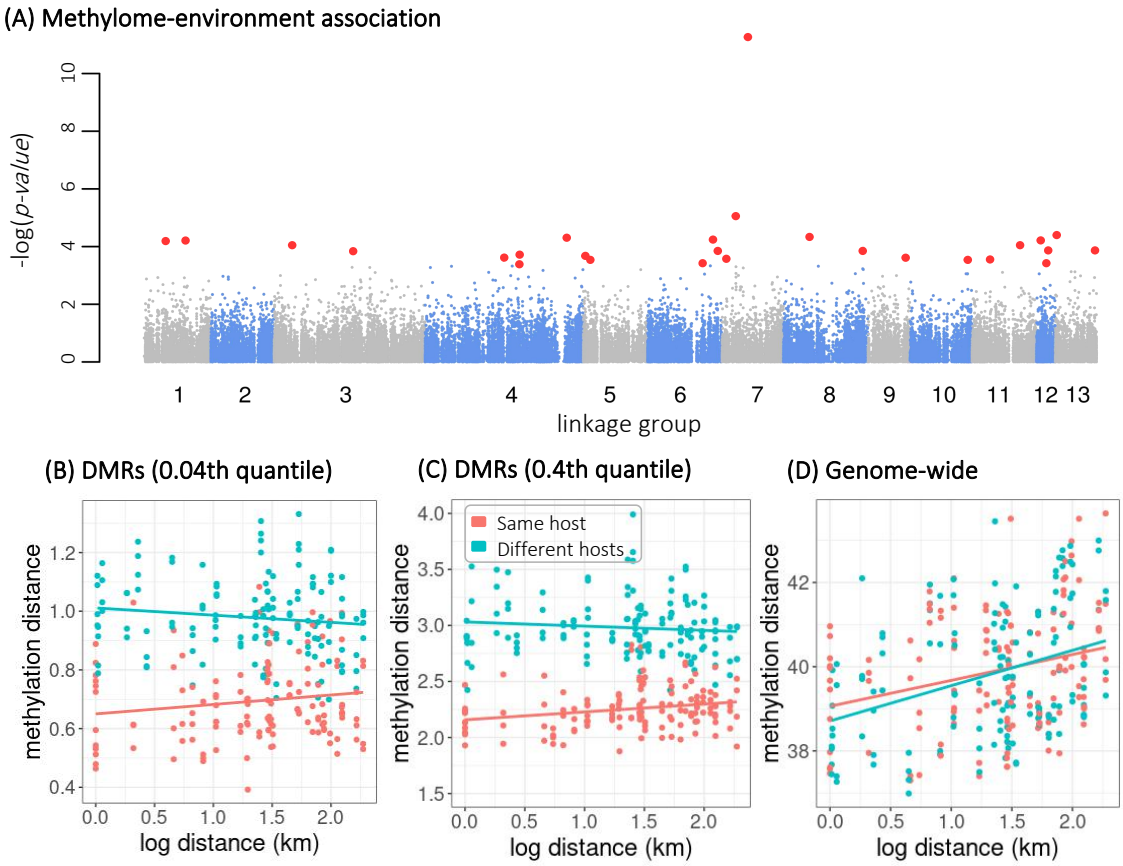

Fig. 5. Evidence for association between methylation patterns of specific genetic regions and host plant use (*i.e.*, differentially methylated regions, DMRs, between ecotypes). (**A**) Manhattan plot showing association between methylation variation and ecotype across all 24 samples, for 1 kilobase-pair (kbp) tiling windows using MACAU. (Lea *et al.*, 2015). Red points represent DMRs delimited by the 0.04th quantile of the empirical p-value distribution ($P < 0.0004$, see Table S15 for details on other quantiles). Pairwise methylation distances vary mostly according to ecotype in DMRs in different quantiles of the empirical *p-value* distribution, here, represented by the (**B**) 0.04th quantile; and (**C**) 0.4th quantile (also see Fig. S4, Table S16). (**D**) Genome-wide trends vary according to the geographical distances in an isolation-by-distance pattern, whereas DMRs show more of a host association. Methylation distances were obtained with Euclidean distances using the 1kbp windows in MACAU (also see Table S17).