ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Environmental stress during larval development induces DNA methylation shifts in the migratory painted lady butterfly (*Vanessa cardui*)

Jesper Boman[1] | Yishu Zhu[2] | Lars Höök[1] | Roger Vila[3] | Gerard Talavera[4] | Niclas Backström[1]

[1]Evolutionary Biology Program, Department of Ecology and Genetics (IEG), Uppsala University, Uppsala, Sweden

[2]Animal Ecology Program, Department of Ecology and Genetics (IEG), Uppsala University, Uppsala, Sweden

[3]Institut de Biologia Evolutiva (CSIC-Univ. Pompeu Fabra), Barcelona, Spain

[4]Institut Botànic de Barcelona (IBB), CSIC-Ajuntament de Barcelona, Barcelona, Spain

Correspondence
Jesper Boman, Evolutionary Biology Program, Department of Ecology and Genetics (IEG), Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden.
Email: jesper.boman@ebc.uu.se

Abstract

Seasonal environmental fluctuations provide formidable challenges for living organisms, especially small ectotherms such as butterflies. A common strategy to cope with harsh environments is to enter diapause, but some species avoid unsuitable conditions by migrating. Despite a growing understanding of migration in the life cycles of some butterfly species, it remains unknown how individuals register and store environmental cues to determine whether and where to migrate. Here, we explored how competition and host plant availability during larval development affect patterns of DNA methylation in the migratory painted lady (*Vanessa cardui*) butterfly. We identify a set of potentially functional methylome shifts associated with differences in the environment, indicating that DNA methylation is involved in the response to different conditions during larval development. By analysing the transcriptome for the same samples used for methylation profiling, we also uncovered a non-monotonic relationship between gene body methylation and gene expression. Our results provide a starting point for understanding the interplay between DNA methylation and gene expression in butterflies in general and how differences in environmental conditions during development can trigger unique epigenetic marks that might be important for behavioural decisions in the adult stage.

KEYWORDS
DNA methylation, environmental stress, migration, painted lady butterfly, *Vanessa cardui*

1 | INTRODUCTION

Organisms are forced to adapt to seasonal variation in both abiotic and biotic factors. Butterflies and moths (lepidopterans) have adopted strategies to cope with unfavourable conditions by either entering a state of dormancy and metabolic inactivity (diapause) or by migrating to regions with more favourable conditions, or a combination thereof (Urquhart, 1976; Williams, 1930). While the mechanistic underpinnings of diapause in lepidopterans have been investigated in some detail (e.g., Denlinger, 2002; Hasegawa, 1957; Kozak et al., 2019; Leal et al., 2018; Lehmann et al., 2016; Süess et al., 2022; Watanabe, 1924; Yamashita, 1996), the molecular

machinery underlying migratory behaviour is somewhat less known, with research focused on a few species of moths and the monarch (*Danaus plexippus*) butterfly (Chowdhury et al., 2022; Jiang et al., 2011; Reppert & de Roode, 2018).

The painted lady (*Vanessa cardui*) is an emerging model system for investigating migratory behaviour in insects (García-Berro et al., 2023; Menchetti et al., 2019; Näsvall et al., 2023; Shipilina et al., 2022; Stefanescu et al., 2013, 2017, 2021; Suchan et al., 2019; Talavera et al., 2018; Wiklund & Friberg, 2022). Besides being the most widespread of all butterflies—the species distribution is almost worldwide—the painted lady is also the butterfly migrating the longest distances (Talavera et al., 2018; Talavera & Vila, 2017). Painted ladies undertake an annual migratory cycle that can span ~6–10 generations. The most well-studied circuit runs from tropical Africa to northern Europe and back (Menchetti et al., 2019; Talavera et al., 2018; Talavera & Vila, 2017). Since the migratory movement of the painted lady is multi-generational, the direction and duration of migratory flight cannot be 'learned' by conspecifics. Instead, individuals must use environmental cues to determine whether and where to migrate. It is not known at what life stage the decisions are made. An important environmental cue is likely host plant abundance (Hu et al., 2021; López-Mañas et al., 2022; Stefanescu et al., 2021), which hypothetically could be sensed already in the early larval stages. In the armyworm (*Mythimna separata*) moth, for example, periodic starvation during development increases migration tendencies in the adults (Cao et al., 1996). Both the abundance of food and the density of competing conspecifics could inform the developing larvae about the expected future availability of host plants. Higher larval rearing densities seem to enhance flight phenotypes in most, but not all, Lepidoptera and other insects showing migration tendencies (Du et al., 2022; Jiang et al., 2011; Lü et al., 2022; Parker & Gatehouse, 1985).

Information about environmental signals such as density of conspecifics or food scarcity needs to be stored, processed and transmitted during metamorphosis to influence migratory behaviour of the adult butterfly. One category of molecular mechanisms that fulfils the requirements of information storage and transmission is epigenetic modifications, especially 5′-CpG-3′ DNA methylation (hereafter 'methylation'), which can be transmitted to daughter strands during DNA replication with high fidelity (Catania et al., 2020). In plants and vertebrates, methylation of promoter sequences has been linked to transcriptional repression (Bird, 2002; Sarkies, 2022; Suzuki & Bird, 2008; Yi, 2017), while the functional role of gene body methylation is still under debate (Jjingo et al., 2012; Neri et al., 2017; Zilberman, 2017). In arthropods, patterns of methylation are less characterized and potential functional relationships with transcription have been investigated to some extent but results remain inconclusive (Bonasio et al., 2012; Hunt et al., 2013; Jones et al., 2018; Lewis et al., 2020; Liu et al., 2019; Oldroyd & Yagound, 2021; Xiang et al., 2010, 2013; Xu et al., 2018, 2021). A study in the crustacean *Daphnia pulex* revealed that DNA methylation patterns induced by a stressful environment can be transgenerationally inherited (Harney et al., 2022). It remains to be shown if this is a general pattern in

arthropods. Recent initiatives suggest that the genome-wide levels of methylation can vary from moderate (~20%) in some arthropod species to almost undetectable (~0%) in others (Bewick et al., 2017; Lewis et al., 2020; Oldroyd & Yagound, 2021). An analysis of silk gland tissue of silk moth (*Bombyx mori*) larvae revealed a low genome-wide methylation level (0.11%) and a positive relationship between gene expression and gene body methylation (Xiang et al., 2010). An investigation of potential relationships between methylation and expression in adult cotton bollworm moths (*Helicoverpa armigera*) with different flight performance phenotypes pinpointed some differentially methylated CpG sites (Jones et al., 2018). However, the experiment covered only 16 genes and the data were not sufficient to establish if differentially expressed genes were enriched with differentially methylated CpGs. A general aim of our study was therefore to cover this gap by studying genome-wide associations between gene expression and DNA methylation in general. We combine that with specific analyses of methylation and expression differences between individuals reared under different environmental conditions.

Here we investigate the DNA methylome of the painted lady using whole-genome bisulphite sequencing of experimental cohorts exposed to different environmental cues that potentially can affect the propensity of individuals to migrate. Our study consists of three components. First, we map general patterns of methylation among genomic features. Second, we combine the methylation data with gene expression profiling of the same samples to uncover the relationship between methylation and gene expression at an unprecedented scale in butterflies. Third, we test if DNA methylation can be a mechanism to transfer environmental signals experienced during larval development to the imaginal stage in female butterflies by splitting larvae into three treatment groups with different larval densities and food availabilities. Since we hypothesise that neurogenetics has a major role in 'decision-making', we analyse heads (including antennae), where the nervous system of butterflies is concentrated. Because of previously established associations between food scarcity and density during larval development on adult flight phenotypes in other migratory Lepidoptera, we expect a molecular response to the treatments. Based on this expectation we ask three questions: (i) does environmental stress during larval development produce head methylome profile shifts in adults, (ii) if so, what genes and biological processes are involved and, (iii) what is the relationship between methylome- and transcriptome profile shifts between treatments?

2 | MATERIALS AND METHODS

2.1 | Experimental setup

Mated female painted lady (*Vanessa cardui*) butterflies were collected in Catalonia, Spain, in spring 2020 and kept in individual cages at 25°C, 18:6 h light: dark regime, with access to a host plant (*Malva sylvestris*) and 10% sugar water. F_1 offspring were raised on ad libitum (AL) *M. sylvestris*, and individually marked as adults and released

in a monitored common cage with other F_1 adults. F_1 females were separated after first mating and put in individual flasks with *M. sylvestris* for egg-laying. We used a split-brood design were F_2 eggs from each mated female were randomly split into three different treatment groups representing larval environmental conditions: LDAL (low density, ad libitum food), HDAL (high density, ad libitum food) and HDLI (high density, limited food). LD represented larvae kept individually, and HD represent larvae raised in groups of 10. In the LI treatment, food was only replaced every other day, leading to a 1:1 day food: starvation regime. Female F_2 adults were flash frozen in liquid nitrogen immediately after eclosure as adults and stored at −80°C until nucleic acid-extraction. The heads (including antennae) of nine individuals (total $n=9$, Table S1) were homogenized individually on ice and separated in aliquots for RNA and DNA extraction. Sample sizes were $n=3$ per treatment group, except for RNA-seq of HDLI ($n=2$) due to inadequate RNA quality in one sample.

2.2 | DNA extraction and bisulphite sequencing

DNA was extracted using a standard phenol-chloroform extraction protocol. Library preparation and whole-genome bisulphite sequencing were performed by the SciLife SNP&SEQ Technology Platform in Uppsala, Sweden. Sequencing libraries were prepared from 4 to 200 ng of DNA according to the SPLAT method (Raine et al., 2017). Prior to bisulphite conversion, 0.5–1% of lambda phage was spiked-in. The bisulphite conversion error rate was low and varied from 0.2 to 0.3% between samples. Paired-end sequencing (150 bp read length) was performed using the NovaSeq 6000 technology and v1.5 sequencing chemistry in a single S4 flow cell.

2.3 | RNA extraction and sequencing

The RNeasy Mini Kit (Qiagen) was used to extract RNA following the general instructions in the kit manual. Library preparations (Illumina TruSeq Stranded mRNA polyA selection kit) and RNA sequencing were performed by the National Genomics Infrastructure (NGI) in Stockholm. Libraries were sequenced on two lanes on one S4 flow cell on the NovaSeq S6000 with 150 bp paired-end reads (in multiplex with additional samples used for another study, data not shown).

2.4 | Processing of bisulphite sequence reads and methylation calling

All processing steps from adapter filtering to read mapping, and methylation calling were performed using the reproducible Nextflow v21.04.1 nf-core Methylseq v1.5 pipeline (Di Tommaso et al., 2017). Reads were trimmed and adapters were removed using Trim Galore! v0.6.4_dev, employing Cutadapt v2.9 (Martin, 2011). Reads were trimmed based on a maximum allowed error rate of 10% and bases with a Phred quality score of at least 20 were removed.

Further clipping of reads was performed using the Epignom profile (8 bp from both 5′ and 3′ ends of both reads in a pair). Trimmed reads were aligned using Bismark v0.22.3 (Krueger & Andrews, 2011) with Bowtie 2 as aligner using directional alignment to a *Vanessa cardui* reference genome (Lohse et al., 2021). The range of the percentage of mapped reads was 35.5–42.3% (Table S1). Duplicated reads were removed and the median read depth after deduplication ranged from 9 to 28X (Table S1). Bismark was also used for methylation calling with (−-no_overlap) specified. Quality control of output from the Methylseq pipeline was performed using MultiQC v1.8 (Ewels et al., 2016).

2.5 | Gene expression quantification and differential expression analysis

To assess potential relationships between gene expression and methylation, we sequenced the transcriptome of 8 of the same samples for which methylation levels were scored and quantified the gene expression levels in each individual. All processing steps from adapter filtering to read mapping and transcript quantification were performed using the reproducible Nextflow v21.10.6 nf-core rnaseq v3.8.1 pipeline (Di Tommaso et al., 2017). Reads were trimmed using Trim Galore! v0.6.7, employing Cutadapt v3.4 (Martin, 2011). Reads were mapped to the reference using STAR v2.7.10a (Dobin et al., 2013) and a previously available gene annotation consisting of 13,161 genes (Shipilina et al., 2022). Read quantification was performed using salmon v1.5.2 (Patro et al., 2017), and we used transcripts per million reads (TPM) values as a measure of gene expression level. Quality control of the RNAseq pipeline was performed using MultiQC v1.11 (Ewels et al., 2016). Differential expression (DE) analyses were performed to investigate if the larval treatments resulted in transcriptional profile shifts and whether they were associated with differentially methylated regions. To identify DE genes, we used deseq2 v1.28.0 in R v4.2.1 taking into account sibling relationships using the design: ~Family+Treatment (Love et al., 2014). A false discovery rate of 0.1 was used to correct for multiple testing.

2.6 | Methylation level

When using bulk samples of DNA, we need to measure methylation as a proportion, since chromatids differ in methylation status dependent on cell type and methyltransferase dynamics. To be precise, it is the combined level of 5′-CpG-3′ methylation and hydroxymethylation that is measured since bisulphite sequencing cannot distinguish these marks (Yong et al., 2016). Here, we measured methylation level by summarizing the number of methylated x_{mCpG} reads mapping to both strands of a reference CpG dinucleotide and dividing by the total number of reads $x_{uCpG}+x_{mCpG}$. To measure the methylation level covering n CpG dinucleotides, we used the average proportion across the individual dinucleotides (i),

$$\text{Methylation level} = \frac{\sum_{i=1}^{n} \frac{x_{mCpG}}{x_{mCpG} + x_{uCpG}}}{n}.$$

To decrease the influence of measurement error, we only included CpG dinucleotides with at least six and less than 200 reads (i.e., $5 < x_{uCpG} + x_{mCpG} < 200$). This measure of the methylation level was used for all methylation analyses except for the inference of differentially methylated regions (see section 2.8).

2.7 | Gene profile

We selected genes having at least one transcript with 5′- and one with a 3′ untranslated region (UTR), such that an average TSS and TTS could be defined with more confidence. This retained 3685 out of 13,161 annotated genes. Regions 5 kb up- and downstream of each gene were split into 100 bp non-overlapping segments. Gene bodies were also split into 100 bp segments, then normalized jointly in 99 windows based on rank along the gene profile. For the analysis of the relationship between gene expression and DNA methylation, genes with 5′ and 3′ UTRs were divided into three rank categories based on expression level: low (0-10th percentile), medium (10-90th percentile) and high (90-100th percentile). DNA methylation along the gene profile was plotted for these three expression categories. This was done separately for each individual, meaning that the same gene could be in different categories based on the expression level in a specific sample. We assessed statistical significance between expression levels by fitting local polynomial regression curves for each sample and expression category combination and then observing whether their 95% (CI) confidence intervals overlap (Cumming, 2009). We also performed Spearman rank correlations along the gene profile and assessed where the 95% CI did not overlap $\rho = 0$. To check for potential confounding effects on the relationship between DNA methylation and gene expression, we measured GC and CpG content along the gene profile.

2.8 | Differential methylation analysis

To identify candidate regions associated with environmental conditions, we performed a genome-wide search for differentially methylated regions (DMRs). We identified DMRs using BSmooth (Hansen et al., 2012) in the bsseq R package. The assumption with this method is that DNA methylation varies continuously approximated by a second-degree polynomial. In the analysis, BSmooth fits a local-likelihood smoother on at least 70 CpGs in a window of at least 2 kb to fit a curve of the methylation level. Thus, information on the smoothed methylation level at a particular CpG is based on a larger region. We defined DMRs as regions at the top two percentile of genome-wide absolute methylation differences, keeping only CpG sites where at least two samples each, for both groups, had a coverage >= 2, following the standard protocol for the programme (Hansen, 2022). This is done to ensure that the smoothed curve has

some read information at the focal point of the DMR. We further required DMRs to contain at least three CpGs with a mean methylation difference >= 10 percentage points. In addition, gaps between CpGs inside the same DMR were not allowed to be larger than 300 bp (Hansen et al., 2012). We investigated two contrasts in the DMR analysis: (1) between the treatments HDAL and LDAL, to find regions associated with larval density in ad libitum food conditions, and (2) between the treatments HDAL and HDLI, to find regions associated with larval food limitation in high-density conditions.

To test for enrichment of DMRs at annotated features, we used a resampling method to shuffle the DMRs $n = 1000$ times across the genome to obtain an expected amount of overlap with annotated features. We calculated a two-tailed empirical p-value as $2r/n$ for $r/n <= 0.5$ and $2(1 - r/n)$ for $r/n > 0.5$, where r is the number of replicates with an overlap greater than or equal to the overlap for the observed data (Ewens, 2003). Enrichment was defined as the following odds ratio:

$$\text{Odds ratio} = \frac{\frac{\text{Overlap between DMRs and annotation}}{\text{Total length of DMRs}}}{\frac{\text{Total length of annotation}}{\text{Genome length}}}.$$

Gene and transposable element (TE) annotation were obtained from Shipilina et al. (2022).

We tested whether DE genes were enriched in DMRs, which is expected if DMRs in *cis* (locally) are the main regulatory cause of DE genes or correlated with a change in expression. We did this by investigating if there was a significant overlap between DE and gene body as well as promoter DMRs.

2.9 | Gene ontology enrichment analysis and candidate gene identification

We performed a gene ontology enrichment (GO) analysis to investigate whether DMRs overlapping gene features and promoters were associated with genes involved in certain biological processes. We also performed GO analyses on DE genes and compared with GO analyses on DMRs. A list of all genes and associated gene ontology terms was used as reference set (Shipilina et al., 2022). All GO analyses were performed using the 'weight01' algorithm in topGO (Alexa et al., 2006). We used the Benjamini–Hochberg method to correct for multiple testing to a family-wise error rate (FWER) of 0.1. To find candidate genes, we focused on genes with multiple DMRs, which are more likely to be biologically relevant (i.e., there is a low probability that many DMRs appear in the same gene by chance). Under the conservative assumption that all identified DMRs overlap CDSs, the expected frequency is ~0.2 DMRs per gene. We selected genes with four or more DMRs overlapping the coding sequence in each comparison.

2.10 | Statistical analysis

Statistical analyses except permutations were performed in R v.4.1.3 and/or v.4.2.1 (R Core Team, 2020).

—WILEY—

3 | RESULTS

3.1 | DNA methylation is concentrated in genic regions

The genome-wide CpG methylation level in *V. cardui* in all nine samples was low (~0.8%). However, a detailed analysis of variation in methylation across genomic regions revealed a considerably higher level in genic regions; both coding sequences (CDS), 3′ untranslated regions (UTRs) and promoters had ~10% methylation (Figure 1). We found that the levels of methylation also varied within gene regions with a considerable dip in methylation around the transcription start site (TSS; Figure 2). Most TEs had low levels of methylation (~0.6%; Figure 1), just slightly higher than the unmethylated cytosine non-conversion fraction (~0.2–0.3% per sample). Overall methylation levels varied between treatments and more stressful conditions resulted in slightly higher methylation levels, especially in promoters, 3′ UTRs and CDS (Figure 1). See section 3.3 for a detailed analysis of differential methylation between treatments.

3.2 | Methylation levels along the gene body vary with gene expression levels

All genes, regardless of expression level, were unmethylated close to the TSS (Figure 2). Genes with a higher expression level also had on average higher methylation levels in the promoters and the gene bodies. However, just downstream of the TSS, genes with medium expression levels showed the highest degree of methylation. Local polynomial regressions were fit to each combination of sample and expression category, which confirmed that these patterns were significantly different (Figure S1). Spearman rank correlations along the gene profile showed a slight negative relationship at the 5′ end of the gene body shifting through a section without a monotonic relationship (corresponding to the region where genes with medium
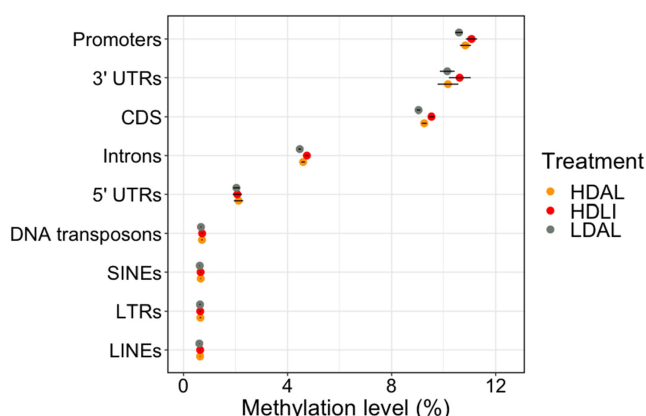
expression showed a higher methylation level), towards a positive relationship at the 3′ end (Figure S2). To check for a potential confounding variables, we plotted the CpG density and GC content gene profiles for the three expression categories (Figure S3). Both variables where on average highest for genes with the lowest expression indicating that neither can explain the relationship between expression and methylation observed here (Figure 2).

3.3 | Most differentially methylated regions show higher methylation levels under stress

We identified DMRs to pinpoint genomic regions involved in the response to environmental stress during larval development. Most of the ~3000 DMRs in each contrast showed a higher methylation level in the more stressful condition (Binomial tests $p < .0005$; Table S2; Figure 3). We also analysed where in the genome DMRs were localized and found a significant enrichment (Monte-Carlo $p < .003$, 1000 replicates) within different gene regions (Figure S4), with 3′ UTRs, CDSs and promoters particularly enriched. This is expected to some extent given the higher average methylation level in these regions which translates to more opportunities for substantial changes in methylation levels.

3.4 | Functional analysis of genes enriched in differentially methylated regions between treatments reveals candidate pathways and genes involved in environmental response

We performed GO analysis to investigate what genes and biological processes were involved in response to environmental stress during larval development. For DMRs overlapping CDSs in both comparisons, we found enrichment of genes involved in peptide transport (Tables S3 and S4). In addition, for the HDAL-LDAL contrast, we also found enrichment of genes involved in several processes related to translation (Table S4). To find candidate genes, we focused on genes with four or more DMRs (Tables S5 and S6). We found several genes with known expression in head, antennae, and central nervous system in *Drosophila* (Tables S5 and S6). In the HDAL-HDLI comparison, we found two candidate genes involved in flight behaviour: *rigour mortis*, which regulates mRNA splicing, and *mettl3*, an enzyme with the ability to methylate adenosines in mRNA (Table S6).

3.5 | Differentially expressed genes are not enriched in differential methylation

We performed DE analysis to investigate the relationship between methylome- and transcriptome profile shifts between treatments. In both contrasts, 35 genes were DE. Eight of those genes were DE in both contrasts. Subsequent GO analysis based on DE genes revealed no significant enrichment of any specific gene categories for

FIGURE 1 Average methylation levels within different genomic elements for each treatment group. Black lines through dots represent the 95% confidence interval of the mean methylation level per annotation feature and treatment.

the HDAL-LDAL comparison. In the HDAL-HDLI contrast, several biological processes related to translation and nonsense-mediated decay of mRNA were significantly enriched (Table S7). Translational
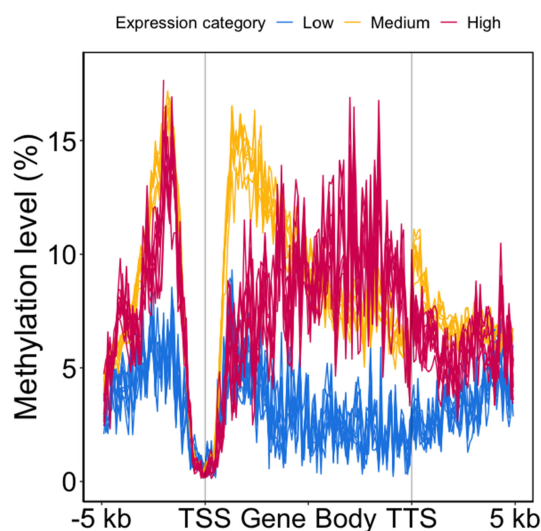

FIGURE 2 Methylation profiles of all genes with UTR annotation ($n = 3685$). The three categories represent genes with low (blue; 0–10th percentile), medium (yellow; 10–90th percentile) and high (red; 90–100th percentile) expression levels. TSS and TTS are the inferred transcription start- and termination sites respectively.
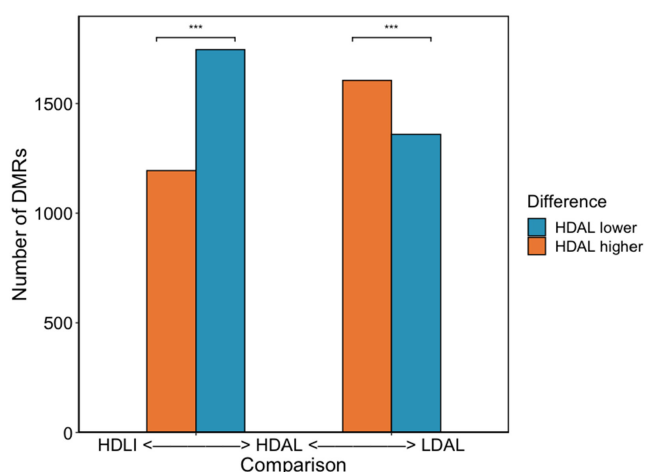

FIGURE 3 Differentially methylated regions in the two contrasts HDAL-LDAL and HDAL-HDLI respectively. Treatment groups are sorted based on level of stress, with the most stressful condition to the left (HDLI) and the least stressful (LDAL) to the right. Note that the most stressful condition in each contrast shows a significant (***$p < .0005$) excess of differentially methylated regions (DMRs). HDAL is used as the reference for each contrast.

regulation was also found in the GO analysis of DMRs but in the HDAL-LDAL contrast, perhaps indicating some functional concordance with DMRs. The differentially methylated genes *rigour mortis* and *mettle3* were not significantly DE. We also tested whether DE genes were enriched in DMRs. The few overlapping regions we identified did not significantly (Fisher's exact test; $p > .05$) deviate from the random expectation (Table 1). In addition, we found no DMRs in promoter regions of DE genes. Consequently, DE genes were not enriched in DMRs.

4 | DISCUSSION

In this study, we characterized the DNA methylome of a nearly cosmopolitan migratory butterfly species and tested whether methylation could be a mechanism to store or react to distinct environmental signals during larval development.

4.1 | Diversification of methylation within Lepidoptera revealed by novel relationship with expression

Our results showed that the association between methylation and gene expression in *V. cardui* is different from what has previously been observed in *B. mori* (Xiang et al., 2010), indicating that DNA methylation patterns have diversified within Lepidoptera. Methylation across the gene body has been shown to be positively associated with gene expression in *B. mori*, while we observed a more complex pattern in *V. cardui*, with intermediately expressed genes having the highest methylation at the 5′ end of genes and a positive relationship at the 3′ end. Our results support that the gene body methylation level in *V. cardui* can be determined by the occupancy level of RNA polymerase, as has been suggested in humans (Jjingo et al., 2012). According to this model, the opportunity of methylation should be greatest when chromatin is open and RNA polymerase occupancy is low, which likely reflects the situation in genes with intermediate expression. For highly expressed genes, the chromatin is probably also highly accessible, but RNA polymerase occupancy levels are high enough to prevent DNA methyltransferase activity. Evidence from several studies indicates that RNA polymerase occupancy is especially high at the 5′ end of expressed genes (Adelman & Lis, 2012). This could explain the observed shifting patterns of methylation levels between 5′ and 3′ ends of the gene body for different levels of gene expression in *V. cardui*. Several hypotheses could explain why the previous studies on *B. mori* may have failed

TABLE 1 Overlap between DMRs and DE genes

Comparison	Odds ratio	#DE genes with gene body DMRs	#DE genes	#genes with DMRs
HDAL-LDAL	0.54 (0.11–1.71)	3	35	2091
HDAL-HDLI	0.74 (0.19–2.09)	4	35	2020

Note: Parentheses in the first column show the 95% confidence interval in a Fisher's exact test.

to detect the complex pattern here described for *V. cardui*. (i) There is a difference between these lineages in the mechanism regulating methylation. (ii) The mechanism varies according to the tissue and/or developmental stage analysed. (iii) The pattern varies according to the different gene sets used. Here we restricted the analysis to the 3685 genes with annotated 5′ and 3′ UTRs to increase accuracy in the inference of TSS and TTS. As previously mentioned, the function of gene body methylation is still debated (Neri et al., 2017; Zilberman, 2017). Targeted partial demethylation at three gene body loci in a *B. mori* embryonic cell line leads to slightly increased levels of gene expression (1.2–1.7 fold) compared to controls (Liu et al., 2019). While only at a few loci, these results challenge other results from *B. mori* (Xiang et al., 2010, 2013; Xu et al., 2021), but in line with the relationship between DNA methylation and gene expression, we observed here among genes. It is also possible that the relationship among genes does not reflect the relationship between treatments at a single gene. More data are needed before any decisive conclusions can be drawn on the function(s) of gene body methylation in Lepidoptera. We also observed both a higher level as well as variation between expression categories of methylation in promoters of *V. cardui* genes compared to *B. mori* genes which appear to predominantly be unmethylated (Xiang et al., 2010). Unlike vertebrates (e.g., Kaluscha et al., 2022), *V. cardui* has a higher promoter methylation in genes with intermediate and high expression levels. This pattern is also seen in the postman butterfly (*Heliconius melpomene*) but its promoter methylation is considerably lower regardless of expression level (Lewis et al., 2020). Consequently, our results show that also promoter methylation levels differ significantly between species/tissues in Lepidoptera. The functional importance of this difference remains to be determined.

4.2 | Stress during development shifts the methylation profile in adults but evidence for a relationship with differential expression remains elusive

Our results showed that environmental signals during larval development can be stored as methylation patterns in the adult. Further confidence for this conclusion comes from the fact that several genes showing multiple DMRs are expressed in homologous structures of the head in *Drosophila melanogaster* (Gramates et al., 2022). Both treatments that induced stress, especially the HDLI treatment, resulted in an increase in methylation in the identified DMRs. This could be an effect of a prolonged development time induced by stress (Kelly & Debinski, 1999), or a more direct effect on, for example, gene expression causing correlated changes in methylation. The identified DMRs could represent regions that either affect regulation at the time of sampling, or at any previous point during development. In contrast, the level of gene expression is likely a much more ephemeral signal (Bird, 2002; Chan et al., 2018). Thus, an absence of enrichment in overlap between DE and DMRs does not mean that differential methylation has not played a regulatory

role. The relationship between gene expression and methylation could alternatively be dominated by *trans* changes or changes at distal regulatory sites. Effects of methylation at *cis*-regulatory elements are transcription-factor dependent in humans (Yin et al., 2017), and largely unexplored in arthropods (Xu et al., 2018). Both the DMRs and DE analyses revealed a concordance with enrichment of genes associated with translational regulation, albeit in two different contrasts that include the HDAL treatment group (see section 4.4). Interestingly, methylated genes in *B. mori*, *D. pulex*, honey bee (*Apis mellifera*) and fire ant (*Solenopsis* invicta) are also enriched in ontologies related to translation (Harney et al., 2022; Hunt et al., 2013; Xiang et al., 2010). Whether or not methylation has a causal role in determining gene expression levels, the concordance on the level of GO between DMRs and DEs, indicates that they are co-regulated to some extent.

4.3 | Differential methylation in an RNA methyltransferase gene: *mettl3*

An interesting candidate gene for further study focusing on the interplay between expression and DNA methylation is *mettl3*, in which we found four DMRs overlapping CDS in the HDAL-HDLI contrast. A recent study (Deng et al., 2022) using human cell cultures showed that *mettl3*-mediated RNA methyladenosine formation leads to demethylation of DNA in neighbouring genomic loci of methylated transcripts, which affects chromatin accessibility and gene expression. If a similar mechanism acts in butterfly cells, it is possible that *mettl3* autoregulates by methylating its own transcript, which in turn leads to increased DNA demethylation at the *mettle3* locus. However, *mettl3* was not DE in the HDAL-HDLI contrast, indicating that differences in *mettle3* activity caused by food availability probably occur prior to the adult stage.

4.4 | Caveats

Separation of the marginal effects of density and food limitation was not possible in our experimental design. For example, in both the HDAL-LDAL and HDAL-HDLI comparisons we observed that peptide transport function was enriched for genes with CDS DMRs. This could be due to similar molecular systems being involved in responses to density (in ad libitum food conditions) as well food limitation (in high-density conditions), but to confidently assess this we would need to investigate marginal effects of each treatment which requires a 2 ×2 factorial design. We also cannot exclude the effects of the same HDAL samples being reused in both comparisons. For example, eight of 35 genes were DE in both contrasts as well as 513 of ~3000 DMRs (when requiring >80% reciprocal overlap). These subsets could either represent true biological signals or correlated false positives in both comparisons. However, it is reassuring for our conclusions that this set represents a minority of DE genes and DMRs. In addition, while the results presented here point to

methylation being part of the response to larval treatments, they do not provide causal evidence. Shifts in methylation may instead be downstream effects of changes in gene expression and/or other epigenetic mechanisms, such as histone modifications (Näsvall et al., 2023; Xu et al., 2021).

4.5 | Outlook

Whether the methylation and expression differences between treatment groups we observe here influence migratory behaviour remains to be proven. However, we have shown, as a first step, that high density and food limitation induce molecular programmes in the heads of larvae that echoes throughout development. We did uncover two interesting candidate genes, both involved in flight behaviour in *Drosophila melanogaster*. Future studies could investigate their expression and methylation profiles across ontogenetic stages and assess their potential functions associated with migratory behaviour using, for example, RNAi or CRISPR-Cas9 methodologies (Gudmunds et al., 2022).

5 | CONCLUSION

Here, we provide a detailed characterization of genome-wide methylation profiles in the painted lady butterfly, a key species for insect migration research. Besides giving a broad sense of how methylation is distributed in a butterfly, we also show that gene-body methylation is variably associated with gene expression in Lepidoptera. Most importantly, we contrast multiple treatment groups and identify differentially methylated regions that provide a first glimpse into how environmental cues can alter the epigenetic landscape and be transmitted from larval to adult stages. The identified candidate regions will provide important starting points for investigating potential mechanistic links between epigenetic modifications and regulation of migration.

AUTHOR CONTRIBUTIONS

JB wrote and revised the original manuscript, and carried out experimental design and data analysis. NB was involved in writing, experimental design, funding and laboratory work. LH and YZ were involved in writing, laboratory work and experimental design. GT carried out writing, funding and data acquisition. RV carried out writing and funding.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

Bisulphite and RNA-seq data are available at the European Nucleotide Archive under study id ERP142326. Scripts are available on Github in the following repository: https://github.com/JesperBoman/DNA-methylation-of-the-Painted-Lady.

ORCID

Jesper Boman https://orcid.org/0000-0002-0537-8219
Roger Vila https://orcid.org/0000-0002-2447-4388
Gerard Talavera https://orcid.org/0000-0003-1112-1345
Niclas Backström https://orcid.org/0000-0002-0961-8427

REFERENCES

Adelman, K., & Lis, J. T. (2012). Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. *Nature Reviews Genetics*, *13*(10), 720–731. https://doi.org/10.1038/nrg3293

Alexa, A., Rahnenführer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, *22*(13), 1600–1607. https://doi.org/10.1093/bioinformatics/btl140

Bewick, A. J., Vogel, K. J., Moore, A. J., & Schmitz, R. J. (2017). Evolution of DNA methylation across insects. *Molecular Biology and Evolution*, *34*(3), 654–665. https://doi.org/10.1093/molbev/msw264

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes and Development*, *16*(1), 6–21. https://doi.org/10.1101/gad.947102

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., Zhang, G., & Reinberg, D. (2012). Genome-wide and caste-specific DNA Methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Current Biology, 22*(19), 1755–1764. https://doi.org/10.1016/J.CUB.2012.07.042

Cao, Y., Luo, L., & Guo, J. (1996). Performance of adult reproduction and flight in relation to larval nutrition in the oriental armyworm, *Mythimna separata* (Walker). *Acta Entomologica Sinica*, *39*(1), 105–108. http://www.insect.org.cn

Catania, S., Dumesic, P. A., Pimentel, H., Nasif, A., Stoddard, C. I., Burke, J. E., Diedrich, J. K., Cook, S., Shea, T., Geinger, E., Lintner, R., Yates, J. R., Hajkova, P., Narlikar, G. J., Cuomo, C. A., Pritchard, J. K., & Madhani, H. D. (2020). Evolutionary persistence of DNA methylation for millions of years after ancient loss of a De novo methyltransferase. *Cell*, *180*(2), 263–277.e20. https://doi.org/10.1016/j.cell.2019.12.012

Chan, L. Y., Mugler, C. F., Heinrich, S., Vallotton, P., & Weis, K. (2018). Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. *eLife*, *7*, e32536. https://doi.org/10.7554/eLife.32536

Chowdhury, S., Zalucki, M. P., Amano, T., Poch, T. J., Lin, M.-M., Ohwaki, A., Lin, D.-L., Yang, L., Choi, S.-W., Jennions, M. D., & Fuller, R. A. (2022). Trends and progress in studying butterfly migration. *Integrative Conservation*, *1*(1), 8–24. https://doi.org/10.1002/INC3.13

Cumming, G. (2009). Inference by eye: Reading the overlap of independent confidence intervals. *Statistics in Medicine*, *28*(2), 205–220. https://doi.org/10.1002/sim.3471

Deng, S., Zhang, J., Su, J., Zuo, Z., Zeng, L., Liu, K., Zheng, Y., Huang, X., Bai, R., Zhuang, L., Ye, Y., Li, M., Pan, L., Deng, J., Wu, G., Li, R., Zhang, S., Wu, C., Lin, D., … Zheng, J. (2022). RNA m6A regulates transcription via DNA demethylation and chromatin accessibility. *Nature Genetics*, *54*(9), 1427–1437. https://doi.org/10.1038/s41588-022-01173-1

Denlinger, D. L. (2002). Regulation of diapause. *Annual Review of Entomology*, *47*, 93–122. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303–0139, USA. https://doi.org/10.1146/annurev.ento.47.091201.145137

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology*, *35*(4), 316–319. https://doi.org/10.1038/nbt.3820

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Du, B., Ding, D., Ma, C., Guo, W., & Kang, L. (2022). Locust density shapes energy metabolism and oxidative stress resulting in divergence of flight traits. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(1), e2115753118. https://doi.org/10.1073/PNAS.2115753118/SUPPL_FILE/PNAS.2115753118.SAPP.PDF

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/BIOINFORMATICS/BTW354

Ewens, W. J. (2003). On estimating P values by Monte Carlo methods. *The American Journal of Human Genetics*, *72*(2), 496–498. https://doi.org/10.1086/346174

García-Berro, A., Talla, V., Vila, R., Wai, H. K., Shipilina, D., Chan, K. G., Pierce, N. E., Backström, N., & Talavera, G. (2023). Migratory behaviour is positively associated with genetic diversity in butterflies. *Molecular Ecology*, *32*(3), 560–574. https://doi.org/10.1111/MEC.16770

Gramates, L. S., Agapite, J., Attrill, H., Calvi, B. R., Crosby, M. A., dos Santos, G., Goodman, J. L., Goutte-Gattat, D., Jenkins, V. K., Kaufman, T., Larkin, A., Matthews, B. B., Millburn, G., Strelets, V. B., the FlyBase Consortium, Perrimon, N., Gelbart, S. R., Agapite, J., Broll, K., … Lovato, T. (2022). FlyBase: A guided tour of highlighted features. *Genetics*, *220*(4), iyac035. https://doi.org/10.1093/GENETICS/IYAC035

Gudmunds, E., Wheat, C. W., Khila, A., & Husby, A. (2022). Functional genomic tools for emerging model species. *Trends in Ecology and Evolution*, *37*(12), 1104–1115. Elsevier Current Trends. https://doi.org/10.1016/j.tree.2022.07.004

Hansen, K. D. (2022). *Analyzing WGBS data with bsseq*. Bioconductor. https://www.bioconductor.org/packages/release/bioc/vignettes/bsseq/inst/doc/bsseq_analysis.html

Hansen, K. D., Langmead, B., & Irizarry, R. A. (2012). BSmooth: From whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biology*, *13*(10), R83. https://doi.org/10.1186/gb-2012-13-10-R83

Harney, E., Paterson, S., Collin, H., Chan, B. H. K., Bennett, D., & Plaistow, S. J. (2022). Pollution induces epigenetic effects that are stably transmitted across multiple generations. *Evolution Letters*, *6*(2), 118–135. https://doi.org/10.1002/EVL3.273

Hasegawa, K. (1957). The diapause hormone of the silkworm, *Bombyx mori*. *Nature*, *179*(4573), 1300–1301. Nature Publishing Group. https://doi.org/10.1038/1791300b0

Hu, G., Stefanescu, C., Oliver, T. H., Roy, D. B., Brereton, T., Van Swaay, C., Reynolds, D. R., & Chapman, J. W. (2021). Environmental drivers of annual population fluctuations in a trans-Saharan insect migrant. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(26), e2102762118. https://doi.org/10.1073/pnas.2102762118

Hunt, B. G., Glastad, K. M., Yi, S. V., & Goodisman, M. A. D. (2013). Patterning and regulatory associations of DNA methylation are mirrored by histone modifications in insects. *Genome Biology and Evolution*, *5*(3), 591–598. https://doi.org/10.1093/gbe/evt030

Jiang, X., Luo, L., Zhang, L., Sappington, T. W., & Hu, Y. (2011). Regulation of migration in *Mythimna separata* (Walker) in China: A review integrating environmental, physiological, hormonal, genetic, and molecular factors. *Environmental Entomology*, *40*(3), 516–533. https://doi.org/10.1603/EN10199

Jjingo, D., Conley, A. B., Yi, S. V., Lunyak, V. V., & King Jordan, I. (2012). On the presence and role of human gene-body DNA methylation. *Oncotarget*, *3*(4), 462–474. https://doi.org/10.18632/oncotarget.497

Jones, C. M., Lim, K. S., Chapman, J. W., & Bass, C. (2018). Genome-wide characterization of DNA methylation in an invasive lepidopteran Pest, the cotton bollworm *Helicoverpa armigera*. *G3: Genes, Genomes, Genetics*, *8*(3), 779–787. https://doi.org/10.1534/G3.117.1112

Kaluscha, S., Domcke, S., Wirbelauer, C., Burger, L., & Schübeler, D. (2022). Direct inhibition of transcription factor binding is the dominant mode of gene and repeat repression by DNA methylation. *Nature Genetics*, *54*(12), 1895–1906. https://doi.org/10.1038/s41588-022-01241-6

Kelly, L., & Debinski, D. M. (1999). Effects of larval food-limitation on *Vanessa cardui* linnaeus (Lepidoptera: Nymphalidae). *American Midland Naturalist*, *141*(2), 315–322. https://doi.org/10.1674/0003-0031(1999)141[0315:EOLFLO]2.0.CO;2

Kozak, G. M., Wadsworth, C. B., Kahne, S. C., Bogdanowicz, S. M., Harrison, R. G., Coates, B. S., & Dopman, E. B. (2019). Genomic basis of circannual rhythm in the European corn borer moth. *Current Biology*, *29*(20), 3501–3509.e5. https://doi.org/10.1016/J.CUB.2019.08.053

Krueger, F., & Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for bisulfite-Seq applications. *Bioinformatics*, *27*(11), 1571–1572. https://doi.org/10.1093/bioinformatics/btr167

Leal, L., Talla, V., Källman, T., Friberg, M., Wiklund, C., Dincă, V., Vila, R., & Backström, N. (2018). Gene expression profiling across ontogenetic stages in the wood white (*Leptidea sinapis*) reveals pathways linked to butterfly diapause regulation. *Molecular Ecology*, *27*(4), 935–948. https://doi.org/10.1111/mec.14501

Lehmann, P., Pruisscher, P., Posledovich, D., Carlsson, M., Käkelä, R., Tang, P., Nylin, S., Wheat, C. W., Wiklund, C., & Gotthard, K. (2016). Energy and lipid metabolism during direct and diapause development in a pierid butterfly. *Journal of Experimental Biology*, *219*(19), 3049–3060. https://doi.org/10.1242/jeb.142687

Lewis, S. H., Ross, L., Bain, S. A., Pahita, E., Smith, S. A., Cordaux, R., Miska, E. A., Lenhard, B., Jiggins, F. M., & Sarkies, P. (2020). Widespread conservation and lineage-specific diversification of genome-wide DNA methylation patterns across arthropods. *PLoS Genetics*, *16*(6), e1008864. https://doi.org/10.1371/journal.pgen.1008864

Liu, Y., Ma, S., Chang, J., Zhang, T., Chen, X., Liang, Y., & Xia, Q. (2019). Programmable targeted epigenetic editing using CRISPR system in *Bombyx mori. Insect Biochemistry and Molecular Biology*, *110*, 105–111. https://doi.org/10.1016/J.IBMB.2019.04.013

Lohse, K., Wright, C., Talavera, G., & García-Berro, A. (2021). The genome sequence of the painted lady, Vanessa cardui Linnaeus 1758. *Wellcome Open Research*, *6*, 324. https://doi.org/10.12688/wellcomeopenres.17358.1

López-Mañas, R., Pascual-Díaz, J. P., García-Berro, A., Bahleman, F., Reich, M. S., Pokorny, L., Bataille, C. P., Vila, R., Domingo-Marimon, C., & Talavera, G. (2022). Erratic spatiotemporal vegetation growth anomalies drive population outbreaks in a trans-Saharan insect migrant. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(19), e2121249119. https://doi.org/10.1073/pnas.2121249119

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lü, W., Meng, L., Jiang, X., Cheng, Y., & Zhang, L. (2022). Larval crowding did not enhance adult migration propensity in Spodoptera frugiperda. *Insects*, *13*(7), 581. https://doi.org/10.3390/insects13070581

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Menchetti, M., Guéguen, M., & Talavera, G. (2019). Spatio-temporal ecological niche modelling of multigenerational insect migrations. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1910), 20191583. https://doi.org/10.1098/rspb.2019.1583

Näsvall, K., Shipilina, D., Vila, R., Talavera, G., Backstrom, N., & Backström, N. (2023). Resource availability affects activity profiles of regulatory elements in a long-distance butterfly migrant. Authorea Preprints https://doi.org/10.22541/AU.167827909.99815237/V1

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., & Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. *Nature*, *543*(7643), 72–77. https://doi.org/10.1038/nature21373

Oldroyd, B. P., & Yagound, B. (2021). The role of epigenetics, particularly DNA methylation, in the evolution of caste in insect societies. *Philosophical Transactions of the Royal Society B*, *376*(1826), 20200115. https://doi.org/10.1098/RSTB.2020.0115

Parker, W. E., & Gatehouse, A. G. (1985). The effect of larval rearing conditions on flight performance in females of the African armyworm, *Spodoptera exempta* (Walker) (Lepidoptera: Noctuidae). *Bulletin of Entomological Research*, *75*(1), 35–48. https://doi.org/10.1017/S0007485300014152

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

R Core Team. (2020). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. https://www.r-project.org/

Raine, A., Manlig, E., Wahlberg, P., Syvänen, A. C., & Nordlund, J. (2017). SPlinted ligation adapter tagging (SPLAT), a novel library preparation method for whole genome bisulphite sequencing. *Nucleic Acids Research*, *45*(6), 36. https://doi.org/10.1093/nar/gkw1110

Reppert, S. M., & de Roode, J. C. (2018). Demystifying monarch butterfly migration. *Current Biology*, *28*(17), R1009–R1022. https://doi.org/10.1016/J.CUB.2018.02.067

Sarkies, P. (2022). Encyclopaedia of eukaryotic DNA methylation: From patterns to mechanisms and functions. *Biochemical Society Transactions*, *50*(3), 1179–1190. https://doi.org/10.1042/BST20210725

Shipilina, D., Näsvall, K., Höök, L., Vila, R., Talavera, G., & Backström, N. (2022). Linkage mapping and genome annotation give novel insights into gene family expansions and regional recombination rate variation in the painted lady (*Vanessa cardui*) butterfly. *Genomics*, *114*(6), 110481. https://doi.org/10.1016/j.ygeno.2022.110481

Stefanescu, C., Páramo, F., Åkesson, S., Alarcón, M., Ávila, A., Brereton, T., Carnicer, J., Cassar, L. F., Fox, R., Heliölä, J., Hill, J. K., Hirneisen, N., Kjellén, N., Kühn, E., Kuussaari, M., Leskinen, M., Liechti, F., Musche, M., Regan, E. C., … Chapman, J. W. (2013). Multi-generational long-distance migration of insects: Studying the painted lady butterfly in the Western Palaearctic. *Ecography*, *36*(4), 474–486. https://doi.org/10.1111/J.1600-0587.2012.07738.X

Stefanescu, C., Puig-Montserrat, X., Samraoui, B., Izquierdo, R., Ubach, A., & Arrizabalaga, A. (2017). Back to Africa: Autumn migration of the painted lady butterfly *Vanessa cardui* is timed to coincide with an increase in resource availability. *Ecological Entomology*, *42*(6), 737–747. https://doi.org/10.1111/een.12442

Stefanescu, C., Ubach, A., & Wiklund, C. (2021). Timing of mating, reproductive status and resource availability in relation to migration in the painted lady butterfly. *Animal Behaviour*, *172*, 145–153. https://doi.org/10.1016/j.anbehav.2020.12.013

Suchan, T., Talavera, G., Sáez, L., Ronikier, M., & Vila, R. (2019). Pollen metabarcoding as a tool for tracking long-distance insect migrations. *Molecular Ecology Resources*, *19*(1), 149–162. https://doi.org/10.1111/1755-0998.12948

Süess, P., Dircksen, H., Roberts, K. T., Gotthard, K., Nässel, D. R., Wheat, C. W., Carlsson, M. A., & Lehmann, P. (2022). Time- and temperature-dependent dynamics of prothoracicotropic hormone and ecdysone sensitivity co-regulate pupal diapause in the green-veined white butterfly Pieris napi. *Insect Biochemistry and Molecular Biology*, *149*, 103833. https://doi.org/10.1016/J.IBMB.2022.103833

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. *Nature Reviews Genetics*, *9*(6), 465–476. Nature Publishing Group. https://doi.org/10.1038/nrg2341

Talavera, G., Bataille, C., Benyamini, D., Gascoigne-Pees, M., & Vila, R. (2018). Round-trip across the Sahara: Afrotropical painted lady butterflies recolonize the Mediterranean in early spring. *Biology Letters*, *14*(6), 20180274. https://doi.org/10.1098/RSBL.2018.0274

Talavera, G., & Vila, R. (2017). Discovery of mass migration and breeding of the painted lady butterfly Vanessa cardui in the sub-Sahara: The Europe–Africa migration revisited. *Biological Journal of the Linnean Society*, *120*(2), 274–285. https://doi.org/10.1111/BIJ.12873

Urquhart, F. A. (1976). Found at last: The monarch's winter home. *National Geographic*, *150*(2), 161–173.

Watanabe, K. (1924). Studies on the voltinism of the silkworm, *Bombyx mori. Bull Sericult Exp Sta (Tokyo)*, *6*, 411–455.

Wiklund, C., & Friberg, M. (2022). Testing the migration syndrome: Comparative fecundity of migratory and non-migratory nymphaline butterflies. *Ecological Entomology*, *47*(6), 1061–1067. https://doi.org/10.1111/een.13195

Williams, C. B. (1930). *The migration of butterflies*. Oliver & Boyd.

Xiang, H., Li, X., Dai, F., Xu, X., Tan, A., Chen, L., Zhang, G., Ding, Y., Li, Q., Lian, J., Willden, A., Guo, Q., Xia, Q., Wang, J., & Wang, W. (2013). Comparative methylomics between domesticated and wild silkworms implies possible epigenetic influences on silkworm domestication. *BMC Genomics*, *14*(1), 1–11. https://doi.org/10.1186/1471-2164-14-646/FIGURES/4

Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., Zhang, H., Zhang, G., Li, D., Dong, Y., Zhao, L., Lin, Y., Cheng, D., Yu, J., Sun, J., Zhou, X., Ma, K., He, Y., Zhao, Y., … Wang, J. (2010). Single base–resolution

methylome of the silkworm reveals a sparse epigenomic map. *Nature Biotechnology*, *28*(5), 516–520. https://doi.org/10.1038/nbt.1626

Xu, G., Lyu, H., Yi, Y., Peng, Y., Feng, Q., Song, Q., Gong, C., Peng, X., Palli, S. R., & Zheng, S. (2021). Intragenic DNA methylation regulates insect gene expression and reproduction through the MBD/Tip60 complex. *iScience*, *24*(2), 102040. https://doi.org/10.1016/j.isci.2021.102040

Xu, G., Zhang, J., Lyu, H., Song, Q., Feng, Q., Xiang, H., & Zheng, S. (2018). DNA methylation mediates BmDeaf1-regulated tissue- and stage-specific expression of BmCHSA-2b in the silkworm, Bombyx mori. *Epigenetics and Chromatin*, *11*(1), 1–14. https://doi.org/10.1186/s13072-018-0202-4

Yamashita, O. (1996). Diapause hormone of the silkworm, Bombyx mori: Structure, gene expression and function. *Journal of Insect Physiology*, *42*(7), 669–679. https://doi.org/10.1016/0022-1910(96)00003-0

Yi, S. V. (2017). Insights into epigenome evolution from animal and plant Methylomes. *Genome Biology and Evolution*, *9*(11), 3189–3201. https://doi.org/10.1093/GBE/EVX203

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P. K., Kivioja, T., Dave, K., Zhong, F., Nitta, K. R., Taipale, M., Popov, A., Ginno, P. A., Domcke, S., Yan, J., Schübeler, D., Vinson, C., & Taipale, J. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science*, *356*(6337), eaaj2239. https://doi.org/10.1126/science.aaj2239

Yong, W. S., Hsu, F. M., & Chen, P. Y. (2016). Profiling genome-wide DNA methylation. *Epigenetics and Chromatin*, *9*(1). BioMed Central Ltd. https://doi.org/10.1186/s13072-016-0075-3

Zilberman, D. (2017). An evolutionary case for functional gene body methylation in plants and animals. *Genome Biology*, *18*(1), 1–3. BioMed Central Ltd. https://doi.org/10.1186/s13059-017-1230-2

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.