Received: 16 October 2017 | Revised: 17 January 2018 | Accepted: 19 January 2018

DOI: 10.1111/mec.14501

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Gene expression profiling across ontogenetic stages in the wood white (*Leptidea sinapis*) reveals pathways linked to butterfly diapause regulation

Luis Leal[1,2] | Venkat Talla[1] | Thomas Källman[3] | Magne Friberg[4] | Christer Wiklund[5] | Vlad Dincă[6] | Roger Vila[7] | Niclas Backström[1]

[1]Department of Evolutionary Biology, Evolutionary Biology Centre (EBC), Uppsala University, Uppsala, Sweden

[2]Department of Plant Ecology and Evolution, Evolutionary Biology Centre (EBC), Uppsala University, Uppsala, Sweden

[3]Department of Medical Biochemistry and Microbiology, Uppsala Biomedical Centre (BMC), Uppsala, Sweden

[4]Department of Biology, Biodiversity Unit, Lund University, Lund, Sweden

[5]Department of Zoology, Division of Ecology, Stockholm University, Stockholm, Sweden

[6]Department of Ecology and Genetics, University of Oulu, Oulu, Finland

[7]Institut de Biologia Evolutiva (CSIC-UPF), Barcelona, Spain

Correspondence
Niclas Backström, Department of Evolutionary Biology, Evolutionary Biology Centre (EBC), Uppsala University, Uppsala, Sweden.
Email: niclas.backstrom@ebc.uu.se

Funding information
SciLife Sweden Biodiversity Program, Grant/Award Number: Backstrom_2014; Swedish Research Council (VR), Grant/Award Number: 2013-4508; Knut and Alice Wallenberg Foundation

Abstract

In temperate latitudes, many insects enter diapause (dormancy) during the cold season, a period during which developmental processes come to a standstill. The wood white (*Leptidea sinapis*) is a butterfly species distributed across western Eurasia that shows photoperiod-induced diapause with variation in critical day-length across populations at different latitudes. We assembled transcriptomes and estimated gene expression levels at different developmental stages in experimentally induced directly developing and diapausing cohorts of a single Swedish population of *L. sinapis* to investigate the regulatory mechanisms underpinning diapause initiation. Different day lengths resulted in expression changes of developmental genes and affected the rate of accumulation of signal molecules, suggesting that diapause induction might be controlled by increased activity of monoamine neurotransmitters in larvae reared under short-day light conditions. Expression differences between light treatment groups of two monoamine regulator genes (*DDC* and *ST*) were observed already in instar III larvae. Once developmental pathways were irreversibly set at instar V, a handful of genes related to dopamine production were differentially expressed leading to a significant decrease in expression of global metabolic genes and increase in expression of genes related to fatty acid synthesis and sequestration. This is in line with a time-dependent (hour-glass) model of diapause regulation where a gradual shift in the concentration of monoamine neurotransmitters and their metabolites during development of larvae under short-day conditions leads to increased storage of fat, decreased energy expenditures, and ultimately developmental stasis at the pupal stage.

KEYWORDS
developmental plasticity, diapause, gene expression, hour-glass model, Lepidoptera, monoamine neurotransmitter

1 | INTRODUCTION

Characterization of the genetic underpinnings of adaptive phenotypic traits is fundamental for understanding the mechanisms of evolution and is a cornerstone of modern evolutionary biology research

Luis Leal and Venkat Talla authors contributed equally to this study.

(Ellegren & Sheldon, 2008; Pardo-Diaz, Salazar, & Jiggins, 2015). Knowledge about genes underlying adaptive phenotypes has recently increased as a result of technological advancements, allowing for the generation of genome-wide marker sets for almost any organism of interest (Byers, Xu, & Schluter, 2017). Still, the progress in identifying genotype-phenotype links is basically restricted to a handful of laboratory model organisms, because data from natural study populations rarely include both quantitative estimates of adaptive phenotypes and extensive genomic resources (Pardo-Diaz et al., 2015). Key to get a deeper understanding of the processes underlying evolutionary change in natural populations is to expand the field into a more diverse array of study systems and to combine different technologies and analytical approaches (Nadeau & Jiggins, 2010; Pardo-Diaz et al., 2015; Pavey, Collin, Nosil, & Rogers, 2010).

To respond to seasonal variation in temperature and availability of resources, many insects in temperate regions enter a state of dormancy – diapause – over the cold season, a period of developmental arrest marked by a sharp decrease in metabolic activity and heightened stress tolerance (e.g. Koštál, 2006). This adaptive response contributes to appropriate timing of the life cycle to seasonal climatic variation, optimization of propagative productivity and avoidance of local extinction in temporally adverse environments (Saunders, 2002, 2009; Van Dyck, Bonte, Puls, Gotthard, & Maes, 2014). Intriguingly, different insect species enter diapause at different stages in the life-cycle and many widespread insects show intraspecific variance in the number of directly developing generations across the distribution range (e.g. Koštál, 2006; Seppänen, 1969). Evidently, this shows that, despite plausibly being under considerable selective constraint, diapause is a flexible life-history trait with high evolvability. The initiation and termination of diapause are triggered by external environmental cues, predominantly the length of the photoperiod, with temperature and diet also playing significant roles in some species (Beck, 1980; Friberg & Wiklund, 2010; Kivelä, Svensson, Tiwe, & Gotthard, 2015; Saunders, 2002). This trait has long been the subject of scientific research (e.g. Beck, 1980; Saunders, 2002), and both genetic techniques and high-throughput genomic studies have been used to identify genes and functional categories that may be involved in diapause regulation (e.g. Denlinger, 2002; Denlinger & Armbruster, 2014; Meuti & Denlinger, 2013). Concurrently, the key endocrine mechanisms responsible for initiating and maintaining diapause have been described in detail (e.g. Denlinger, Yocum, & Rinehart, 2012; Hahn & Denlinger, 2011). Studies of reproductive diapause in *Drosophila melanogaster* have for example highlighted the role played by the insulin signaling pathway (Williams et al., 2006) and ecdysteroid modulators (Schmidt et al., 2008). However, our knowledge is still incomplete regarding the exact molecular mechanisms and pathways employed to process and store light and temperature data, and how this information is used for the induction and termination of diapause and/or direct development (Denlinger, Hahn, Merlin, Holzapfel, & Bradshaw, 2017; Emerson, Bradshaw, & Holzapfel, 2009; Koštál, 2011; Lehmann et al., 2018; Meuti & Denlinger, 2013).

The circadian clock – a biochemical oscillator that regulates daily metabolic processes in response to external cues (Hardin, Hall, & Rosbash, 1990; Young & Kay, 2001) – has been suggested to play an important role in diapause regulation (Bünning, 1936) with photoperiodic induction being controlled either by a single (external coincidence) or multiple (internal coincidence) circadian oscillators (Koštál, 2011; Saunders, 2012). Molecular cascades linking circadian clock genes and photoperiod-dependent developmental processes have been exposed in great detail for some systems and traits, e.g. the timing of flowering in *Arabidopsis* (Andrés & Coupland, 2012). While evidence has been accumulated suggesting circadian clock genes may be implicated in the photoperiodic switch mechanism in insects (e.g. Meuti & Denlinger, 2013; Saunders, Lewis, & Warman, 2004), identification of the specific pathways linking the circadian clock functional module and insect diapause has so far remained elusive (Bradshaw & Holzapfel, 2010b; Denlinger et al., 2017; Emerson et al., 2009; Koštál, 2011; Meuti & Denlinger, 2013). It is possible that circadian clock genes regulate daily rhythms and diapause pleiotropically, triggering distinct pathways, and/or that their role is tissue dependent (Bajgar, Jindra, & Dolezel, 2013; Bradshaw & Holzapfel, 2010a). Furthermore, lineage specific mechanisms might have evolved during the repeated incidences of diapause evolution, both in the long-term (between taxa with distinct diapause strategies) and the short-term (intraspecific variation in voltinism) perspective (Saunders & Bertossa, 2011).

One alternative to the circadian model of diapause regulation, the hour-glass hypothesis, asserts that photoperiodism is underpinned by a set of biochemical reactions taking place during nighttime, with diapause being induced when specific metabolites reach a certain threshold (Koštál, 2011; Lees, 1953, 1973; Veerman, 2001). The validity of this model rests on the identification of biomolecules involved in the cascade and in showing how concentration levels of end products can lead to the onset of diapause. Studies on variation in monoamine concentration carried out in the large cabbage white butterfly (*Pieris brassicae*), the silkworm (*Bombyx mori*) and the Chinese oak silkmoth (*Antheraea pernyi*) suggest that dopamine, an amine neurotransmitter and neuromodulator whose concentration varies according to light conditions, is highly implicated in processes leading to, maintaining, and terminating diapause (Isabel, Gourdoux, & Moreau, 2001; Noguchi & Hayakawa, 2001; Wang, Egi, Takeda, Oishi, & Sakamoto, 2015). A study based on the cotton bollworm (*Helicoverpa armigera*) associated diapause to reduced tricarboxylic acid activity (Xu, Lu, & Denlinger, 2012), and a recent investigation of the metabolome in green-veined whites (*Pieris napi*) found a link between diapause termination and alanine concentration (Lehmann et al., 2018). It is unclear whether the circadian clock and hourglass models of diapause regulation are competing or complementary hypotheses. It might be that observed changes in metabolite content occur downstream from a circadian trigger signal. It is also possible that circadian and hourglass components contribute additively to photoperiodism and, as the specifics of diapause regulation often differ among insects (Denlinger, 2002), the relative importance of each component could therefore be lineage specific.

The Eurasian wood-white butterfly (*Leptidea sinapis*) is an attractive system for studying local adaptation in general and diapause regulation in particular. *Leptidea sinapis* is subdivided into multiple geographically distinct ecotypes with different habitat utilization preferences and life-history characteristics reflecting variation in environmental conditions across the distribution range (Friberg, Bergman, Kullberg, Wahlberg, & Wiklund, 2008; Friberg & Wiklund, 2010). Similar to many other insect species (Aalberg Haugen & Gotthard, 2015; Bradshaw & Holzapfel, 2001), *L. sinapis* individuals in the northern part of the distribution range enter diapause during their pupal stage after only one round of reproduction (univoltine), while populations in the southern part have several generations (multivoltine) per season (Friberg et al., 2008; Tolman, 1997). Northern and southern *L. sinapis* populations also have different critical day lengths (the day length when 50% enter diapause and 50% go into direct development), with northern specimens requiring a substantially longer photoperiod to develop directly (Friberg et al., 2008). Populations originating in the southern range, but reared in conditions similar to the northern range, do not adjust their life-cycle accordingly, instead continuing to produce several broods before the onset of diapause (Friberg & Wiklund, personal communication). These results raise the hypothesis that regulatory mechanisms underpinning diapause and direct development in *L. sinapis* have a genetic basis and might be under different selective regimes in the different ecotypes. However, there is currently no data available on the genetic pathways underlying the switch that determines if individuals will enter diapause or opt for direct development. With the main aim of identifying key genes and pathways associated with diapause regulation, we compared gene expression profiles across ontogenetic stages in *L. sinapis* cohorts exposed to artificial treatments with distinct light regimes triggering either diapause or direct development.

2 | METHODS

2.1 | Experimental design

Three mated female *Leptidea sinapis* (Swedish ecotype) were collected in south central Sweden (Uppsala region) during spring 2016 and kept in captivity in separate 50*50*50 cm cages in a common garden with constant day light (18 h light, 6 h dark) and access to sugar water and stands of a commonly utilized host plant, birdsfoot trefoil (*Lotus corniculatus*), for egg laying. This population is generally univoltine, but a smaller second generation can sometimes be observed in nature and direct development can be triggered by long-day light treatment (Friberg et al., 2008). Eggs collected from each of the three unrelated females (families, from here on) were divided into two cohorts and reared in growth chambers with constant temperature (23°C), under either short-day (12 h daylight + 12 h dark) or long-day (22 h daylight + 2 h dark) light conditions (Fig. S1). Lights-off switch time was synchronized between treatments. Based on previous data in other butterfly species with a pupal overwintering stage, including the related pierid, green-veined white (*Pieris napi*), the critical point for diapause or direct development induction takes place during larval instar-IV (Friberg, Aalberg Haugen, Dahlerus, Gotthard, & Wiklund, 2011). Developing individuals were therefore monitored daily and samples were harvested the day after molting into larval stages instar-III and instar-V, and one day after pupation. Adult samples were also collected for the long-day light treatment the day after eclosion. Apart from instar-III, six samples (three males and three females, one from each family) were collected for each developmental stage and light treatment. Only three samples were collected for instar-III larvae (one from each family) since sexing based on morphology is impracticable at that stage (Friberg & Wiklund, personal observation). Hence, in total 36 samples were harvested representing three (short-day treatment) or four (long-day treatment) developmental stages (Fig. S1). At least ten additional individuals from the same families were raised under each light treatment to control that short-day and long-day treatments triggered diapause induction and direct development, respectively. All ten individuals raised under long-day light conditions went into direct development and enclosed after 7–10 days as pupae. All ten individuals reared under short-day light conditions entered diapause upon arriving to the pupal stage (and remained in diapause for several months).

2.2 | RNA extraction, library preparation and sequencing

Total RNA was extracted from the whole body using a modified acid phenol-chloroform protocol (Chomczynski & Mackey, 1995; Chomczynski & Sacchi, 1987), and individual libraries were generated using the Illumina TruSeq Stranded mRNA protocol (Illumina Inc.) with Poly-A selection and using dUTP for second strand cDNA synthesis. Libraries were sequenced on an Illumina HiSeq 2500 platform using a 126 bp paired-end approach (see Supplementary Methods for more details). General information about the *L. sinapis* genome, RNA samples, library preparation, and sequencing technology is shown in Table S1.

2.3 | Transcriptome assembly, read mapping and gene count consolidation

After initial quality assessment and removal of low quality sequence reads (Supplementary Methods), a de novo transcriptome was assembled using TRINITY, version 2.3.2 (Grabherr et al., 2011), implementing default parameters apart from the minimum contig length, which was set at 126 bp (read length). Multiple transcriptomes were initially produced with the aim of evaluating the robustness in transcriptome sampling and assembly. Separate assemblies were generated for each sample, then by experiment, developmental stage, lane, family, sex, and using all samples. All assemblies were characterized in terms of number of contigs, average GC content, and contig ExN50. The latter is similar to the standard N50 contig statistic but is computed using only the most highly expressed transcripts (e.g. Miller, Koren, & Sutton, 2010). ExN50 values were computed as

described in the TRINITY website (https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-Transcript-Quantification; accessed 2017-08-08). Finally, the samples with the highest ExN50 values and percentage of mapped reads (n = 25, Table S2), were used to create an optimized, consensus transcriptome. TRINITY generally produces transcriptomes with a number of contigs that far exceeds the estimated number of genes/gene-variants expected to be present in the genome (Haas, 2017). Contig trimming was carried out during downstream analysis on the basis of transcript coverage. Mapping of reads to the optimized transcriptome and quantification of transcript abundance was done using KALLISTO, version 0.43.0 (Bray, Pimentel, Melsted, & Pachter, 2016), with default parameter settings. After in silico functional annotation of the assembly (Supplementary Methods), all transcripts associated to the same gene were clustered into one unique entry, as were the associated GO terms. For each library, transcript counts were consolidated into counts-per-gene using the R/BIOCONDUCTOR package TXIMPORT, version 1.4.0 (Soneson, Love, & Robinson, 2015). Genes lacking annotation information and detected in just one sample were omitted from subsequent analysis.

After the extensive quality control and filtering procedure (Supplementary Methods), the average number of reads across libraries was close to 16 million (Table S3). Around 90% of these could be mapped to the *L. sinapis* genome assembly (Talla et al., 2017), indicating that this fragment set was suitable for transcriptome assembly. In total, 122 different transcriptomes were assembled, including an optimized reference transcriptome (Table S4). The total number of transcripts and predicted genes in the assemblies increased as we increased the number of libraries used (Table S5) and contig ExN50 values ranged between 1.28–1.85 kb (Fig. S2). Computing ExN50 allowed us to estimate the number of genes expressed in each sample more accurately (Haas, 2016), and the observed number of unique transcripts in experimental cohorts grouped by developmental stage ranged from 12,782 in adults to 16,092 in instar-V larvae (Fig. S3). This is within range of what could be expected based on the number of genes identified in other Lepidoptera taxa (Challis, Kumar, Dasmahapatra, Jiggins, & Blaxter, 2017; Cong et al., 2016; Nowell et al., 2017; Shen et al., 2016; The Heliconius Genome Consortium 2012). This analysis also showed that transcriptomes based on individual treatments or development stages had a smaller number of transcripts than when libraries were combined – the optimized transcriptome for example consisted of 21,189 transcripts (Fig. S3). On average, 85–96% of all reads in each of the libraries mapped to the corresponding transcriptome (Fig. S4).

2.4 | Replicate variability and clustering analyses

The biological variance across the three replicates associated to each treatment was assessed by selecting the 5,000 most highly expressed genes for each treatment and computing the Pearson correlation coefficient between each pair of replicates based on their expression profiles (expression levels were normalized by library size and log-10 transformed). For each set of three replicates, three inter-replicate correlation coefficients were thus computed, which were subsequently averaged to characterize each set with a single correlation value. The visualization of gene expression variance associated to each treatment was also done using principal component analysis (PCA) based on the *pca* function in MATLAB (https://se.mathworks.com/), based on the 5,000 genes with the highest expression variance across treatments. Expression levels were normalized by library size, log10 transformed, and zero-centered prior to PCA analysis. Hence, for each gene, the mean expression value was subtracted from the expression values associated to the different treatments. After this operation, down-regulated (up) genes had negative (positive) expression values and the average expression value across treatments was zero (for each gene).

2.5 | Differential gene expression (DGE) analysis

DGE analysis was carried out using DESEQ2, version 1.16.1 (Love, Huber, & Anders, 2014), a R/BIOCONDUCTOR package compatible with the *KALLISTO/tximport* suites (Love, Anders, Kim, & Huber, 2016). Counts per gene were rounded to the closest integer and low coverage genes with *baseMean* (count average across all samples) < 1 were removed prior to the DGE analysis. Batch effects were removed using the R/BIOCONDUCTOR package SVA, version 3.24.4 (Leek, Johnson, Parker, Jaffe, & Storey, 2012). SVA detects hidden technical, biological, and environmental batch effects, which are modeled with surrogate variables that are used to adjust expression values prior to downstream analysis (Leek et al., 2012). In the present study, the rationale behind using *sva* was to detect known sources of expression co-variation, such as family and sex, as well as additional unknown artifacts present in the data. The number of surrogate variables was estimated by *sva*, and varied between two and four across comparisons. An alternative way of dealing with batch effects would be to explicitly model biases as fixed effects using a linear model (i.e. 'expression ~ sex + family + daylength'). However, we had no information concerning the sex of instar-III samples, and family effects were not uniform. Hence, we decided to use SVA as sex and family effects are adjusted for individually in each sample (Leek et al., 2012). Furthermore, this procedure can detect unknown surrogate variables that would have been difficult to quantify in advance and incorporate in a linear model.

DGE analysis of SVA-adjusted expression values was carried out using the *DESeq* function (Love et al., 2014), after counts per gene were normalized by library size. The DESEQ2 protocol calculates average expression levels across all samples (*baseMean*), expression ratios across treatments (*log2FoldChange*) and significance values (p-value) for individual genes using the Wald test. The Benjamini-Hochberg adjustment method (Benjamini & Hochberg, 1995) as implemented in DESEQ2 was applied to account for multiple hypotheses testing, control the false discovery rate (FDR) (*alpha* parameter was set to 0.05) and calculate an adjusted significance value (p-adj). Only genes with p-adj < .05, *baseMean* > 10, and $|log2FoldChange| > 1$ were considered to be differentially expressed.

2.6 | Enrichment analysis of gene ontology terms

Enrichment analysis of GO terms was performed using the R/BIOCONDUCTOR package TOPGO, version 2.28.0 (Alexa & Rahnenfuhrer, 2016). The list of genes of interest was defined as the set of differentially expressed genes identified previously using DESEQ2. The 'gene universe' included all annotated genes with *baseMean* > 6, as well as the list of genes of interest. GO term significance was computed using topGO's Fisher-elim algorithm (see Supplementary Methods for more details) since it provides the lowest rate of false positives (Alexa, Rahnenfuhrer, & Lengauer, 2006), and indirectly corrects for multiple testing (Alexa & Rahnenfuhrer, 2016).

3 | RESULTS

3.1 | Replicate variability and clustering analyses

The degree of similarity between biological replicates varied considerably across treatments. In some cases, replicates linked to a specific treatment had very similar expression profiles (e.g. long-day instar-V females) while in other cases the replicates turned out to be rather different (e.g. long-day pupae) (Fig. S5). Concurrently, we found consistent significant correlations (*Pearson's r* = 0.31–0.93, *p*-value < .001) between the expression profiles of replicates originating from the same family, even when these were sampled from different developmental stages (Fig. S6). These results strongly suggest that expression values are likely to be modulated by several factors, including those associated to family traits. PCA analysis of these samples further confirmed that expression levels were influenced by multiple factors, including developmental stage, family, and sex (Fig. S7). All three factors had an influence on sample clustering but the relative importance of family and sex varied across ontogenetic stages. For example, the clustering of instar-V short-day samples was shaped by both sex and family factors, while for long-day treatment family was the most important factor. As expected, clustering by sex was more obvious during pupal and adult stages (Fig. S7). Taken together, the results from the correlation and clustering analyses suggest that controlling for the effect of both known (developmental stage, family, sex) and unknown confounding factors was necessary prior to carrying out DGE analysis aimed at identifying the key genes involved in the regulation of diapause and direct development.

3.2 | Differential gene expression analysis

Surrogate variables identified with *sva* were related to both family and sex. For example, one of the surrogate variables detected in the contrast of instar-III samples was significantly correlated with family membership (*F*-test, adjusted R^2 = .80, *p*-value < .05). Likewise, when assessing pupal and adult samples subjected to long-day light treatment, a surrogate variable was significantly correlated with sex (R^2 = .88, *p*-value < $1e^{-4}$). Hence, by employing *sva* we successfully identified surrogate variables highly correlated to confounding effects flagged during the PCA and correlation analyses. After correcting for batch effects, the number of differentially expressed genes for instar-III, instar-V and pupa in short-day versus long-day contrasts were 102, 67 and 168, respectively (Figure 1). The total number of genes differentially expressed across treatments, including annotation and GO information when available, is shown in Table 1. Genes differentially expressed across light treatments showed limited overlap across developmental stage contrasts; four genes for instar-III vs. instar-V (2.5% of all differentially expressed genes in both contrasts), and four genes (1.8%) for the instar-V vs. pupa comparison (Figure 2). There was a higher degree of overlap in differentially expressed genes across developmental stages when comparing treatment groups; 59 of 449 (13.1%) for development from instar-III to instar-V and 2,118 of 4,946 (42.8%) from instar-V to pupa (Figure 3). The 25 most significantly (lowest *p*-adj) differentially expressed genes across light treatments for instar-III, instar-V and pupa are listed in Table S6.

3.3 | Gene ontology enrichment analysis across light treatments

The GO enrichment analysis was performed separately for all eight sets of differentially expressed genes. The most significant GO categories for each set of differentially expressed genes across light treatments for different developmental stages are shown in Figure 4 (full list provided in Table S7). For instar-III, there were several terms associated to developmental processes, such as regulation of chitin and cuticle biosynthesis. Larvae raised under short-day conditions developed at a slower rate than those reared under long-day conditions and some of the GO categories here flagged as over-represented reflect differences in expression for genes involved in the regulation of metabolic and catabolic processes that control general

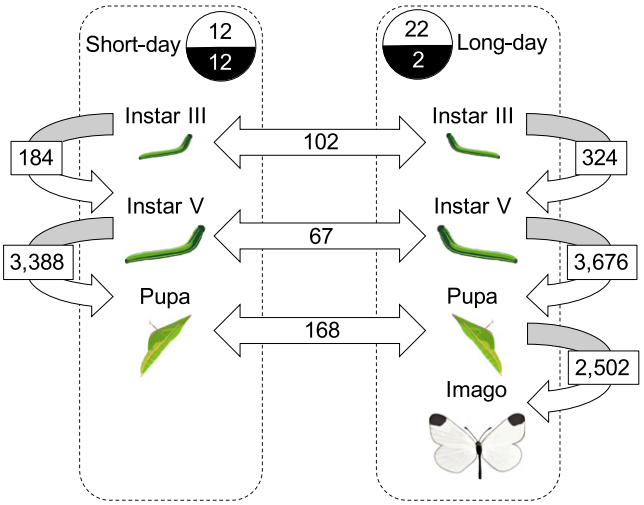

FIGURE 1 Total number of genes differentially expressed across developmental stages and light treatments after correcting for batch effects. Only genes with *p*-adj < .05, *baseMean* > 10, and |*log2FoldChange*| > 1 were considered to be differentially expressed. Left columns show short-day treatment and right columns show long-day treatment

TABLE 1 Total number of differentially expressed genes, annotated genes and genes with associated GO terms, across treatments. Expression values were corrected for batch effects prior to DGE analysis. Only genes with *p*-adj < .05, *baseMean* > 10, and |*log2FoldChange*| > 1 were considered to be differentially expressed

	Instar-III SD to LD	Instar-V SD to LD	Pupa SD to LD
# genes (*p*-adj < .05)	**102**	**67**	**168**
Up-regulated	25	50	76
Down-regulated	77	17	92
Annotated	76	35	106
With GO terms	63	31	70
# genes (*p*-adj < .01)	52	44	73

	SD Instar-III to V	SD Instar-V to pupa
# genes (*p*-adj < .05)	184	3,388
Up-regulated	119	1,331
Down-regulated	65	2,057
Annotated	121	2,454
With GO terms	96	2,189
# genes (*p*-adj < .01)	117	761

	LD Instar-III to V	LD Instar-V to pupa	LD Pupa to imago
# genes (*p*-adj < .05)	324	3,676	2,502
Up-regulated	269	1,239	1,361
Down-regulated	55	2,437	1,141
Annotated	242	2,591	1,727
With GO terms	200	2,255	1,438
# genes (*p*-adj < .01)	135	971	905

SD, short-day; LD, long-day.

growth and development. For example, *SUG* (Sugarbabe) which has been shown to repress genes involved in fat catabolism (Zinke, Schütz, Katzenberger, Bauer, & Pankratz, 2002), was up-regulated in short-day treatment larvae, and *AMYP* (Alpha-amylase), a gene implicated in the breakdown of starch (Powell & Andjelković, 1983), was down-regulated in short-day treatment larvae. *DESAT1* (Desaturase 1) has been shown to be involved in lipid metabolic processes favoring fat storage (Parisi et al., 2013) and was up-regulated in short-day light conditions, and so were *P49010* (*Beta-GlcNAcase*), *CHT2* (Chitinase 2), and *CHT10* (Chitinase 10), all genes coding for chitinases involved in chitin catabolism and nutrient recycling during ecdysis (Merzendorfer & Zimoch, 2003). A second set of significantly over-represented GO terms was related to processes involving neurotransmitters and neuromodulators (Figure 4, Table S7). Particularly well represented were categories linked to monoamine metabolic processes, such as those involved in synthesis of dopamine and serotonin. Two main genes were associated with these GO terms; *DDC* (Aromatic-L-amino-acid decarboxylase) and *ST* (Scarlet). Both were up-regulated in instar-III specimens under short-day treatment.

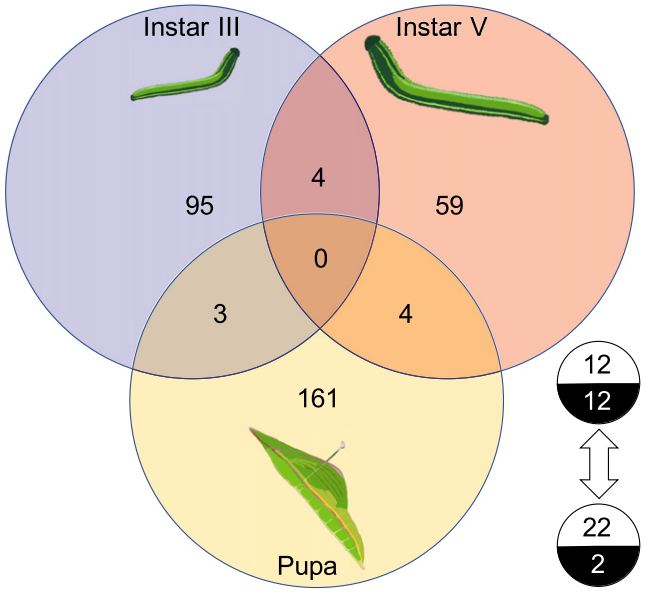

FIGURE 2 Venn diagram showing the total number of differentially expressed genes across light treatments after correcting for batch effects. The four genes shared between instar-III and instar-V contrasts were: *CG31548*, a putative oxidoreductase gene (DOWN-DOWN); *KGM_07406* (immune-related Hdd1 protein) (DOWN-DOWN); *CG9701*, a gene involved in carbohydrate metabolic processes (UP-UP); and *DDC*, an amine regulator (DOWN-DOWN), [DOWN = down-regulated; UP = up-regulated]. All four genes shared between instar-V and pupa contrasts lack annotation

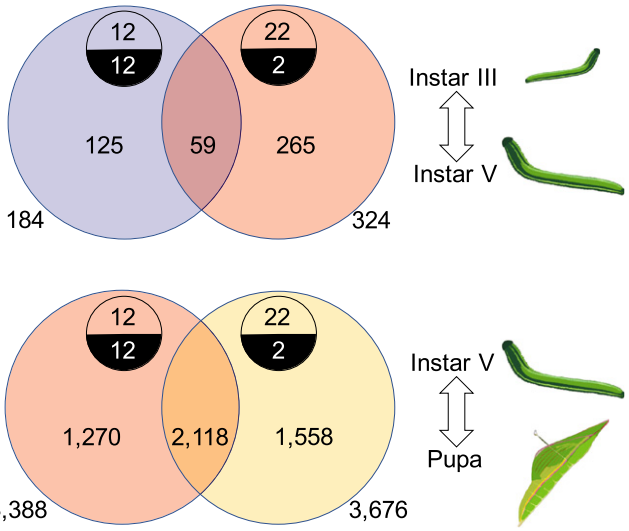

FIGURE 3 Venn diagrams showing numbers of differentially expressed genes across developmental stages after correcting for batch effects for the comparisons between instar-III and V (top) and between instar-V and pupa (bottom), for short-day and long-day light treatments

For instar-V samples, again multiple metabolic and biosynthetic processes that reflect developmental processes, such as molting, were overrepresented (Figure 4, Table S7). By this stage, specimens

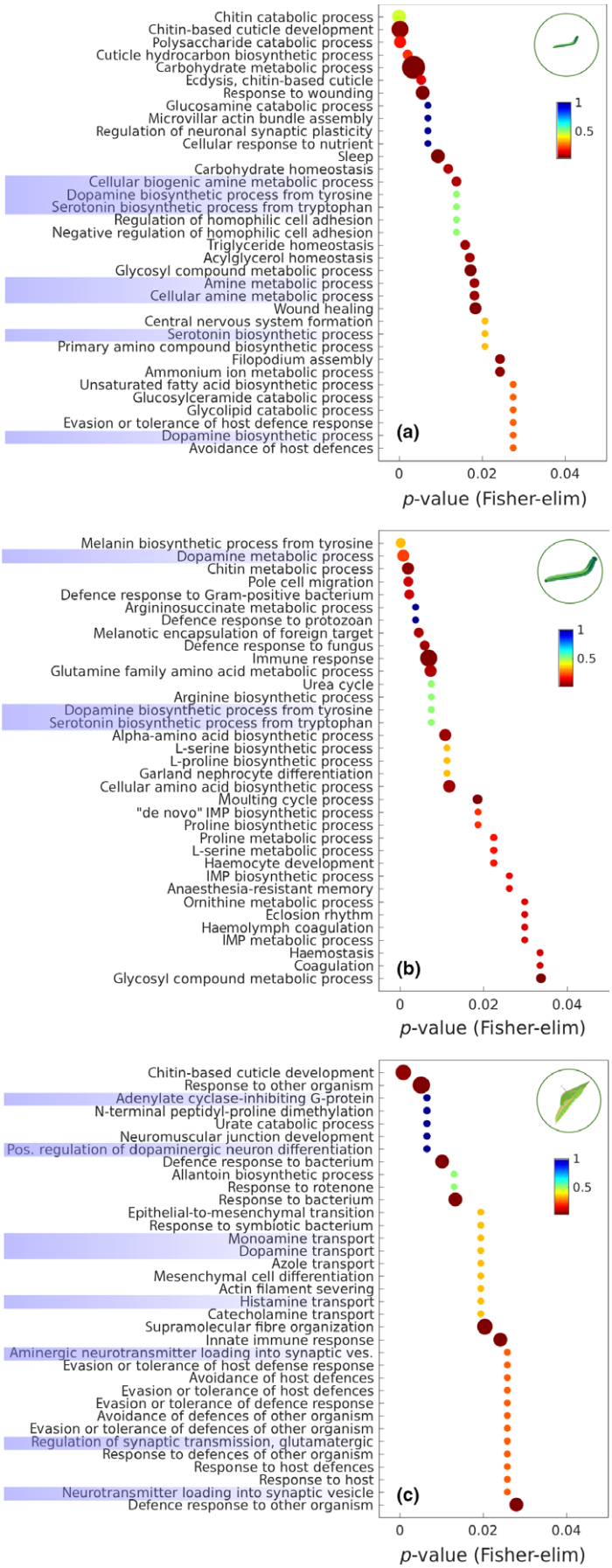

FIGURE 4 Top-35 most significant GO categories, for the set of differentially expressed genes across light treatments for instar-III (a), instar-V (b) and pupal (c) samples. Symbol size is proportional to the number of significantly expressed genes assigned to a GO category (significant genes). The smallest symbol size corresponds to one gene. Symbol color reflects the ratio between the number of significant genes associated to a particular GO term and the total number of annotated genes associated to that same GO category (enrichment score). Terms highlighted in blue are associated to processes involving monoamine neurotransmitters

should already have opted to go either into diapause or direct development, and activation of direct development pathways should accelerate sexual maturation. In line with this, one of the significant GO categories was 'pole cell migration' (Figure 4, Table S7). This category was underpinned by the gene *ZFH1* (Zink finger homeodomain 1), up-regulated in long-day instar-V larvae and known to be involved in gonadogenesis in *Drosophila* (Howard, 1998). Readjustment of the immune system has been observed in insects undergoing diapause (Lee, Horodyski, Valaitis, & Denlinger, 2002; Prasai & Karlsson, 2011; Ragland, Denlinger, & Hahn, 2010) or subject to food restriction (e.g. Adamo, Davies, Easy, Kovalko, & Turnbull, 2016), likely due to the close link between immune system pathways and pathways involved in metabolism and stress resistance regulation (Becker et al., 2010). Our results show that functional categories associated to immune system response were already being differentially regulated during instar-V. Dopamine and serotonin-related GO terms were again significantly overrepresented; *DDC* was asymmetrically expressed between light treatments (up-regulated in short-day conditions, like in instar-III). Another amine-regulator was up-regulated in the long-day treatment, *PPO2* (Prophenoloxidase 2), an enzyme that has been shown to act on dopamine substrates (Asada & Sezaki, 1999).

For the light-treatment contrast of pupal samples, most GO terms were linked to cytoskeleton development and organization, immune response and, once again, to biological processes involving monoamine neurotransmitters, including dopamine, histamine, and monoamine transport (Figure 4, Table S7), but the underlying genes were different from those involved at larval stages. For example, *VMAT* (vesicular monoamine transporter), a gene involved in dopamine and histamine storage and transport in *Drosophila* (Sang et al., 2007), and *CG32447*, a gene implicated in the G-protein coupled glutamate receptor signaling pathway (Brody & Cravchik, 2000) were both up-regulated in long-day light treatment pupae.

3.4 | Gene ontology enrichment analysis across developmental stages

Analysis of GO terms associated to the sets of genes differentially expressed across developmental stages revealed the presence of several genes linked to monoamine or hormonal regulation not detected in our previous analysis of light treatment contrasts (Fig. S8, Table S8). Most of these genes had expression profiles that did not differ significantly across light treatments. Besides the five genes flagged through direct comparison across light treatments (*DDC*, *ST*, *PPO2*, *VMAT*, and *CG32447*), three monoamine/hormonal-related genes were inferred to be modulated differently across light treatments: *CYP15C1* (farnesoate epoxidase), *MANF* (mesencephalic astrocyte-derived neurotrophic factor), and *HN* (henna) (Fig. S8). In short, *DDC* and *ST* were up-regulated in instar-III samples under short-day light treatment (Fig. S9 and S10). By instar-V, the gap in expression levels across treatments for these genes had been reduced substantially, although the difference remained statistically significant for *DDC*. Average expression levels for *CYP15C1* increased across developmental stages for individuals under short-day light treatment, while the opposite was observed in individuals in the direct development path (Fig. S9 and S10). During instar-V, specimens under long-day light treatment showed up-regulated expression levels for *PPO2* (Fig. S11), and reduced expression for *DDC* and *MANF* (Fig. S9, S10 and S11). Two genes (*VMAT* and *CG32447*) showed reduced expression levels during the pupal stage in short-day treatment while *HN* was up-regulated (Fig. S11 and S12).

3.5 | Expression patterns of circadian clock genes

Since previous studies have indicated that genes associated to the circadian clock may be implicated in diapause regulation, we specifically looked into the expression levels of six circadian clock related genes: *CLK* (clock), *CYC* (cycle), *PER* (period), *TIM* (timeless), *CRY1* (cryptochrome 1), and *CRY2* (cryptochrome 2). All six genes were generally lowly expressed (Fig. S13). Differentials between weakly expressed genes rarely attain statistical significance and are consequently difficult to investigate using DGE analysis. For instance, for instar-V samples, *TIM* was differentially expressed across light treatments at the unadjusted level (p-value = .002), but the FDR-adjusted p-value was = 0.15. There were, however, two genes whose expression showed a statistically significant difference across developmental transitions. *CYC* was down-regulated between instar-V and pupal stages under short-day light treatment (p-adj = .013), while *PER* was up-regulated (p-adj = .03) (Figs S13, S14). To conclude, expression differences of circadian clock genes in *Leptidea* were only supported across ontogenetic stages after the decision had been made on whether to enter diapause or direct development.

4 | DISCUSSION & CONCLUSIONS

4.1 | General

In insects, the diapause-stage (egg, larva, pupa or adult) is highly species-specific (Koštál, 2006), suggesting that diapause has evolved independently multiple times (e.g. Ragland et al., 2010). In addition, diapause is a flexible facultative trait that needs to be regulated over comparatively short time scales as species distribution ranges contract and expand in response to oscillations in environmental conditions. This indicates that tuning of regulatory pathways, rather than utilization of different functional variants, is the most efficient way to optimize reproductive output and avoid adverse conditions in regions with considerable seasonal climatic variation. A recent meta-analysis of transcriptomic data across multiple insect orders indicates that the transcriptomic response to diapause partly is shared, with expression profiles of genes involved in circadian rhythm, insulin and Wnt signaling revealing patterns that suggest these are similarly regulated across species (Ragland & Keep, 2017). However, there was no phylogenetic signal in the data and lineage-specific effects could have been underestimated as the analysis was restricted to 1:1 orthologs (≈4,800 genes) across the sample set and included taxa

that enter diapause at different stages (Ragland et al., 2010). A general conclusive support for specific pathways involved in diapause has hence not yet been established and it is likely that different molecular mechanisms are involved in initiation, maintenance and termination (Lehmann et al., 2018). In this study, we used gene expression profiling to investigate the gene expression differences between directly developing and diapausing lineages of the common wood-white, *Leptidea sinapis*. Our experimental set-up was designed to focus on the mechanisms involved in diapause initiation. The results suggest that gene expression cascades are differentially tuned under discrete light treatments already in instar-III larvae, resulting in different levels of regulatory molecules involved in the induction of diapause or direct development. At the stage when the developmental pathway had been irreversibly set (instar-V), a new set of cascades was activated depending on whether diapause or direct development had been induced, affecting both developmental pathways and metabolic processes differently in post-decision larvae and pupae. The general pattern emerging from the DGE and GO enrichment analysis suggests that light conditions induce changes in expression in three major gene categories as specimens develop through the different ontogenetic stages. During instar-III, a large cohort of genes involved in the regulation of carbohydrate and lipid metabolic processes was asymmetrically expressed across light treatments. Neuronal/hormonal pathways and chitin metabolism were likewise differently modulated, trends that to a large extent were also observed during instar-V and the pupal stage, although involving a different set of genes. Differences in expression for genes associated to gonad development and immune response became more pronounced during Instar-V, while in pupae reared under different light treatments there was a marked expression differential for genes involved in inter-cellular communication and organization.

4.2 | Gene expression differences across light treatments and ontogenetic stages

We found that *ST* and *DDC* were up-regulated in instar-III larvae under short-day treatment. Both *ST* and *DDC* have been shown to be involved in the regulation and recycling of monoamine neurotransmitters in arthropods (Borycz, Borycz, Kubów, Lloyd, & Meinertzhagen, 2008; Livingstone & Tempel, 1983; Stuart, Borycz, & Meinertzhagen, 2007). In *Drosophila*, *ST* is involved in histamine recycling in neuroglia (Stuart et al., 2007). Histamine is a monoamine neurotransmitter, released by photoreceptors in arthropods, with histamine levels modulated according to light intensity. *ST* affects not only histamine levels but, similar to *DDC*, expression of this gene is also associated with an increase in dopamine and serotonin levels in the brain of fruit flies (Borycz et al., 2008; Livingstone & Tempel, 1983), suggesting that average dopamine and serotonin concentrations were likely higher in wood-white larvae exposed to short-day conditions. Monoamine neurotransmitters play several important regulatory roles. In *Drosophila*, neural dopamine is required for entrainment of daily activity rhythms under low ambient light levels (Hirsh et al., 2010). Studies of larvae of the drosophilid fly *Chymomyza costata*, showed that serotonin and dopamine levels increased by up to 20% at the onset of darkness (Koštál, Noguchi, Shimada, & Hayakawa, 2000), while in *Bombyx mori* larvae, the amount of melatonin in the head increased up to 5-fold during nighttime (Itoh, Hattori, Nomura, Sumi, & Suzuki, 1995). Of particular relevance to the present study, dopamine and serotonin levels have been shown to diverge between diapausing and non-diapausing groups in some butterfly and moth species. In larvae of a fellow pierid butterfly, the large cabbage white (*Pieris brassicae*), the amount of dopamine by the end of the photosensitive period under short-day light treatment was twice as high as under long-day treatment, mostly because catabolism of dopamine had ceased to occur (Isabel et al., 2001). Dopamine levels were also higher in diapausing individuals during the pupal stage, and were maintained at high levels until the termination of diapause. In *B. mori*, dopamine concentrations are consistently higher in diapause-type larvae and pupae than in those set for direct development (Noguchi & Hayakawa, 2001). In the Chinese oak silkmoth (*Antheraea pernyi*), both dopamine and melatonin have been shown to be involved in diapause termination (Wang, Hanatani, Takeda, Oishi, & Sakamoto, 2015; Wang, Egi, et al., 2015). In *Drosophila*, low concentration of octopamine, a dopamine catabolite functionally equivalent to the vertebrate neurotransmitter norepinephrine (Roeder, 2004), has been linked with increased body fat, higher starvation resistance, reduced glucose release (energy expenditures), and a general slowdown of metabolic processes (Li et al., 2016). High dopamine concentration and a sharp reduction on the rate of dopamine metabolism have been associated to the onset of diapause in some insects (Houk & Beck, 1977; Isabel et al., 2001). In addition, in *P. brassicae*, high serotonin levels have been proposed to lead to progressive inhibition of cerebral metabolism and endocrinological activity (Isabel et al., 2001).

Our results also show that several metabolic processes are modulated asymmetrically between light treatments during instar-III. The *SUG* gene, up-regulated in pre-diapausing individuals under short-day treatment, has been shown to inhibit fat catabolism and promote fatty acid synthesis (Zinke et al., 2002). *AMY-P*, a gene involved in carbohydrate metabolism (Powell & Andjelković, 1983), was down-regulated in short-day instar-III samples. The *DESAT1* gene, up-regulated in larvae subject to short-day light conditions, codes for an enzyme that promotes lipid metabolic processes favoring fat storage (Parisi et al., 2013). Finally, our data also suggests that expression of *CYP15C1* was reduced in larvae subjected to short-day light treatment. *CYP15C1* has been shown to be required for the production of the juvenile hormone (*JH*) in *Bombyx mori* (Daimon et al., 2012). Juvenile hormones, in turn, are known to play a critical role in the regulation of larval molting during late instars, in signaling the onset of metamorphosis, and in the regulation of reproductive maturation (Jindra, Palli, & Riddiford, 2013; Riddiford, 2012). In insects undergoing diapause as adults, fat synthesis and accumulation occur only if juvenile hormone levels are kept low (Liu, Li, Zhu, Lei, & Wang, 2016; Sim & Denlinger, 2008). Taken together, these results suggest that metabolic activities in instar-III larvae under short-day treatment

are being redirected towards fatty acid synthesis and sequestration, a hallmark of pre-diapausing behavior (Arrese & Soulages, 2009), while glucose usage is being curtailed.

After molt into instar-V, developmental paths are most likely irreversible (Friberg et al., 2011). At this stage, monoamine regulation continued to differ significantly between treatments, although including a different group of genes. While *DDC* was still up-regulated under short day treatment, *ST* expression was similar for both light conditions. Instead, *PPO2*, a gene weakly expressed during instar-III under both light treatments, was now up-regulated in long-day larvae. *PPO2* is a multifunctional enzyme used by arthropods in melanin synthesis and sclerotization, a process during which dopamine is used as a precursor (Sugumaran, 2002). *PPO2* was up-regulated in larvae under long-day light treatment, suggesting a higher metabolic rate, as well as a higher catabolic rate of dopamine. The expression level of *MANF*, a gene involved in positive dopamine regulation (Lindholm & Saarma, 2010), dropped between instar-III and instar-V in samples under long-day treatment. All three genes (*DDC*, *PPO2* and *MANF*) are directly or indirectly involved in monoamine regulation, which indicates that high dopamine levels are maintained in instar-V larvae kept under diapause inducing conditions. In addition, indicative of more specific pathways triggered in the different developmental paths, expression levels for genes involved in gonad development and immune system regulation diverged between treatments.

In pupae, a new set of differentially expressed monoamine regulators were identified. *DDC* and *PPO2* were only weakly expressed, and differences in expression across light treatments for *MANF* were not meaningful. Instead, significant expression differences were observed for *VMAT*, *CG32447*, and *HN*. *VMAT*, a monoamine transporter known to decrease overall dopamine levels when overexpressed (Sang et al., 2007), was up-regulated in pupae under long-day light treatment. *CG32447* is part of the metabotropic G-protein coupled glutamate receptor family and acts as a dopamine modulator (Brody & Cravchik, 2000). It was down-regulated in pupae reared under short-day conditions. In *Drosophila*, *HN* is an enzyme homologous to the human *PAH* (Phenylalanine-4-hydroxylase) and has been suggested to be part of the metabolic pathway involved in the synthesis of catecholamine neurotransmitters such as dopamine (Craig, Buckle, Lamouroux, Mallet, & Craig, 1988). *HN* was only weakly expressed in long-day light treatment pupae.

5 | CONCLUSIONS

These results suggest that the process leading to diapause induction starts early during development, likely at the beginning of the photosensitive period. In other words, the switch to diapause is the culmination of a process that builds up slowly in larvae under short-day conditions. Under long-day light treatment, these diapause-inducing conditions never take a hold and the (ancestral) direct development pathway remains activated. Short-day conditions trigger an increase in dopamine and other monoamines neurotransmitters while, under long-day conditions, monoamine levels remain low after each short night (Itoh et al., 1995). Hence, we hypothesize that sustained high dopamine levels instigate metabolic changes in larvae reared under short-day conditions that favor fatty acid storage and reduce glucose usage. The progressive slowdown of general metabolic activities reaches its end point by the end of the photosensitive period, which can be seen either as a decision point (one or more biochemical reactions mark the switch to diapause) or the end of a window of opportunity. This is consistent with experiments carried out in the green-veined white (*Pieris napi*), the speckled wood (*Pararge aegeria*) and the map (*Araschnia levana*) butterflies, where developmental pathway decisions were observed to be made in late instar larvae in an asymmetrical fashion – inducing a switch from diapause to direct development in late instar larvae was possible, but not the other way around (Friberg et al., 2011). From our results, *DDC* and *ST* emerge as the key genes whose expression promotes a sustained high level of dopamine during the duration of the photosensitive period. Expression of monoamine modulators active during the last instar and pupation, such as *PPO2*, *MANF*, *VMAT*, and *CG32447*, might be subservient to the determination of a specific developmental pathway triggered earlier.

5.1 | Outlook and some caveats

The notable variability in expression profiles associated with family in our data demonstrates that there is natural variability with a genetic basis upon which natural selection can act. This agrees well with the differences in diapause induction documented between natural *Leptidea* populations at different latitudes (Friberg & Wiklund, unpublished data), and the potential necessity of fine-tuning as a part of local adaptation. This raises the interesting hypothesis that diapause regulation has diverged among different common wood-white butterfly ecotypes, and that these differences might be observable at the genomic level, for example by allele frequency shifts in regulatory regions of genes involved in diapause induction. To get more detailed information about the general effects of photoperiod length and genes and pathways involved in diapause regulation, a follow-up study would ideally include populations with different critical day-length thresholds for diapause induction. Furthermore, crossing experiments with individuals from different *Leptidea* populations with distinct life-history strategies could also provide a framework to map the genetic basis of diapause induction. Having access to a handful of candidate genes related to diapause regulation also opens the door to in vitro experiments for functional verification. An interesting aspect is that Swedish *L. sinapis* diapause propensity also partly depends on host plant. When reared on *Lathyrus linifolius*, a significantly higher proportion of individuals enter diapause development than when reared on *Lathyrus pratensis* or *Lotus corniculatus*, even when specimens are subjected to long-day light treatment (Friberg & Wiklund, 2010). Hence, a similar study with host plant treatments could reveal if food source triggers similar cascades of monoamine regulators or if this triggers a different response.

It should be noted that this study was based on RNA collected from the whole body of specimens. If some of the key genes and pathways directly or indirectly involved in diapause regulation are confined to a specific organ or tissue, for example brain or gut (Bajgar et al., 2013; Goto & Denlinger, 2002; Linn, Poole, Roelofs, & Wu, 1995), the associated transcriptional signal could have been diluted, or even lost. In *Drosophila melanogaster* for example, differentially expressed genes between diapausing and non-diapausing female lines are highly distinct for head and ovary (Zhao et al., 2016). As shown in the results, circadian clock genes were usually weakly expressed and, as befitting a group of genes responsible for regulating circadian cellular activities, their expression levels are expected to vary throughout the day following a 24-h cycle (Hardin et al., 1990; Meuti & Denlinger, 2013; Young & Kay, 2001). This makes detection and interpretation of expression differentials rather problematic. Synchronization between light treatments could be done by enforcing simultaneous lights-on at specific times. This would ensure that the expression of circadian clock genes would be in phase at the beginning of each cycle. A denser temporal sampling procedure would also be necessary to properly quantify variation in circadian gene expression patterns between different day-length conditions. As mentioned above, this was not the set-up in this study where treatments were synchronized at lights-off and we can therefore not explicitly rule out that circadian clock genes may affect diapause induction. For example, in *Leptidea*, circadian regulators may be active at an early stage triggering downstream processes of monoamine regulation that fit the observed hour-glass model.

ACKNOWLEDGEMENTS

This work was supported by a Junior Research Grant from the Swedish Research Council [VR 2013-4508] to N.B. We acknowledge additional funding for library preparation and sequencing from Science for Life Laboratory (SciLife Sweden) Biodiversity Program. The SNP&SEQ Technology Platform and Uppsala Genome Center performed the library preparations and the sequencing supported by Science for Life Laboratory (SciLife, Stockholm); a national infrastructure funded by the Swedish Research Council (VR-RFI) and the Knut and Alice Wallenberg Foundation. The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (Lampa, Dahlö, Olason, Hagberg, & Spjuth, 2013) under Project # b2014034. We thank Berrit Kiehl for help with RNA extractions and the journal editor and three anonymous reviewers for insightful and constructive comments that helped improving the manuscript.

DATA ACCESSIBILITY

All raw sequence reads from this study have been uploaded to the European Bioinformatics Institute (EMBL-EBI) data-base ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) under accession number E-MTAB-6454. In house developed scripts and pipelines are available at: https://github.com/LLN273/RNAseq_Lsinapis (see also Supplementary Methods).

AUTHOR CONTRIBUTIONS

V.T. and N.B. conceived of the study, sampled the specimens, performed the experiments with assistance and advice from M.F., C.W., V.D. and R.V., and did the molecular work. L.L. performed all data analyses with advice from T.K., V.T. and N.B. L.L. and N.B. wrote the manuscript with input from all co-authors. The final version of the manuscript was approved by all co-authors.

ORCID

Niclas Backström http://orcid.org/0000-0002-0961-8427

REFERENCES

Aalberg Haugen, I. M., & Gotthard, K. (2015). Diapause induction and relaxed selection on alternative developmental pathways in a butterfly. *Journal of Animal Ecology*, *84*, 464–472. https://doi.org/10.1111/1365-2656.12291

Adamo, S. A., Davies, G., Easy, R., Kovalko, I., & Turnbull, K. F. (2016). Reconfiguration of the immune system network during food limitation in the caterpillar *Manduca sexta*. *Journal of Experimental Biology*, *219*, 706–718. https://doi.org/doi: 10.1242/jeb.132936

Alexa, A., & Rahnenfuhrer, J. (2016). topGO: enrichment analysis for gene ontology.

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

Andrés, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics*, *13*, 627–639. https://doi.org/10.1038/nrg3291

Arrese, E. L., & Soulages, J. L. (2009). Insect fat body: Energy, metabolism, and regulation. *Annual Review of Entomology*, *55*, 207–225. https://doi.org/10.1146/annurev-ento-112408-085356

Asada, N., & Sezaki, H. (1999). Properties of phenoloxidases generated from prophenoloxidase with 2-propanol and the natural activator in *Drosophila melanogaster*. *Biochemical Genetics*, *37*, 149–158. https://doi.org/10.1023/A:1018730404305

Bajgar, A., Jindra, M., & Dolezel, D. (2013). Autonomous regulation of the insect gut by circadian genes acting downstream of juvenile hormone signaling. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 4416–4421. https://doi.org/10.1073/pnas.1217060110

Beck, S. D. (1980). *Insect photoperiodism*, 2nd ed.. New York: Academic Press.

Becker, T., Loch, G., Beyer, M., Zinke, I., Aschenbrenner, A. C., Carrera, P., … Hoch, M. (2010). FOXO-dependent regulation of innate immune homeostasis. *Nature*, *463*, 369. https://doi.org/doi: 10.1038/nature08698

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, *57*, 289–300. https://doi.org/10.2307/2346101

Borycz, J., Borycz, J. A., Kubów, A., Lloyd, V., & Meinertzhagen, I. A. (2008). *Drosophila* ABC transporter mutants white, brown and scarlet have altered contents and distribution of biogenic amines in the

brain. *Journal of Experimental Biology*, *211*, 3454–3466. https://doi.org/10.1242/jeb.021162

Bradshaw, W. E., & Holzapfel, C. M. (2001). Genetic shift in photoperiodic response correlated with global warming. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 14509–14511. https://doi.org/10.1073/pnas.241391498

Bradshaw, W. E., & Holzapfel, C. M. (2010a). Circadian clock genes, ovarian development and diapause. *BMC Biology*, *8*, 115. https://doi.org/10.1186/1741-7007-8-115

Bradshaw, W. E., & Holzapfel, C. M. (2010b). What season is it anyway? Circadian tracking vs. photoperiodic anticipation in insects. *Journal of Biological Rhythms*, *25*, 155–165. https://doi.org/10.1177/0748730410365656

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*, 525–527. https://doi.org/10.1038/nbt.3519

Brody, T., & Cravchik, A. (2000). *Drosophila melanogaster* G-protein–coupled receptors. *Journal of Cell Biology*, *150*, F83–F88. https://doi.org/10.1083/jcb.150.2.F83

Bünning, E. (1936). Die endogene tagesrhythmik als grundlage der photoperiodischen reaktion. *Berichte der Deutschen Botanischen Geschellschaft*, *53*, 590–607.

Byers, K. J., Xu, S., & Schluter, P. M. (2017). Molecular mechanisms of adaptation and speciation: Why do we need an integrative approach? *Molecular Ecology*, *26*, 277–290. https://doi.org/10.1111/mec.13678

Challis, R. J., Kumar, S., Dasmahapatra, K. K., Jiggins, C. D., & Blaxter, M. (2017). Lepbase - the lepidopteran genome database. *BioRxiv*, 056994. https://doi.org/10.1101/056994

Chomczynski, P., & Mackey, K. (1995). Modification of the TRIZOL reagent procedure for isolation of RNA from polysaccharide-and proteoglycan-rich sources. *BioTechniques*, *19*, 942–945.

Chomczynski, P., & Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, *162*, 156–159. https://doi.org/10.1006/abio.1987.9999

Cong, Q., Shen, J., Warren, A. D., Borek, D., Otwinowski, Z., & Grishin, N. V. (2016). Speciation in cloudless sulphurs gleaned from complete genomes. *Genome Biology and Evolution*, *8*, 915–931. https://doi.org/10.1093/gbe/evw045

Craig, S. P., Buckle, V. J., Lamouroux, A., Mallet, J., & Craig, I. W. (1988). Localization of the human dopamine beta hydroxylase (*DBH*) gene to chromosome 9q34. *Cytogenetic and Genome Research*, *48*, 48–50 https://www.ncbi.nlm.nih.gov/pubmed/3180847.

Daimon, T., Kozaki, T., Niwa, R., Kobayashi, I., Furuta, K., Namiki, T., … Shinoda, T. (2012). Precocious metamorphosis in the juvenile hormone–deficient mutant of the silkworm, *Bombyx mori*. *PLoS Genetics*, *8*, e1002486. https://doi.org/10.1371/journal.pgen.1002486

Denlinger, D. L. (2002). Regulation of diapause. *Annual Review of Entomology*, *47*, 93–122. https://doi.org/10.1146/annurev.ento.47.091201.145137

Denlinger, D. L., & Armbruster, P. A. (2014). Mosquito diapause. *Annual Review of Entomology*, *59*, 73–93. https://doi.org/10.1146/annurev-ento-011613-162023

Denlinger, D. L., Hahn, D. A., Merlin, C., Holzapfel, C. M., & Bradshaw, W. E. (2017). Keeping time without a spine: What can the insect clock teach us about seasonal adaptation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*, e20160257. https://doi.org/doi: 10.1098/rstb.2016.0257

Denlinger, D. L., Yocum, G. D., & Rinehart, J. P. (2012). Hormonal control of diapause. In L. I. Gilbert (Ed.), *Insect endocrinology* (pp. 430–463). San Diego, CA, USA: Academic Press. https://doi.org/10.1016/B978-0-12-384749-2.10010-X

Ellegren, H., & Sheldon, B. C. (2008). Genetic basis of fitness differences in natural populations. *Nature*, *452*, 169–175. https://doi.org/10.1038/nature06737

Emerson, K. J., Bradshaw, W. E., & Holzapfel, C. M. (2009). Complications of complexity: Integrating environmental, genetic and hormonal control of insect diapause. *Trends in Genetics*, *25*, 217–225. https://doi.org/10.1016/j.tig.2009.03.009

Friberg, M., Aalberg Haugen, I. M., Dahlerus, J., Gotthard, K., & Wiklund, C. (2011). Asymmetric life-history decision-making in butterfly larvae. *Oecologia*, *165*, 301–310. https://doi.org/10.1007/s00442-010-1804-0

Friberg, M., Bergman, M., Kullberg, J., Wahlberg, N., & Wiklund, C. (2008). Niche separation in space and time between two sympatric sister species - a case of ecological pleiotropy. *Evolutionary Ecology*, *22*, 1–18. https://doi.org/10.1007/s10682-007-9155-y

Friberg, M., & Wiklund, C. (2010). Host-plant-induced larval decision-making in a habitat/host-plant generalist butterfly. *Ecology*, *91*, 15–21. https://www.ncbi.nlm.nih.gov/pubmed/20380190.

Goto, S. G., & Denlinger, D. L. (2002). Short-day and long-day expression patterns of genes involved in the flesh fly clock mechanism: Period, timeless, cycle and cryptochrome. *Journal of Insect Physiology*, *48*, 803–816. https://doi.org/10.1016/S0022-1910(02)00108-7

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652. https://doi.org/10.1038/nbt.1883

Haas, B. (2016). *Transcriptome contig Nx and ExN50 stats*. Retrieved from https://github.com/trinityrnaseq/trinityrnaseq/wiki/Transcriptome Contig Nx and ExN50 stats. Accessed 2017-08-08.

Haas, B. (2017). *Trinity documentation: Trinity FAQ*. Retrieved from https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-FAQ. Accessed 2017-08-08

Hahn, D. A., & Denlinger, D. L. (2011). Energetics of insect diapause. *Annual Review of Entomology*, *56*, 103–121. https://doi.org/10.1146/annurev-ento-112408-085436

Hardin, P. E., Hall, J. C., & Rosbash, M. (1990). Feedback of the *Drosophila Period* gene product on circadian cycling of its messenger RNA levels. *Nature*, *343*, 536–540. https://doi.org/10.1038/343536a0

Hirsh, J., Riemensperger, T., Coulom, H., Iché, M., Coupar, J., & Birman, S. (2010). Roles of dopamine in circadian rhythmicity and extreme light sensitivity of circadian entrainment. *Current Biology*, *20*, 209–214. https://doi.org/10.1016/j.cub.2009.11.037

Houk, E. J., & Beck, S. D. (1977). Distribution of putative neurotransmitters in the brain of the European corn borer, *Ostrinia nubilalis*. *Journal of Insect Physiology*, *23*, 1209–1217. https://doi.org/10.1016/0022-1910(77)90155-X

Howard, K. (1998). Organogenesis: *Drosophila* goes gonadal. *Current Biology*, *8*, R415–R417. https://doi.org/10.1016/S0960-9822(98)70266-0

Isabel, G., Gourdoux, I. G. L., & Moreau, R. (2001). Changes of biogenic amine levels in haemolymph during diapausing and non-diapausing status in *Pieris brassicae* L. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *128*, 117–127. https://doi.org/10.1016/S1095-6433(00)00284-1

Itoh, M., Hattori, A., Nomura, T., Sumi, Y., & Suzuki, T. (1995). Melatonin and arylalkylamine N-acetyltransferase activity in the silkworm, *Bombyx mori*. *Molecular and Cellular Endocrinology*, *115*, 59–64. https://doi.org/10.1016/0303-7207(95)03670-3

Jindra, M., Palli, S. R., & Riddiford, L. M. (2013). The juvenile hormone signaling pathway in insect development. *Annual Review of Entomology*, *58*, 181–204. https://doi.org/10.1146/annurev-ento-120811-153700

Kivelä, S. M., Svensson, B., Tiwe, A., & Gotthard, K. (2015). Thermal plasticity of growth and development varies adaptively among alternative developmental pathways. *Evolution*, *69*, 2399–2413. https://doi.org/10.1111/evo.12734

Koštál, V. (2006). Eco-physiological phases of insect diapause. *Journal of Insect Physiology*, *52*, 113–127. https://doi.org/doi: 10.1016/j.jinsphys.2005.09.008

Koštál, V. (2011). Insect photoperiodic calendar and circadian clock: Independence, cooperation, or unity? *Journal of Insect Physiology*, *57*, 538–556. https://doi.org/10.1016/j.jinsphys.2010.10.006

Koštál, V., Noguchi, H., Shimada, K., & Hayakawa, Y. (2000). Circadian component influences the photoperiodic induction of diapause in a drosophilid fly, *Chymomyza costata*. *Journal of Insect Physiology*, *46*, 887–896. https://doi.org/10.1016/S0022-1910(99)00195-X

Lampa, S., Dahlö, M., Olason, P. I., Hagberg, J., & Spjuth, O. (2013). Lessons learned from implementing a national infrastructure in Sweden for storage and analysis of next-generation sequencing data. *Gigascience*, *25*, 9. https://doi.org/10.1186/2047-217X-2-9

Lee, K. Y., Horodyski, F. M., Valaitis, A. P., & Denlinger, D. L. (2002). Molecular characterization of the insect immune protein hemolin and its high induction during embryonic diapause in the gypsy moth, *Lymantria dispar*. *Insect Biochemistry and Molecular Biology*, *32*, 1457–1467. https://doi.org/10.1016/S0965-1748(02)00066-8

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, *28*, 882–883. https://doi.org/10.1093/bioinformatics/bts034

Lees, A. D. (1953). The significance of the light and dark phases in the photoperiodic control of diapause in *Metatetranychus ulmi-koch*. *Annals of Applied Biology*, *40*, 487–497. https://doi.org/10.1111/j.1744-7348.1953.tb02388.x

Lees, A. D. (1973). Photoperiodic time measurement in the aphid *Megoura viciae*. *Journal of Insect Physiology*, *19*, 2279–2316. https://doi.org/10.1016/0022-1910(73)90237-0

Lehmann, P., Pruisscher, P., Koštál, V., Moos, M., Simek, P., Nylin, S., & Agren, R. (2018). Metabolome dynamics of diapause in the butterfly *Pieris napi*: Distinguishing maintenance, termination and post-diapause phases. *Journal of Experimental Biology Early Online Version*, https://doi.org/10.1098/rstb.2016.0257

Li, Y., Hoffmann, J., Li, Y., Stephano, F., Bruchhaus, I., Fink, C., & Roeder, T. (2016). Octopamine controls starvation resistance, life span and metabolic traits in *Drosophila*. *Scientific Reports*, *6*, srep35359. https://doi.org/10.1038/srep35359

Lindholm, P., & Saarma, M. (2010). Novel *CDNF/MANF* family of neurotrophic factors. *Developmental Neurobiology*, *70*, 360–371. https://doi.org/10.1002/dneu.20760

Linn, C. E., Poole, K. R., Roelofs, W. L., & Wu, W.-Q. (1995). Circadian changes in melatonin in the nervous system and hemolymph of the cabbage looper moth, *Trichoplusia* N. *Journal of Comparative Physiology A*, *176*, 761–771. https://doi.org/10.1007/BF00192624

Liu, W., Li, Y. L., Zhu, F., Lei, C.-L., & Wang, X.-P. (2016). Juvenile hormone facilitates the antagonism between adult reproduction and diapause through the methoprene-tolerant gene in the female *Colaphellus bowringi*. *Insect Biochemistry and Molecular Biology*, *74*, 50–60. https://doi.org/10.1016/j.ibmb.2016.05.004

Livingstone, M. S., & Tempel, B. L. (1983). Genetic dissection of monoamine neurotransmitter synthesis in *Drosophila*. *Nature*, *303*, 67–70. https://doi.org/10.1038/303067a0

Love, M. I., Anders, S., Kim, V., & Huber, W. (2016). RNA-seq workflow: Gene-level exploratory analysis and differential expression. *F1000 Research*, *4*, 1070 https://doi.org/10.12688/f1000research.7035.1.

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

Merzendorfer, H., & Zimoch, L. (2003). Chitin metabolism in insects: Structure, function and regulation of chitin synthases and chitinases. *Journal of Experimental Biology*, *206*, 4393–4412. https://doi.org/10.1242/jeb.00709

Meuti, M. E., & Denlinger, D. L. (2013). Evolutionary links between circadian clocks and photoperiodic diapause in insects. *Integrative and Comparative Biology*, *53*, 131–143. https://doi.org/10.1093/icb/ict023

Miller, J. R., Koren, S., & Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. *Genomics*, *95*, 315–327. https://doi.org/10.1016/j.ygeno.2010.03.001

Nadeau, N. J., & Jiggins, C. D. (2010). A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics*, *26*, 484–492. https://doi.org/10.1016/j.tig.2010.08.004

Noguchi, H., & Hayakawa, Y. (2001). Dopamine is a key factor for the induction of egg diapause of the silkworm, *Bombyx mori*. *European Journal of Biochemistry*, *268*, 744–780. https://doi.org/10.1046/j.1432-1327.2001.01933.x

Nowell, R. W., Elsworth, B., Oostra, V., Zwaan, B. J., Wheat, C. W., Saastamoinen, M., … Blaxter, M. (2017). A high-coverage draft genome of the mycalesine butterfly *Bicyclus anynana*. *Gigascience*, *6*, 1–7. https://doi.org/10.1093/gigascience/gix035

Pardo-Diaz, C., Salazar, C., & Jiggins, C. D. (2015). Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution*, *6*, 445–464. https://doi.org/10.1111/2041-210X.12324

Parisi, F., Riccardo, S., Zola, S., Lora, C., Grifoni, D., Brown, L. M., & Bellosta, P. (2013). DMyc expression in the fat body affects DILP2 release and increases the expression of the fat desaturase *Desat1* resulting in organismal growth. *Developmental Biology*, *379*, 64–75. https://doi.org/10.1016/j.ydbio.2013.04.008

Pavey, S. A., Collin, H., Nosil, P., & Rogers, S. M. (2010). The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences*, *1206*, 110–129. https://doi.org/10.1111/j.1749-6632.2010.05765.x

Powell, J. R., & Andjelković, M. (1983). Population genetics of *Drosophila* amylase IV. Selection in laboratory populations maintained on different carbohydrates. *Genetics*, *103*, 675–689. https://www.ncbi.nlm.nih.gov/pubmed/6189764

Prasai, K., & Karlsson, B. (2011). Variation in immune defence in relation to developmental pathway in the green-veined white butterfly, *Pieris napi*. *Evolutionary Ecology Research*, *13*, 295–305.

Ragland, G. J., Denlinger, D. L., & Hahn, D. A. (2010). Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14909–14914. https://doi.org/10.1073/pnas.1007075107

Ragland, G. J., & Keep, E. (2017). Comparative transcriptomics support evolutionary convergence of diapause responses across Insecta. *Physiological Entomology*, *42*, 246–256. https://doi.org/10.1111/phen.12193

Riddiford, L. M. (2012). How does juvenile hormone control insect metamorphosis and reproduction? *General and Comparative Endocrinology*, *179*, 477–484. https://doi.org/10.1016/j.ygcen.2012.06.001

Roeder, T. (2004). Tyramine and octopamine: Ruling behavior and metabolism. *Annual Review of Entomology*, *50*, 447–477. https://doi.org/10.1146/annurev.ento.50.071803.130404

Sang, T.-K., Chang, H.-Y., Lawless, G. M., Ratnaparkhi, A., Mee, L., Ackerson, L. C., … Jackson, G. R. (2007). A *Drosophila* model of mutant human parkin-induced toxicity demonstrates selective loss of dopaminergic neurons and dependence on cellular dopamine. *Journal of Neuroscience*, *27*, 981–992. https://doi.org/10.1523/JNEUROSCI.4810-06.2007

Saunders, D. S. (2002). *Insect clocks*, 3rd ed.. Amsterdam, NL: Elsevier.

Saunders, D. S. (2009). Circadian rhythms and the evolution of photoperiodic timing in insects. *Physiological Entomology*, *34*, 301–308. https://doi.org/10.1111/j.1365-3032.2009.00699.x

Saunders, D. S. (2012). Insect photoperiodism: Seeing the light. *Physiological Entomology*, *37*, 207–218. https://doi.org/10.1111/j.1365-3032.2012.00837.x

Saunders, D. S., & Bertossa, R. C. (2011). Deciphering time measurement: The role of circadian 'clock' genes and formal experimentation in

insect photoperiodism. *Journal of Insect Physiology*, *57*, 557–566. https://doi.org/10.1016/j.jinsphys.2011.01.013

Saunders, D. S., Lewis, R. D., & Warman, G. R. (2004). Photoperiodic induction of diapause: Opening the black box. *Physiological Entomology*, *29*, 1–15. https://doi.org/10.1111/j.1365-3032.2004.0369.x

Schmidt, P. S., Zhu, C.-T., Das, J., Batavia, M., Yang, L., & Eanes, W. F. (2008). An amino acid polymorphism in the couch potato gene forms the basis for climatic adaptation in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 16207–16211. https://doi.org/10.1073/pnas.0805485105

Seppänen, E. (1969). Suurperhostemme talvehtimisasteet/Die überwinterungstadien unserer Grosschmetterlinge. *Annales Entomologici Fennici*, *35*, 129–152.

Shen, J., Cong, Q., Kinch, L. N., Borek, D., Otwinowski, Z., & Grishin, N. V. (2016). Complete genome of *Pieris rapae*, a resilient alien, a cabbage pest, and a source of anti-cancer proteins. *F1000 Research*, *5*, 2631. https://doi.org/10.12688/f1000research.9765.1

Sim, C., & Denlinger, D. L. (2008). Insulin signaling and FOXO regulate the overwintering diapause of the mosquito *Culex pipiens*. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6777–6781. https://doi.org/10.1073/pnas.0802067105

Soneson, C., Love, M. I., & Robinson, M. D. (2015). Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000 Research*, *4*, 1521. https://doi.org/10.12688/f1000research.7563.1

Stuart, A. E., Borycz, J., & Meinertzhagen, I. A. (2007). The dynamics of signaling at the histaminergic photoreceptor synapse of arthropods. *Progress in Neurobiology*, *82*, 202–227. https://doi.org/10.1016/j.pneurobio.2007.03.006

Sugumaran, M. (2002). Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. *Pigment Cell Research*, *15*, 2–9. https://doi.org/10.1034/j.1600-0749.2002.00056.x

Talla, V., Suh, A., Kalsoom, F., Dincă, V., Vila, R., Friberg, M., … Backström, N. (2017). Rapid increase in genome size as a consequence of transposable element hyperactivity in wood-white (*Leptidea*) butterflies. *Genome Biology and Evolution*, *9*, 2491–2505. https://doi.org/doi: 10.1093/gbe/evx163

The Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*, 94–98. https://doi.org/10.1038/nature11041

Tolman, T. (1997). *Butterflies of Britain and Europe*. London, UK: Harper Collins.

Van Dyck, H., Bonte, D., Puls, R., Gotthard, K., & Maes, D. (2014). The lost generation hypothesis: Could climate change drive ectotherms into a developmental trap? *Oikos*, *124*, 54–61. https://doi.org/10.1111/oik.02066

Veerman, A. (2001). Photoperiodic time measurement in insects and mites: A critical evaluation of the oscillator-clock hypothesis. *Journal of Insect Physiology*, *47*, 1097–1109. https://doi.org/10.1016/S0022-1910(01)00106-8

Wang, Q., Egi, Y., Takeda, M., Oishi, K., & Sakamoto, K. (2015). Melatonin pathway transmits information to terminate pupal diapause in the Chinese oak silkmoth *Antheraea pernyi* and through reciprocated inhibition of dopamine pathway functions as a photoperiodic counter. *Entomological Science*, *18*, 74–84. https://doi.org/10.1111/ens.12083

Wang, Q., Hanatani, I., Takeda, M., Oishi, K., & Sakamoto, K. (2015). D2-like dopamine receptors mediate regulation of pupal diapause in Chinese oak silkmoth *Antheraea pernyi*. *Entomological Science*, *18*, 193–198. https://doi.org/10.1111/ens.12099

Williams, K. D., Busto, M., Suster, M. L., Anthony, S. K. C., Ben-Shahar, Y., Leevers, S. J., & Sokolowski, M. B. (2006). Natural variation in *Drosophila melanogaster* diapause due to the insulin-regulated PI3-kinase. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 15911–15915. https://doi.org/10.1073/pnas.0604592103

Xu, W. H., Lu, Y. X., & Denlinger, D. L. (2012). Cross-talk between the fat body and brain regulates insect developmental arrest. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 14687–14692. https://doi.org/10.1073/pnas.1212879109

Young, M. W., & Kay, S. A. (2001). Time zones: A comparative genetics of circadian clocks. *Nature Reviews Genetics*, *2*, 702–715. https://doi.org/10.1038/35088576

Zhao, X., Bergland, A. O., Behrman, E. L., Gregory, B. D., Petrov, D. A., & Schmidt, P. S. (2016). Global transcriptional profiling of diapause and climatic adaptation in *Drosophila melanogaster*. *Molecular Biology and Evolution*, *33*, 707–720. https://doi.org/10.1093/molbev/msv263

Zinke, I., Schütz, C. S., Katzenberger, J. D., Bauer, M., & Pankratz, M. J. (2002). Nutrient control of gene expression in Drosophila: Microarray analysis of starvation and sugar-dependent response. *EMBO Journal*, *21*, 6162–6173. https://doi.org/10.1093/emboj/cdf600

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Leal L, Talla V, Källman T, et al. Gene expression profiling across ontogenetic stages in the wood white (*Leptidea sinapis*) reveals pathways linked to butterfly diapause regulation. *Mol Ecol*. 2018;00:1–14. https://doi.org/10.1111/mec.14501