Transcriptomic analysis reveals pathways underlying environmental stress response in a migratory butterfly

Shipilina D, (Hook L, Talla V, Parkes E, Nasvall K, Palahi Torres A, JB?) order TBD , Villa R (?), Talavera G, Backstrom N

Abstract

Migration is a complex behavior that involves the synchronization of a multitude of physiological and behavioral processes in response to environmental cues. In this study, we used the Painted Lady butterfly (*Vanessa cardui*) as a model and examined the gene expression profiles in response to stressors introduced both early in development (larval crowding and lack of food resources) and at the adult stage (host plant absence).
Our examination of the adult response to host plant availability revealed significant modifications in gene regulation, particularly in the expression of ecdysone esterase and juvenile hormone. We hypothesized that the ecdysone pathway may play a crucial role in the rapid switch between dispersal and reproduction. In addition, our functional analysis of candidate genes revealed significant enrichment in lipid, carbohydrate, and vitamin biosynthesis pathways, as well as the metabolism of pathogens itself and expression of multiple immune genes.
Our findings showed that both environmental stress treatments during early life stages resulted in a significant change in gene expression patterns, when larval instar V stage appear to be of a particular importance. In particular, genes involved in development, reproduction, and metabolism were significantly enriched in response to crowding and starvation stress. Furthermore, in adult individuals raised in food limitation, the expression of genes involved in environmental sensing was altered, whereas, in response to crowding, immune genes were differentially activated.
Our study offers a comprehensive insight into the genetic basis of environmental response in migratory insects and highlights candidate genes which may underpin migratory syndrome.

Introduction

Migration is a widely adapted response to swifts in the environment allowing organisms to avoid unfavorable conditions (Aidley, 1981). Migratory movements are characterized in various groups of organisms, however vertebrates traditionally received more attention, while the field of insect migration and migratory genomics is still in its infancy (Chowdhury et al., 2021; Dingle, 1978, 2001; Johnson, 1963). Migratory behaviour is a faculative trait, imposing that phenotypic plasticity is exhibited as a decision to start migratory movement in response to environmental cues. Therefore, one of the crucial aspects in understanding the genetic basis of migration is

pinpointing gene regulatory networks leading into migratory syndrome initiation in response to the environment (Harringmeyer et al., 2021; Liedvogel et al., 2011). While environmental cues play a vital role in triggering behavioral switches, mechanism of their processing and response on the molecular level has been studied in only a few migratory insect species such as migratory locust (*Locusta migratoria*) and monarch butterfly (*Danaus plexippus*) (Kang et al., 2004; Merlin et al., 2020; Zhu et al., 2009).

Following the process of acquisition of environmental signals, migratory insects commonly face trade-offs in resource allocation (Guerra, 2011; Hanski et al., 2006; Tigreros & Davidowitz, 2019) and need for a rapid physiological response (Bhaumik & Kunte, 2018; Tigreros & Davidowitz, 2019). The key trade-off characterizing migratory syndrome in insects is termed oogenesis-flight syndrome and described as delaying reproduction in favor of migration (Johnson, 1963; Rankin et al., 1986). Different migratory insect species exhibit significant variations in their migration traits, physiological and behavioral integration and degree of strength of exhibiting of oogenesis-flight syndrome. It could vary from complete reproductive arrest during migration (for example in boll weevil (Rankin et al., 1994), beet webworm (Y. X. Cheng et al., 2012) to exhibiting syndrome in certain generations (monarch butterfly (Malcolm et al., 2018)) to lack of reproductive diapause in a sense of migrating with developed eggs (summarized in (Tigreros & Davidowitz, 2019)). Therefore, more detailed look is needed to enhance our understanding of this complex phenomenon. Empirical observation of oogenesis-flight syndrome comes predominantly from phenotypic observations (Oliveira et al., 2006; Rankin et al., 1986), and while this physiological change is described in some detail, our studies of guiding genetic mechanisms is limited those of control of reproductive diapuse in overwintering individuals (Green & Kronforst, 2019; Herman, 1981); therefore opening a need for detailed study of molecular regulation and gene expression facilitating this major process. Moreover, environmental cues triggering this trade off need further investigation.

Reading of the environmental cues is essential for the expression of the migratory syndrome both in adult individuals and during the entire insect development (Angelo & Jr., 1984). In particular two environmental larval stressors were associated with subsequent variation of migratory propensity: larval density and periodic starvation. Larval density leads to heightened competition and the development of migratory tendencies as a strategy to disperse (Bauerfeind & Fischer, 2005). The desert locust is a notable example of this phenomenon, exhibiting a density-dependent phase polyphenism that triggers a transition from a benign, solitary phase to a more gregarious, highly migratory phase (Gasking Butler & Innes, 1936; Kang et al., 2004). In Lepidoptera larval density-dependent migration is observed in fall armyworm (*Spodoptera frugiperda*) (S. Wang et al., 2023), while larval density is associated with outbreaks of the other agricultural pest, beet webworm (*Loxostege sticticalis*) (Y. X. Cheng et al., 2012). The effect of lower food quantity in insects manifests in reduction in body size, fat storage, fecundity and investment in reproduction, which in turn may have major influence in migration capacity and a switch between reproduction and migration (Bauerfeind & Fischer, 2005; Boggs & Freeman, 2005; Chen & Ruberson, 2008; Niitepõld, 2019).

The painted lady (*Vanessa cardui*) is an emerging model species for studying genomic basic of multigenerational long-distance migration (García−Berro et al., 2023; Lohse et al., 2021; Shipilina et al., 2022; Stefanescu et al., 2013). In addition to the longest migratory distance among Lepidoptera (Suchan et al., 2019, Reich, in prep), *V. cardui* is completely lacking a diapause, which highlights the balance between reproduction and migration as a dominating adaptation in the species. The oogenesis-flight syndrome is well pronounced in *V.cardui* (Stefanescu et al., 2021), but phenotypical differences in migration distance and propensity are also recenlty observed (Reich, in prep). Recent availability of the genomic resources made it possible to address questions of genetic basic of migration in the species, however very limited studies available looking at different components of migratory syndrome from genetic and epigenetic perspectives ((Boman et al., 2023; Näsvall et al., 2023) and gene expression studies have not been performed.

As regulation of migratory behaviour is likely as multifaceted as the trait itself it's advantageous to disentangle response to various environmental cues in separate experiments addressing both rapid "decision making" in adults (oogenesis-flight syndrome) and long term effects of larval stress. In this study, we present the first attempt to investigate transcriptomic response to several environmental stress factors in light of migratory behavior, proposing following aims:

1. Characterize differential transcriptomic response to key environmental stressors in migratory butterfly, both in adult females and throughout developmental stages,
2. Identify crucial developmental time points at which the environment cue is read.
3. Attempt to identify candidate genes involved fascilation a swift btwen reproduction and migration
4. Study the effect of population density and food limitation on development, and their connection to migratory propensity.

Our ultimate goal is to advance our understanding of migratory behavior and contribute to comprehensive understanding of migratory syndrome on the new level in the genomic era (Merlin et al., 2020). Furthermore, this study deepens our comprehension of interaction between individuals, their genes and the environment.

Methods

Experimental setup

Painted lady (*Vanessa cardui*) butterfly females were collected in Catalonia, Spain, and individually housed in cages for egg laying at 25°C under an 18:6-hour light-dark regime. The individuals were provided with a host plant (*Malva sylvestris*) and a 10% sugar water solution as a food source. The F1 offspring were raised together in a shared cage under the same conditions and subsequently divided into two separate experiments upon reaching adulthood. The total number of individuals used in this study is summarized in Suppl. table 1.

Two experiments were carried out to analyze the transcriptomic response to different environmental cues in Vanessa cardui. The first experiment investigated the influence of the presence or absence of the host plant (Fig. 1B) as a potential initiator of the trade-off between dispersal and reproduction. The second experiment examined the combined effects of larval density and food limitation (Fig. 1B). The experimental setup was identical to a previously published study by Boman et al. (2023), which included a small number of individuals also included in our study (see Suppl. table 1).

Host plant limitation

F1 adult females were marked individually and released into one of two large cages (808050 cm) with free-flying males. One cage contained an abundance of host plants (9 pots with *M. sylvestris*), while the other cage lacked host plants (Fig. 1A). Both experimental groups were provided with 10% sugar water as a food source, and the temperature and light regime remained unchanged. After five days of emergence, when the predisposition to either mate or migrate is established (Stefanescu et al., 2021), the females were collected, and samples were rapidly frozen in liquid nitrogen.

Larval density and food resource limitation

In the subsequent part of the experiment, F1 females were individually placed in flasks containing *M. sylvestris* after their initial mating. The F2 eggs laid by each female were collected and divided into three treatment groups (Figure 1): LDAL (low density, local food), HDAL (high density, local food), and HDLI (high density, limited food) (Figure 1B). In the LD (low-density) conditions, larvae were individually housed (flask size to be specified), while in the HD (high-density) treatment, 10 larvae were kept together in a single flask. Both density treatment groups had continuous access to food (*M. sylvestris*), which was replaced on a daily basis. In the LI (limited resource) treatment group, the food was replaced every other day, creating a mild starvation regime. Individuals in this treatment group were maintained in groups of 10.

Samples were collected at four developmental stages: larvae (instar III and instar V), pupae, and adults. Larvae were harvested on the day they entered the larval stage, pupae were sexed and collected one day after pupation, and adults were harvested on the day of emergence. Prior to nucleic acid extraction, individuals were rapidly frozen in liquid nitrogen and stored at -80°C.

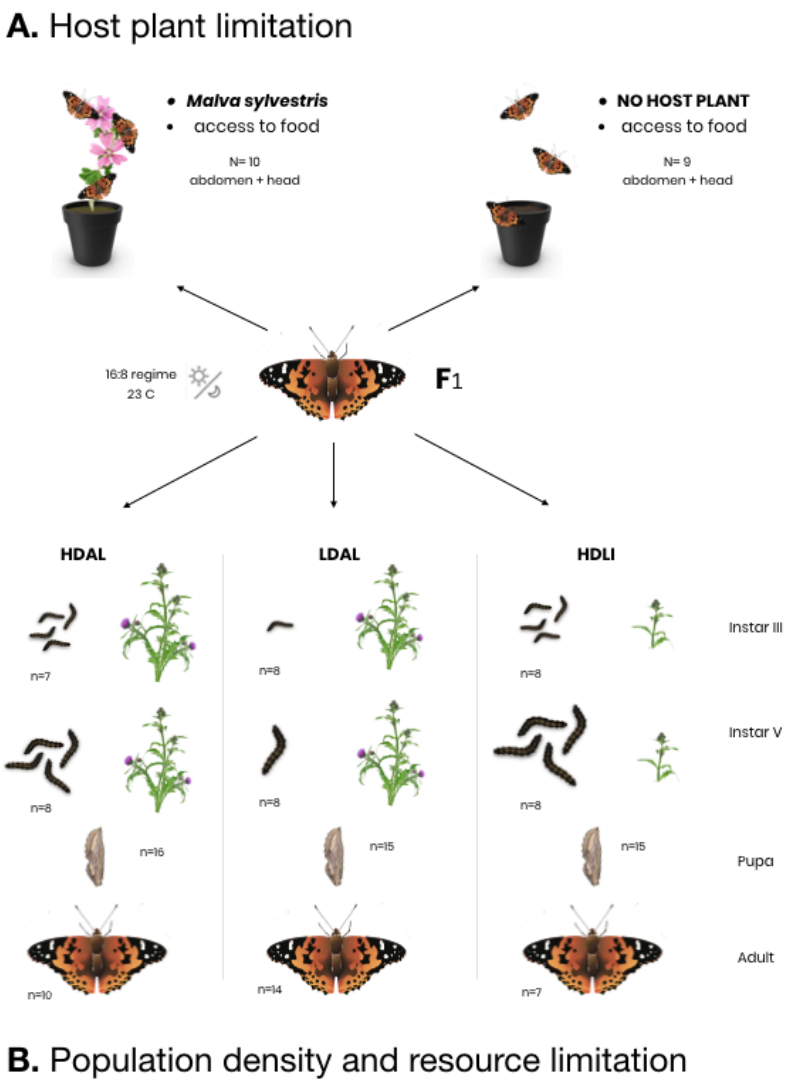

Figure 1. Experimental setup depicting two experiments conducted on offspring of wild-caught Vanessa cardui females (in the centre). Numbers of individuals sampled in the study provided. A) Host plant limitation experiment (*Malva sylvestris*). B) Larval density and food limitation experiment.

RNA extraction and sequencing

Two types of tissue, heads including antennae and abdomens between the 6th-8th segments (see the number of samples in Fig. 1 and Suppl.Table 1), were separated for subsequent RNA extraction and processing. The tissues were homogenized using a micro-pestle in guanidine-isothiocyanate lysis buffer, followed by QiaShredder. RNA extraction was performed using the RNeasy Mini Kit (Qiagen) following the recommended guidelines. RNA integrity and fragment length were assessed using 1% agarose gel electrophoresis, followed by measurements using NanoDrop (ThermoFisher) and Qubit (ThermoFisher). Sequencing

libraries were prepared using the Illumina TruSeq Stranded mRNA polyA selection kit and sequenced by the National Genomics Infrastructure (NGI) in Stockholm. Sequencing was conducted on two lanes of one S4 flow cell on the NovaSeq S6000 platform, generating 150 bp paired-end reads.

Differential expression analysis

For all steps of the processing, from adapter filtering to read mapping and transcript quantification, Nextflow nf-core (Di Tommaso et al., 2017; Ewels et al., 2020) pipeline rnaseq v3.8.1 was employed (Harshil Patel et al., 2023). In brief, raw sequencing reads were trimmed using Trim Galore! v0.6.7, utilizing Cutadapt v.3.4 (Martin, 2011). STAR v2.7.10a (Dobin et al., 2013) was used for mapping the reads to a previously published genome aasmbly (Lohse et al., 2021) with minor modifications. Read quantification was carried out using salmon v1.5.2 (Patro et al., 2017), and gene expression levels were measured in transcripts per million reads (TPM) values. Differential expression analyses were conducted in R v.4.2.1 using DESeq2 v1.28.0 (Love et al., 2014).

To assess differential expression in adult individuals experiencing host plant limitation, we employed the Wald test. Our experimental design incorporated the correction for family effects, considering treatment as the primary variable: ~family+treatment. Due to incomplete family assignment for some samples, we utilized PCA analysis of mapped reads to recover the missing assignments. We applied the same Wald test for differential expression analysis in adult individuals subjected to environmental stressors, specifically food limitation and a dense larval environment. Additionally, in this experiment, we accounted for the potential effect of sex by including it as a variable in the model: ~family+sex+treatment.

The differences in differential expression across developmental stages were assessed using the likelihood ratio test mode of DESeq2 (model = "LRT"). This test compared the fit of a full model (~family + devstage + condition + condition:devstage) with a reduced model that excluded the interactive effect between the treatment ("condition") and developmental stage ("devstage") variables. This analysis aimed to evaluate whether the effect of the condition on gene expression differs across different developmental stages. The same model was applied to both head and abdomen tissues, and the analysis included the treatments of food limitation and population density.

For further analysis, candidate genes were selected based on the criteria of a p-value < 0.05 and a log fold change of 2. The GeneOverlap package (Li Shen, 2017) was used to assess the significance of overlaps between candidate gene sets in different tissues. Since tests using LRT typically result in larger gene sets, clusterProfiler (Wu et al., 2021) was applied to identify functional clusters within all sets of candidate genes across four ontogenetic stages, using parameters (consensusCluster = TRUE, groupDifference = 2) on rlog-transformed data. In the case of the head tissue, where 745 candidate genes were identified, more stringent clusterization parameters were employed with a group difference of 3. We attempted to identify gene names by utilizing previously obtained annotation and conducting additional BLAST (Altschul et al., 1990) searches using an NCBI online tool.

To explore the functional categories represented within the candidate genes, we conducted two types of enrichment analysis: based on Gene Ontology (GO) terms and KEGG pathways utilizing previously obtained functional annotation (Shipilina et al., 2022). Enrichment analysis of GO terms was performed using the TopGO package (Adrian Alexa, 2017), employing the "weight01" algorithm, with a focus on the biological processes category. Enriched terms were further examined, visualized, and analyzed if they reached a statistical significance threshold of $p < 0.01$. Additionally, enrichment analysis of KEGG terms was performed using the enricher module of clusterProfiler.

All analysis and visualization were performed in R (RStudio version v.4.2.1), composite figures assembled using Adobe Illustrator 23.0.4.

Results

Gene expression patterns in response to limited host plant availability

We investigated the differential gene expression in the head and abdomen tissues of *Vanessa cardui* in response to limited host plant availability treatment (Fig. 2). A total of 10 individuals were analyzed for head tissue, while 9 individuals were analyzed for abdomen tissue. In the head tissue, we found 34 upregulated genes (0.30% of the analyzed genes) and 44 downregulated genes (0.39% of the analyzed genes) with an adjusted p-value of less than 0.05. In the abdomen tissue, we observed 45 upregulated genes (0.40% of the analyzed genes) and 26 downregulated genes (0.23% of the analyzed genes) with an adjusted p-value of less than 0.05.

To conduct a detailed analysis and visualization, we selected candidate genes with a fold change difference greater than 5 and a -log10(p-value) higher than 10 (Fig. 2 A,B). Notably, we identified an extreme outlier gene, the ecdysone oxidase gene, which exhibited a remarkable more than 20-fold difference in expression across both treatments. Other genes encompassed a diverse range of functional groups, including immune genes (gloverin, attacin, cytochrome 450, PGRP), metabolic genes (lipase), and genes involved in endoskeleton formation (cuticle protein). The full list of genes can be found in Supplementary Material 1.

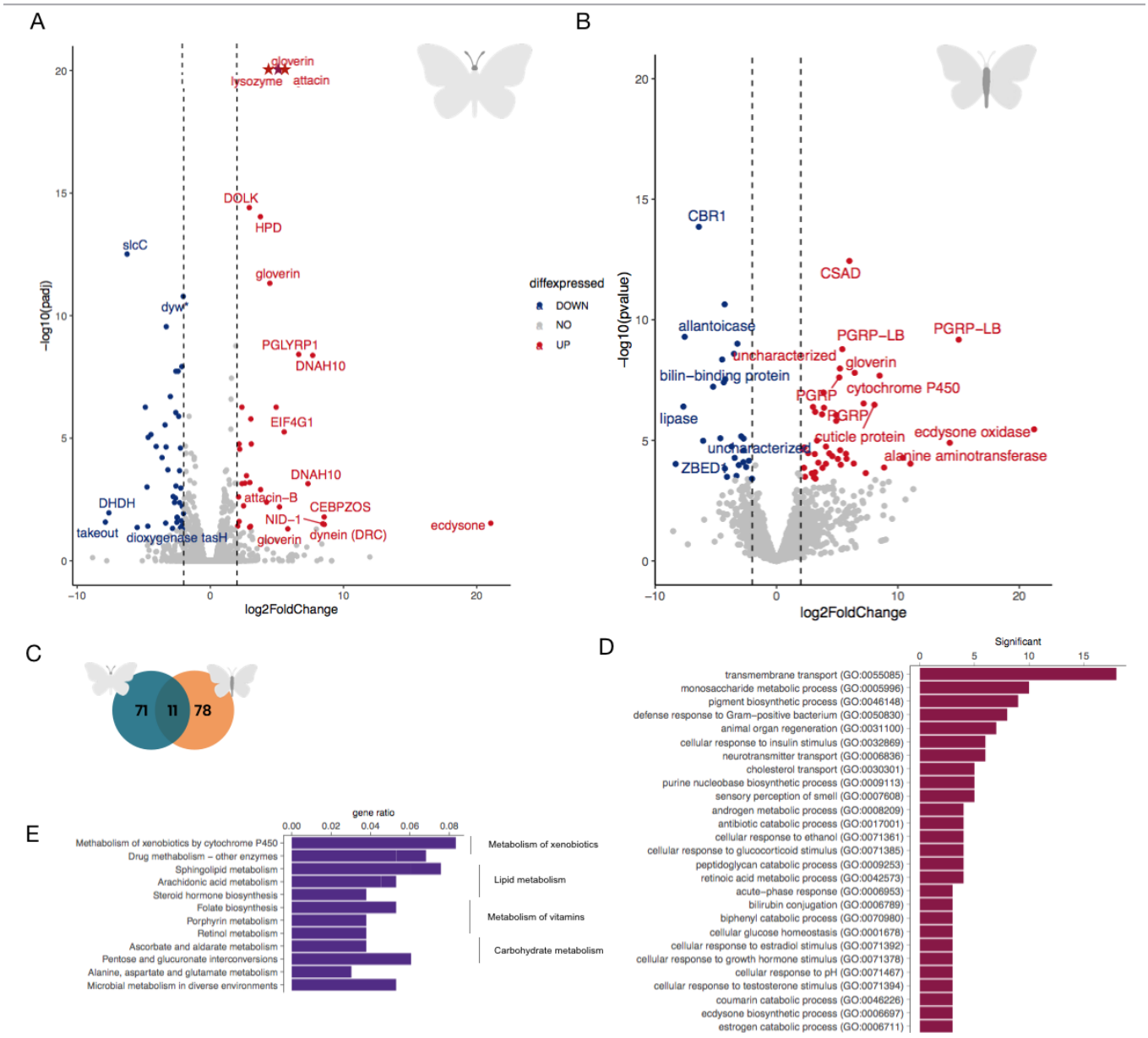

Figure 2. A) Differentially expressed genes in the host plant availability experiment (head tissue). Genes with a fold change difference greater than 2 and p > 0.05 are depicted in red, while genes with a fold change difference less than 2 and p > 0.05 are shown in teal. B) Differentially expressed genes in the host plant availability experiment (abdomen tissue), labels corresponding to A. C) Overlap between gene sets in the two tissues. D) Gene ontology enrichment analysis of the candidate genes (significance threshold p < 0.01). E) KEGG pathway enrichment analysis and their higher hierarchical grouping (displayed on the right).

Multiple genes exhibited differential expression in both tissues, highlighting their widespread significance and underscoring the importance of these genes, such as the ecdysone oxidase (Fig. 2C). To gain further insights, we conducted functional categorization of all candidate genes identified in the analysis, combining data from both tissues. This categorization was performed through enrichment analyses using GOterms and KEGG.

The analysis revealed a range of functional categories associated with the candidate genes. The most prominently represented GOterms encompassed transmembrane transport, various metabolic processes (including ecdysone biosynthesis), and defense response. Consistently, in the separate pathway analysis, we observed a high enrichment of metabolic pathways: lipid, carbohydrate, vitamin, and xenobiotic metabolism were notably enriched.

Gene expression patterns in response to food limitation across developmental stages

Next, we focused on investigating differential gene expression in a food limitation experiment, specifically analyzing the expression patterns in adult individuals raised under stressful conditions and identifying genes that exhibited differential expression patterns across developmental stages. Through a likelihood ratio test, a significant number of such genes were discovered: 745 genes in the head tissue and 321 genes in the abdominal tissue. Notably, the gene sets demonstrated a high degree of similarity (Jaccard index=0.1, p-value=9.6e-30) (Fig. 3A), with a larger number of differentially expressed genes observed in the head tissue. To gain a better understanding of which stages contributed to overall expression patterns, we performed cluster analysis.

Clustering analysis groups genes based on their expression patterns, facilitating the identification of genes with similar profiles and potential functional relationships. In the head tissue, the most prominent cluster comprised 123 genes (16.51% of the differentially expressed genes in this tissue) (Fig. 3B). The majority of expression differences within this cluster were observed during the initial developmental stage, instar III, coinciding with exposure to environmental stress. Similarly, in the abdominal tissue, 149 genes (46.42% of the differentially expressed genes) formed a distinct cluster (Fig. 3C). In contrast to the head tissue, genes within this predominant cluster exhibited differential expression during the second investigated developmental stage, specifically instar V.

Analysis of GOterms of the candidate genes in head tissue (Fig. 3D) showed broad patterns of metabolic processes and regulation, and several more specific terms: epithelial cell development, sarcomere organization, angiogenesis regulation and blood coagulation. While GOterms were supported by a somewhat smaller number of genes, abdomen genes (Fig. 3E) clearly linked to reproductive processes, multiple terms related to neural and immune cell development. Overall, the enriched GO terms indicated effects on the development of nearly all organ systems. We discovered terms such as epithelial cell development, sarcomere organization, regulation of angiogenesis, regulation of dendrite development (specifically, dendritic spine development), and hemocyte proliferation. In addition, a joint KEGG pathway analysis of both tissues indicated the involvement of ribosome biogenesis in eukaryotes and metabolism.

Intriguingly, we observed that the influence of initial environmental stress on gene expression appears to diminish as the organism transitions into the pupal and adult stages. Therefore, we focused on the differences in gene expression of adult individuals that had experienced environmental stress during th larval stage (Fig. 3 F,G). In the head tissue, only six genes

showed significant differential expression (assuming more then 2-fold change difference): Tret1, odorant receptor, UDP-glucosyltransferase, esterase, zinc-finger MYM. In contrast, the abdomen tissue exhibited a surprisingly high number of differentially expressed genes in response to this treatment. Among of the most prominent outliers were cuticle protein, gloverin, glutamine synthetase, tektin, clavesin, gooseberry-neuro, orcokinin, tyrosine.

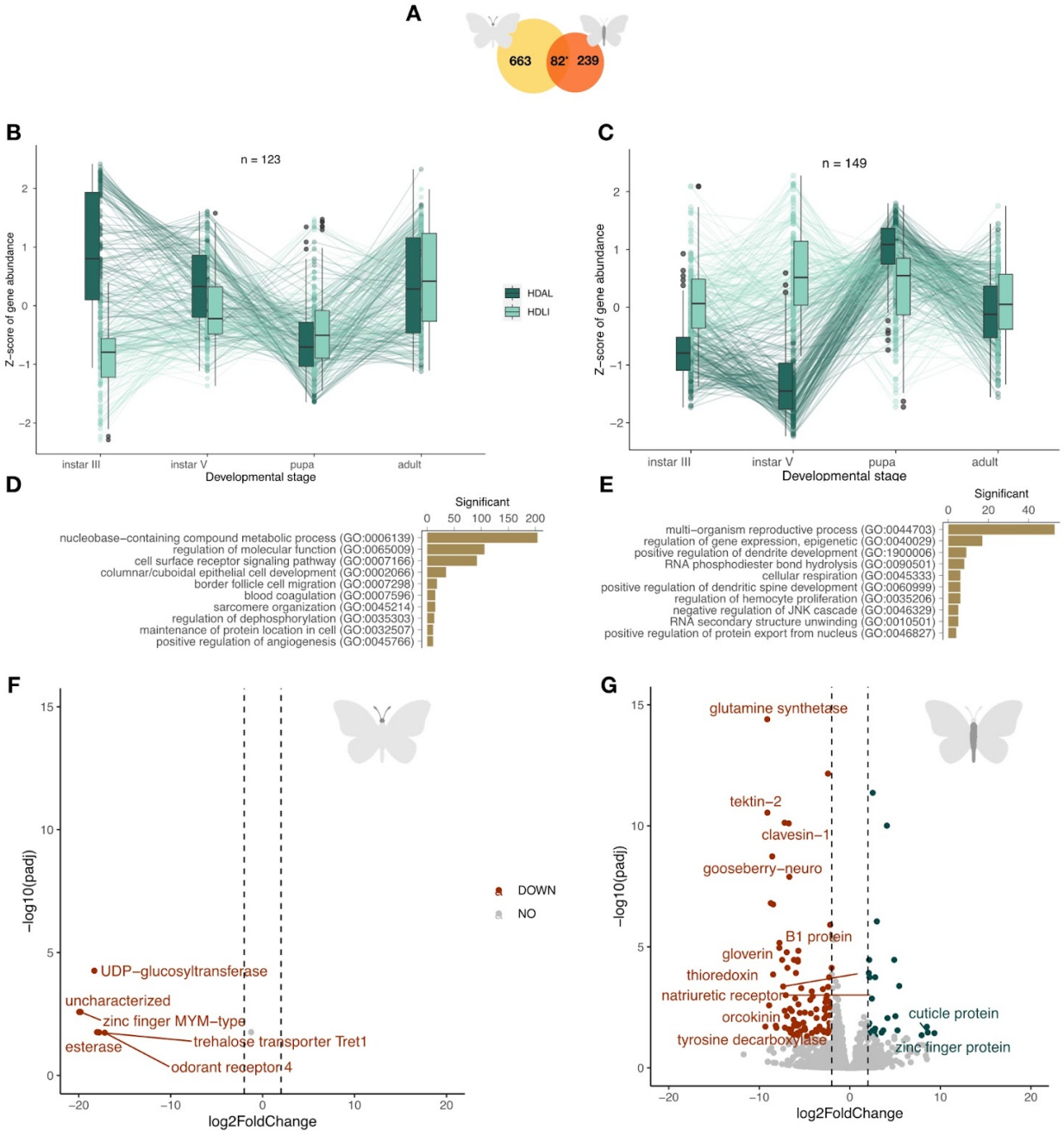

Figure 3. The food limitation experiment. A) Candidate genes differentially expressed in the experiment, overlap between gene sets for head and abdomen tissues is shown, star indicates significance. Patterns of differential expression across ontogenetic stages in head B) and abdomen C) tissue are shown in a standard box plot, outliers are indicated with circles, individual gene expression values are connected with trend lines. The top-10 most significant GOterms ($p < 0.01$) of the differentially expressed genes in head D) and abdomen E). Differential expression analysis of the adult individuals head F) and abdomen G) tissues, genes with fold change difference more than 2 and $p > 0.05$ are shown in red, genes with fold change difference less than 2 and $p > 0.05$ in teal.

Gene expression patterns in response to larval crowding across developmental stages

We found a large number of genes differentially expressed under larval crowding treatment conditionally to the developmental stage. In the head tissue we discovered 222 genes consistent with the proposed model (LRT test), 372 genes were found in the abdomen tissue. These gene sets significantly overlapped (p-value=$1.2e^{-80}$, Jaccard Index=0.2)

In order to have a look at patterns of differential expression of obtained candidate genes throughout the developmental stages we used ClusterProfiler. This pattern analysis revealed two consensus clusters with 143 (38.44% of genes) (Fig. 4B) and 69 (18.54% (Fig. SXX) genes grouped within them. Similarly to the starvation experiment we observed stronger difference between earlier stages (both exposed to environmental stress), difference in gene expression evened out at pupal and adult stages. For abdomen tissue ClusterProfiler revealed one cluster of genes (n=83, 37.39%, Fig. 4C). We observed the same pattern differential expression for this tissue.

We utilized gene ontology enrichment analysis (Fig. 4) to gain insights into the functional roles of obtained candidate genes. In the head tissue, the most enriched functional category was the regulation of filopodia assembly (Fig. 4D). Additionally, we observed enrichment in gene ontology terms related to sperm maturation, ephrin signaling, and several other functional categories. In the abdomen tissue, we found a prevalence of genes involved in reproductive processes, which include various functions such as egg formation, egg laying, and mating development. We also identified several enriched terms associated with signal transduction and the activation of different pathways. These included ephrin signaling, Ras signal transduction (involved in cell growth, division, and differentiation), Notch signaling (associated with neurogenesis), and JNK signaling (known as the "death pathway"). Apart from JNK, its regulation encompassed more general apoptosis processes and the regulation of ubiquitin-dependent processes, indicating that programmed cell death process may play important role in molecular response to this treatment.

Due to the limited amount of differential expression captured by the cluster analysis in adult individuals, we conducted a separate analysis specifically focused on the differential expression during this stage. In the head tissue, we observed differential expression of six genes (Fig. 4F). Out of these, we obtained functional information for five of them. In low-density conditions, we found upregulation of two copies of SUMO ligase and a trehalose transporter. In high-density conditions, we observed upregulation of pickpocket and NADH dehydrogenase. In the abdomen tissue, we identified the differential expression of peptidoglycan recognition protein (PGRP), chitinase, hemocytin, D-arabinitol, and NCAM.

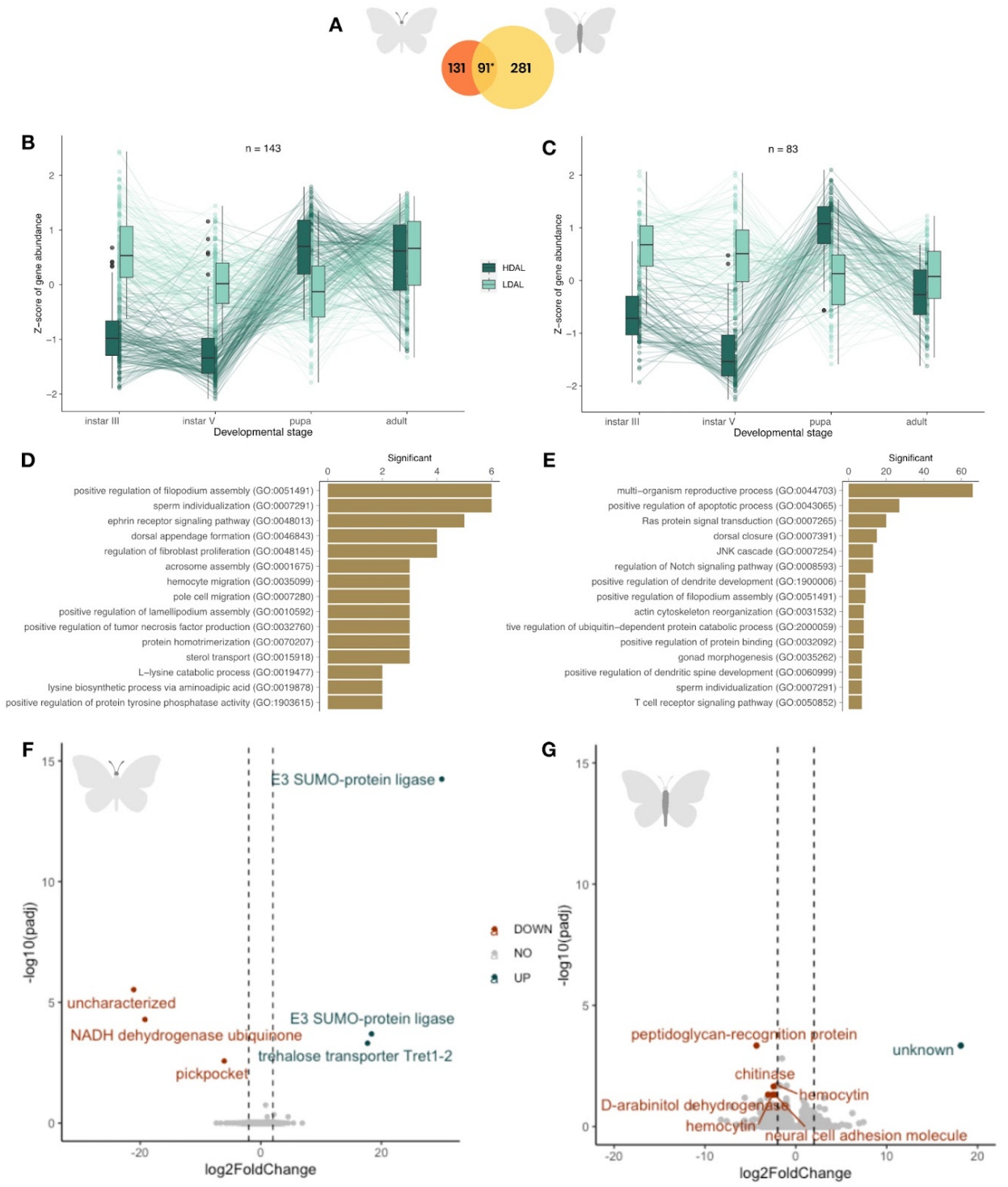

Figure 4. The larval density experiment. A) Candidate genes obtained by linear model comparison, overlap between gene sets are shown, star indicates significance. Patterns of differential expression across ontogenetic stages in head B) and abdomen C) tissue are shown in a standard box plot, outliers are indicated with circles, individual gene expression values are connected with trend lines. The top-10 most significant GOterms ($p < 0.01$) of the differentially expressed genes in head D) and abdomen E). Differential expression analysis of the adult individuals head F) and abdomen G) tissues, genes with fold change difference more than 2 and $p > 0.05$ are shown in red, genes with fold change difference less than 2 and $p > 0.05$ in teal.

Discussion

Transcriptomic response to host plant abundance, migratory behavior initiation

The ability of insect species to rapidly switch between migration and reproduction is likely to be a key adaptation for multiple migratory insects (Johnson, 1963; Rankin et al., 1986; Rankin & Burchsted, 1992) and might be of great significance for non-diapausing long-distance migrants in particular (ex. migratory females during fall migration in monarch butterflies (Malcolm et al., 2018)). In the Monarch butterfly, one of the most extensively studied species, it appears that different generations may have varying needs for resource allocation between flight/dispersal and reproduction. This variation persists irrespective of the presence or absence of the oogenesis-flight syndrome in classical definition, emphasizing the need for investigation of molecular mechanisms underpinning this trade-off. One important environmental trigger for the initiation (or termination) of migratory behavior is the presence and abundance of host plants, as especially observed in *Vanessa cardui* (Stefanescu et al., 2021).

The findings of our study underscored the importance of hormonal regulation for the plastic response to host plant abundance, which is validated by the noticeable change in the expression of multiple genes regulating developmental hormones. Notably, among these genes, ecdysone oxidase stood out due to its prominent role and significant alteration in expression. The steroid hormone ecdysone is crucial for numerous biological processes in metamorphic insects during major developmental transitions, maturation of terminal oocytes, and control of oviposition (D. Cheng et al., 2018; Niwa & Niwa, 2014; Robbins et al., 1968; Swevers & Iatrou, 2003). Ecdysone oxidase (EO) in turn plays a critical role in the ecdysone regulation. By converting ecdysone to 3-dehydroecdysone (3DE), it establishes a reversible pathway crucial for the synthesis of active ecdysteroids, which facilitates a rapid feedback loop controlling the synthesis of the ecdysone-like hormones (C.-F. Wang et al., 2018). We propose that the overexpression of ecdysone oxidase is a crucial factor enabling rapid response to host plant abundance, given that ecdysone levels are typically lower in adult insects (D. Cheng et al., 2018; C.-F. Wang et al., 2018). We hypothesize that the detailed orchestration of this process likely involves coregulation of multiple genes discovered in our study, all potentially part of the same pathway (Dubrovsky, 2005). These candidate genes include the juvenile hormone, possibly controlled by the upstream expression of JHE (Dubrovsky, 2005; Herman, 1981; Leyria et al., 2022), daywake and takeout genes which are juvenile hormone binding proteins, nrf-6, a neuropeptide and hormone receptor (Zinke et al., 2002), and cytochrome P450, which among other functions has been shown to control ecdysone biosynthesis (Namiki et al., 2005). In *Vanessa cardui*, JHE has been demonstrated to be significant and upregulated in butterflies with access to host plants (Nasvall, 2023). JH has been shown to manage the trade-off between reproduction and flight in other migratory species like the monarch butterfly, where higher JH levels lead to reproductive diapause and trigger migratory behavior (Green & Kronforst, 2019; Herman, 1981; Herman & Tatar, 2001; Qiu et al., 2023). Moreover, JH expression has been observed in actively migrating insects (Doyle et al., 2022). Together with the ecdysone pathway,

juvenile hormone plays a pivotal role in the plastic response of insects (Leyria et al., 2022; Qiu et al., 2023), potentially being key components to rapid switching between migration and reproduction.

In addition to the activation of hormonal regulation, our study illustrated an alternation of numerous metabolic processes and a shift in immune potential, as revealed by the differential expression of various genes. Fine-tuning the regulation of immune genes and metabolism appears to be crucial for migration syndrome as a whole Doyle et al., 2022. Lacking the capacity for self-immunity through antibody production (Bangham et al., 2006; H. Jiang et al., 2010) (Lemaitre et al. 1997; Iwanaga and Lee 2005), insects depend on cellular responses to neutralize pathogens they encounter during migration. In our experiment, we likely observe diverse phases of immune reaction, during which numerous genes seem to be upregulated. For example, multiple peptidoglycan-recognition proteins (PGRPs) guide recognition of various pathogens (Bangham et al., 2006; Tong et al., 2022), initiating toll signaling pathway induce production of antimicrobial peptides such as attacin (Buonocore et al., 2021), gloverin (Xu et al., 2015), lysozyme (In Seok & Yoe, 2005), and cecropin (Kim et al., 2015; Lee et al., 2015; M. Wang et al., 2021). We propose that an activated immune response could be indicative of preparedness for dispersal in a host plant experiment. Thus, separating reproductive and migratory phases allows the maintenance of essential physiological functions like immunity during migration (Rankin & Burchsted, 1992). Overall, immune gene evolution is very dynamic in migratory species and may have particular importance in *Vanessa cardui* (Shipilina et al., 2022).

Metabolic efficiency is of the highest importance in migrating species, where fat serves as the most effective energy store available to insects (Arrese & Soulages, 2010; Beenakkers et al., 1981), lipids become the main fuel for flight activity. This metabolic reliance on lipids may be shown by activation of the pathways of lipid and carbohydrate mobilization. Our study illuminates this aspect by identifying differential expression of key enzymes and genes such as lipases (Jones et al., 2015). These components potentially play a critical role in this resource allocation, acting as biological switches that underlie the phenotypic transformation from the migratory to the breeding state (Bhaumik & Kunte, 2020; Brower et al., 2006).

In conclusion, our findings highlight the central role of hormonal regulation in the response to host plant absence in these insect species, which in turn connects to the fundamental principle of tuning the reproduction-flight response. Future research could delve deeper into these regulatory pathways, perhaps with a particular focus on the ecdysone pathway, although achieving a comprehensive understanding of the migratory genotype remains a significant challenge. Nonetheless, it's important to note that the propensity for migration is not exclusively dictated by adult responses to environmental cues. It may also exhibit a degree of plasticity depending on the environmental conditions experienced during the larval stage. We explore this intriguing aspect in the next section.

Larval food limitation and crowding facilitate alternation of *V.cardui* development

The density of the population and nutrition restrictions have been identified as influential factors that can impact the development and morphology of migratory insects (Applebaum & Heifetz, 1999; Bauerfeind & Fischer, 2005; Bhavanam & Trewick, 2022; S. Wang et al., 2023; F. Yang et al., 2015), which in turn could affect the flight response norms and migration propensity in adults (Bhavanam & Trewick, 2022). Our findings demonstrate that both population density and nutritional restrictions exert significant effects on gene expression profiles (Fig. 3,4). Notably, we observed pronounced alterations in developmental pathways, suggesting a potential link between environmental stress and developmental plasticity. Analysis spanning across different developmental stages enabled us to gain initial insights into the timing of gene expression, shedding light on critical periods.

Our experimental setup enabled us to assess the individual effects of each treatment and gain understanding of their potential combined influence, as the treatment groups were allowed to overlap. We noted that the most pronounced peak of gene expression occurred during instar V. This late larval stage represents a crucial developmental point in the life cycle of insects (Truman & Riddiford, 2019; C.-H. Yang et al., 2016), since it provides the last window in response to the environment, while adult insect morphology is irreversible. It is accompanied by rapid growth and physiological changes and shown to have a particular importance for plastic response (Mirth et al., 2021).

Environmental stress caused profound differences in gene expression and showed transcriptomic signatures of development alternation in both periodic starvation and larval crowding experiments. Developmental pathways of nearly all organ systems were affected: epithelial cell and dendrite development, sarcomere organization, angiogenesis and hemocyte proliferation. This observation is in agreement with multitude of observations in Lepidoptera (reviewed in Boggs & Freeman, 2005): individuals experiencing starvation stress often exhibit delayed development (Chen & Ruberson, 2008), may extend their larval stages, and as a consequence face reduction of the body size (Bauerfeind & Fischer, 2005; Boggs & Freeman, 2005; Chen & Ruberson, 2008; Niitepõld, 2019; Niitepõld & Boggs, 2022; F. Yang et al., 2015). In *Vanessa cardui* food limitation it is shown to be associated with developmental delay in pupation and subsequent adult emergence (Kelly & Debinski, 1999). Noticeably, food limitation treatment appeared to specifically affect neural development. For example we observed downregulation of gooseberry-neuro (GsbN) transcription factor responsible for central nervous system development (Bonneaud et al., 2017), glutamine synthetase glial neurotransmitter recycling protein (Brunet Avalos et al., 2019; Huang et al., 2015) and neuropeptide orcokinin (Tanaka, 2016).

Similarly a close link is established between larval density and developmental processes (Applebaum & Heifetz, 1999; Bauerfeind & Fischer, 2005; Bhavanam & Trewick, 2022; Plazio et al., 2020; S. Wang et al., 2023; F. Yang et al., 2015). The larval density experiment triggered differential expression of genes that are related to development in general and development of reproductive system in particular, as illustrated by specific gene ontology terms, such as sperm individualization and gonad morphogenesis. Interestingly, some Lepidoptera such as *Plodia interpunctella*, and *Mythimna separata* increase their sperm production in response to crowding

(Gage, 1995; He & Miyata, 1997). Another specificity of transcriptomic response to larval crowding is an increased signal of changes in programmed cell death: differential expression of E3-type small ubiquitin-like modifier (SUMO) (Enserink, 2015), enrichment of genes in JNK, Ras signaling pathways (Lehembre et al., 2000; Ureña et al., 2016). Programmed cell death is a crucial process for insect development, since metamorphic stages involve multiple self-destructive mechanisms (Tettamanti & Casartelli, 2019).

Another common observation in response to environmental stress is alternation of metabolism, observed both during developmental stages and resulting expression in adults (notably Tret). More than 200 differentially expressed genes in starvation treatment are associated with the compound metabolic process. In adults we see down regulation of trehalose (blood-sugar) transporter (Tret1) responsible for fat transport (Kikawada et al., 2007) and clavesin, known to bind small lipophilic molecules (Smith & Briscoe, 2015). TRET1 is a key gene involved in transporting trehalose (insect blood sugar) synthesized in the fat body into the hemolymph (Kikawada et al., 2007). Among other candidate genes differentially expressed in adults exclusively are metabolic genes NADH dehydrogenase and several chitinases (chitin metabolism).

Finally, two candidate genes pointed to activation of environmental sensing in adult individuals, which have not experienced food stress as larva. Odorant receptor 4 and esterase-like (odorant degrading protein) (Godoy et al., 2021) are responsible for processing of olfactory signals in insects, which can be crucial to the successful mating of insects, locating of host plants for oviposition and food sources. Neural development and sensing are of particular importance for migratory insect development as it can largely influence plasticity of the environmental response in adults (Zhang et al., 2019).

Conclusions

We provided new insights into the complex response to environmental cues and connected them to the migratory syndrome in highly migratory butterflies. Signatures consistent with the oogenesis-flight syndrome were observed in a host plant experiment, which aimed to initiate the trade-off between migration and reproduction. This highlighted the crucial role of hormonal regulation in the response, which exhibits high plasticity.
We examined the early predisposition for migratory plasticity by subjecting larvae to different environmental cues, such as food abundance and larval crowding. This experiment allowed us to closely examine the timing of the environmental cue perception and track this process throughout development. We obtained the first glimpse of the timing of such a response and identified the peak of the response during the last larval stage, indicating its importance in developmental regulation and metabolism. Furthermore, this led us to identify genes and pathways that jointly contribute to the migratory syndrome.

Data access

RNA-seq data are available at the European Nucleotide Archive under XXXXXXXX. Scripts are available on GitHub in the following repository: XXXXX.

Bibliography

Adrian Alexa, J. R. (2017). *TopGO* [Computer software]. Bioconductor. https://doi.org/10.18129/B9.BIOC.TOPGO

Aidley, D. J. (Ed.). (1981). *Animal migration*. Cambridge University Press.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Angelo, M. J., & Jr., F. S. (1984). Body Building by Insects: Trade-Offs in Resource Allocation with Particular Reference to Migratory Species. *The Florida Entomologist*, *67*(1), 22. https://doi.org/10.2307/3494102

Applebaum, S. W., & Heifetz, Y. (1999). DENSITY-DEPENDENT PHYSIOLOGICAL PHASE IN INSECTS. *Annual Review of Entomology*, *44*(1), 317–341. https://doi.org/10.1146/annurev.ento.44.1.317

Arrese, E. L., & Soulages, J. L. (2010). Insect Fat Body: Energy, Metabolism, and Regulation. *Annual Review of Entomology*, *55*(1), 207–225. https://doi.org/10.1146/annurev-ento-112408-085356

Bangham, J., Jiggins, F., & Lemaitre, B. (2006). Insect Immunity: The Post-Genomic Era. *Immunity*, *25*(1), 1–5. https://doi.org/10.1016/j.immuni.2006.07.002

Bauerfeind, S. S., & Fischer, K. (2005). Effects of food stress and density in different life stages on reproduction in a butterfly. *Oikos*, *111*(3), 514–524. https://doi.org/10.1111/j.0030-1299.2005.13888.x