A time course analysis through diapause reveals dynamic temporal patterns of microRNAs associated with endocrine regulation in the butterfly Pieris napi

Kevin Roberts[1], Rachel Steward[1], Philip Süess[1], Philipp Lehmann[2], and Christopher Wheat[1]

[1]Stockholm University
[2]University of Greifswald

April 20, 2024

Abstract

Organisms inhabiting highly seasonal environments must cope with a wide range of environmentally induced challenges. Many seasonal challenges require extensive physiological modification to survive. In winter, to survive extreme cold and limited resources, insects commonly enter diapause, which is an endogenously derived dormant state associated with minimized cellular processes and low energetic expenditure. Due to the high degree of complexity involved in diapause, substantial cellular regulation is required, of which our understanding primarily derives from the transcriptome via messenger RNA expression dynamics. Here we aim to advance our understanding of diapause by investigating microRNA (miRNA) expression in diapausing and direct developing pupae of the butterfly Pieris napi. We identified coordinated patterns of miRNA expression throughout diapause in both head and abdomen tissues of pupae, and via miRNA target identification, found several expression patterns to be enriched for relevant diapause-related physiological processes. We also identified two candidate miRNAs, miR-14-5p and miR-2a-3p, that are likely involved in diapause progression through the ecdysone synthesis pathway, a critical regulator of diapause termination. miR-14-5p targets phantom, a gene in the ecdysone synthesis pathway, and miR-2a-3p, which has been found to be expressed in response to ecdysone. Together, the expression patterns of these two miRNAs match our current understanding of the timing of hormonal regulation of diapause in P. napi and provide interesting candidates to further explore the mechanistic role of microRNAs in diapause regulation.

Introduction

In seasonal environments, organisms are faced with abiotic stressors and limited resources. In order to cope with seasonal challenges, many organisms enter some form of dormancy (Wilsterman, Ballinger, & Williams, 2021). For insects, this dormancy comes in the form of diapause, which is a pre-programmed endogenous dormancy that is associated with cessation of development, increased stress tolerance, and decreased metabolic expenditure (Denlinger, 2022). Diapause is not a single process, as it is composed of several stages associated with different physiological and cellular processes (V. Koštál, 2006). Due to the complex nature of diapause there is a high order of physiological coordination required to initiate and successfully complete diapause (Denlinger, 2002; V. Koštál, Štětina, Poupardin, Korbelová, & Bruce, 2017). Although our understanding of the hormonal regulation of diapause is becoming more clear, much remains unknown about the mechanisms giving rise to relevant hormonal changes (Denlinger, 2022; Denlinger, Yocum, & Rinehart, 2012). Here we investigate the role of post-transcriptional modification throughout diapause by exploring microRNA (miRNA) expression across important diapause transitions. Ultimately our goal is to gain sufficient understanding of the environmentally induced gene regulatory network controlling diapause (Lehmann et al., 2018; Lehmann et al., 2016; Pruisscher, Lehmann, Nylin, Gotthard, & Wheat, 2022), that

such information via intersection with among population, heritable variation in diapause phenotypes (Lees & Archer, 1980; Pruisscher, Nylin, Wheat, & Gotthard, 2021), will begin to reveal how diapause as a phenotype evolves.

miRNAs work as a posttranscriptional gene-regulatory network influencing diverse phenotypes and likely play an important role in adaptation (Biggar & Storey, 2018; Fruciano, Franchini, & Jones, 2021; Lucas, Zhao, Liu, & Raikhel, 2015). miRNAs are 18-22 nucleotides (nt) long segments of RNA that have a characteristic hairpin structure (Wienholds & Plasterk, 2005). They are produced in the nucleus, exported to the cytosol where they are processed into mature sequences that then eventually bind to mRNA 3'-UTR to direct post-transcriptional repression (Yates, Norbury, & Gilbert, 2013). A single miRNA can target several hundred mRNAs, making them an interesting set of candidate loci that can have large impacts across diverse physiological processes (Schnall-Levin, Zhao, Perrimon, & Berger, 2010). Due to the multifaceted nature of diapause, which involves the coordination of the aforementioned diverse physiological processes, there is a need for understanding more about miRNA's role as they are an understudied functional layer that could have major impacts on regulating diapause progression (Reynolds, 2019).

To date, only a handful of studies have investigated the role of miRNAs in the diapause phenotypes of insects. These have either had a direct focus upon specific miRNAs and their correlated expression change across diapause or non-diapause states, or they have been global analyses, looking at all the detectable miRNAs and how their expression changes, but again, these have been limited to a comparison of only diapause vs. non-diapause states (Batz, Goff, & Armbruster, 2017; Meuti, Bautista-Jimenez, & Reynolds, 2018; Reynolds, Peyton, & Denlinger, 2017). Additionally, there have been several studies focusing on miRNA expression patterns during aestivation (T. Duan, Li, Tan, Li, & Pang, 2021; T. F. Duan, Li, Wang, & Pang, 2023). Together, these studies have found a range of miRNAs that are differentially expressed during dormancy, suggesting that miRNA likely play an important role in diapause. Nevertheless, more detailed study of miRNA expression temporal dynamics across diapause, and their role in transitions between diapause stages, is needed given the dynamic expression changes expected and seen at the mRNA level (V. Koštál et al., 2017; Pruisscher et al., 2022). Ultimately, such studies will result in candidate miRNA genes and targets, whose contribution to diapause can be functionally validated (Gudmunds, Wheat, Khila, & Husby, 2022), and whose variation among populations used to reveal how diapause phenotypes evolve.

Here, we focus our efforts on *Pieris napi* (Pieridae, Lepidoptera), an emerging model for understanding the ecological drivers, physiological mechanisms and molecular underpinnings of insect diapause. *Pieris napi* is broadly distributed across Eurasia, spanning a wide range of seasonal environments, with among population variation in critical photoperiod induction and termination length (Lees & Archer, 1980; Pruisscher et al., 2021). In highly seasonal environments pupae can facultatively diapause when cued by a short photoperiod (Forsberg & Wiklund, 1988). There is an extensive understanding of the timing of diapause stages and how sensitive they are to environmental influences (Lehmann, Van Der Bijl, Nylin, Wheat, & Gotthard, 2017), as well as the mechanism of diapause progression via the prothoracicotropic hormone (PTTH)-ecdysone axis (Süess et al., 2022). There is also extensive data on gene expression, metabolome, and lipidome, making it a valuable resource for linking miRNA targets with known physiological function (Lehmann et al., 2018; Lehmann et al., 2016; Pruisscher et al., 2022).

Here we use a time course analysis across the initiation, maintenance, and termination stages of diapause, following time points used in a previous RNA-seq study (Pruisscher et al., 2022). Our goal is to identify candidate miRNAs that may be important in regulating the progression of diapause, allowing future studies to query their role in among population variation in related phenotypes. First, we identify the miRNA genes and their location in the genome. Second, we then look at expression changes through diapause, cluster these changes into distinct patterns, and compare them with mRNA results. Finally, we identify interesting miRNA candidate genes based upon differential expression across multiple tissues during a critical diapause transition, along with some previous candidates from the diapause literature

Materials and Methods

Insect rearing and sampling

Pieris napi eggs were collected from wild plants in Skåne county Sweden (Kullaberg; 56°18'N, 12deg27'E and Vejbystrand; 56deg18'N, 12deg46'E), then reared using a split brood design in the laboratory following previously described methods (Lehmann et al., 2018; Lehmann et al., 2016; Pruisscher et al., 2022). To induce divergent developmental pathways, each cohort of larvae was placed in L:D 10 h:14 h at 20degC to induce diapause or L:D 22 h:2 h at 20degC to induce direct development. Upon pupation, individuals remained in their respective treatments for the first ten days, after which temperatures for the diapausing individuals were reduced to 10degC, and then to 2degC in total darkness after day 17 of pupation. Diapausing pupae remained in 2degC 24 h darkness until day 144, where temperatures were increased to 10degC, and then to 20degC on day 151. Direct developing pupae were sampled on the first day of pupation (day 0), day 3, and day 6 for head tissue. Diapausing pupae were sampled on 0, 3, 6, 24, 114, 144, and 155 after pupation (n = 3-4 individuals per time point for each treatment). When sampled, individuals were flash frozen by being submerged in liquid nitrogen. Only females were sampled to avoid the effect of sex on expression patterns, and all individuals were sampled around the same time of day (˜12:00) to avoid any circadian effects.

RNA extraction, sequencing and data processing

Total RNA was extracted from the head and abdomen of each individual separately by homogenizing each tissue in TRIzol (Thermo Fisher Scientific), then purified using Direct-zol RNA MiniPrep kits (Zymo Research) following manufacturer's recommended protocol. Small non-coding RNA (sncRNA) library preparation and sequencing was done by the Swedish National Genomics Infrastructure. Libraries were constructed using Illumina TruSeq small RNA library kits and then sequenced using Illumina HiSeq2500 50SR. Once libraries were prepared, each sample was split and sequenced in two separate lanes to prevent any lane bias in sequence count.

The two raw sequence files for each sample were merged together prior to data processing. The tool miRTrace (Kang et al., 2018) was used for quality filtering (i) removing reads with a PHRED score below 20, (ii) removing adapters, (iii) removing reads with highly repetitive or ambiguous nucleotides, and (iv) trimming reads by length (removing reads shorter than 18nt); default settings were used with *Bombyx mori*as a reference.

miRNA identification, mapping and quantification

Detection of known and novel miRNAs was done using miRDeep2 (Friedlander et al., 2008; Friedlander, Mackowiak, Li, Chen, & Rajewsky, 2012). First, reads from all samples were merged and mapped to a *P. napi*genome (Lohse, Hayward, Ebdon, of Life, & Consortium, 2021). Reads were compared to known miRNAs from all mature miRNAs and hairpin sequences from *B. mori* and all mature reads from *Heliconius melpomene* from miRbase v22 (Kozomara, Birgaoanu, & Griffiths-Jones, 2019). Remaining unmapped sequences were then used to identify novel miRNAs and their hypothetical secondary structure (Friedlander et al., 2012), giving an output of known and novel miRNAs in all samples in the data set. Identified miRNAs were then discarded if they did not meet threshold requirements of a miRDeep2 score [?] 1, presence in more than 5 samples, and a total read coverage of [?] 10 reads across samples. All identified miRNAs were then used to obtain read counts for each miRNA for each sample using the quantifier.pl script in miRDeep2.

Location of miRNAs in the genome

To identify the location of the closest genomic feature to each miRNA, the miRDeep2 output was compared to the annotated features in the*P. napi* genome using the *closest* command in BEDTools (Quinlan & Hall, 2010). While the chromosome-level genome used for this project was produced by the Darwin Tree of Life (Lohse et al., 2021), an independent functional annotation was used (Steward, Pruisscher, Roberts, & Wheat, *in review*), as the first-generation annotation for this genome had numerous overlapping coding regions and UTRs of predicted genes. The miRNAs inside the coding sequence (CDS) region of a gene were identified as either being in an intron or exon, while miRNAs outside of a CDS region were identified as either closest to a start or stop codon. The identifies miRNA origins were used for subsequent analyses, such as assessing the

functional annotations and gene ontology (GO) terms identified for the coding genes nearest miRNA loci.

Differential expression of miRNAs during diapause

Differential expression analysis was done in R [v4.1.2 (R Core Team, 2021)] using the package DESeq2 [v3.17 (Love, Huber, & Anders, 2014)] on both the head and abdomen separately. Pairwise comparisons were made between expression at each timepoint to expression of miRNAs on diapause day 0, for both head and abdomen tissues. Differentially expressed miRNAs were then identified with a false discovery rate (FDR) of 0.05, and extracted for subsequent analyses. Global expression for differentially expressed miRNAs was visualized using multidimensional scaling of filtered read counts by applying the *plotMDS* function in the EdgeR package [v3.42.4 (Robinson, McCarthy, & Bioinformatics, 2020)].

Differentially expressed miRNAs were then clustered using fuzzy c-means clustering in the Mfuzz package [v3.17 (Kumar & Futschik, 2007)] separately for head and abdomen tissues. Normalized data were first standardized with a mean of zero and a standard deviation of one, then used to estimate the optimal fuzzifier using the *mestimate*function. The optimal numbers of clusters were determined based on minimum centroid distance (Kerr, Ruskin, Crane, & Doolan, 2008). A membership threshold of 0.8 was set for each cluster in each tissue for subsequent analyses.

GO enrichment analysis of clusters

Once targets were detected we ran a gene set enrichment analysis (GSEA) on the targets of the miRNA in each cluster using the parent-child algorithm in topGO [v2.46.0 (Alexa & Rahnenfuhrer, 2020; Grossmann, Bauer, Robinson, & Vingron, 2007)] and a functional annotation for the longest isoforms of each gene in our *P. napi* annotation produced with eggNOG (Rodriguez del Rio et al., 2022).. We checked for enrichment for Biological Processes using the functional annotations of their identified miRNA targets (Wheat et al., *in review*). An FDR of 0.001 was used as a cutoff, as well as a minimum of 2 members present. Due to the number of targets in each cluster the results were dominated by large GO terms that are difficult to interpret. To focus the GSEA analysis, gene sets within cluster were filtered to only include those containing a minimal number of miRNA target sites per gene. An increased number of predicted targets per gene reduces false-positive detection of miRNA targets (Ritchie, Flamant, & Rasko, 2009), and also indicates coordination of targets within a larger set of targeted genes. Each cluster had varying numbers of miRNAs targeting genes, so the minimum number of targeting miRNAs needed per locus to be included per cluster GSEA was determined to be the number that produced a majority of the top ten GO terms containing more than one significant gene.

Hypothesis testing of major transitions during diapause

The timing of termination of diapause in *P. napi* is regulated through the Prothoracicotropic hormone (PTTH)/20-hydroxyecdysone (20E) axis, where cold accumulation is needed in order to transition from diapause to post-diapause quiescence (Suess et al., 2022). A differential expression analysis was run comparing expression in both tissues at 24, 144, and 144 days in diapause, which are timepoints where diapause termination should be occurring based on prior work, and when pupae were also in cold temperatures (Lehmann et al., 2018; Lehmann et al., 2017). The differentially expressed genes (DEGs) from each list were then checked for overlap, and the overlap were considered candidate miRNAs.

Expression of candidate miRNA from literature

Expression profiles of candidate miRNAs identified as differentially expressed between direct developing and diapausing pupae of the fly*Sarcophaga bullata* (Reynolds et al., 2017), pharate larvae of the mosquito *Aedes albopictus* (Batz et al., 2017), and pupae of the moth *Helicoverpa zea* (Reynolds, Nachman, & Denlinger, 2019) were checked for differential expression during diapause in both tissues of*P. napi* . Combined, this comprised 16 candidate miRNA genes: miR-1-3p, miR-9c-5p, miR-13b-3p, miR-14-5p, miR-31a-5p, miR-92b-3p, miR-275-3p, miR-276a-3p, miR-277-3p, miR-282, miR-283-5p, miR-286b-3p, miR-289-5p, miR-305-5p, miR-2942-3p, and bantam-5p.

Results

Identification of known and novel miRNAs

Our sncRNA sequencing returned 634.7 million reads across 73 libraries (8.7M +- 2.1M reads library^{-1}). Of those reads, 505.1 million passed qc (6.9M +- 2.2M reads library^{-1}), and 19.7 million, or 3.1%, were identified as miRNAs (0.27M +- 0.16M reads library^{-1}). Similar small fractions of identified miRNAs are commonly obtained from sncRNA datasets (Jain, Rana, Tridibes, Sunil, & Bhatnagar, 2015). Based on the length distribution of each sample, there is a peak in all abdomen samples ~ 33 nt in length, which were absent in the head tissue samples (Table S1). This peak likely corresponds to tRNA fragment length (Kang et al., 2018), and highlights that there may be an abundance of tRNAs present in these tissues or they may be serving an alternative function (Su, Wilson, Kumar, & Dutta, 2020). Using miRTrace, we identified that all miRNAs detected were 100% insect specific miRNAs, indicating that there is likely no detectable contamination in any of the samples used (Table S1).

In total, we identified 188 candidate miRNAs in *P. napi* , of which 129 were novel and 59 were identified based on known mature read sequences in *B. mori* or *H. melpomene* . Including both 3' and 5' sequences of each miRNAs totaled to 257 expressed miRNAs in our samples (Table S2). Of these 257 miRNAs, 236 were identified in the head and 207 in the abdomen, with 205 shared between the tissues (Table S3).

Location of miRNAs in the genome

Of the 188 miRNAs in *P. napi* , the distribution across autosomes was relatively even, ranging from a single miRNA on chromosome 24 to 12 each on chromosomes 1 and 17, while the Z chromosome, with 35, had many more miRNAs than the autosomes (Figure 1A). The majority were located outside of a gene coding region, with 46 being found before a start codon with an average distance of 15,876 bp (SD +- 25408 bp) and 69 after a stop codon with an average distance of 7,507 bp (SD +- 9924 bp) (Figure 1C). Of the miRNAs located within the boundary of a protein coding gene, the vast majority were located in introns, and 2 on the border of an exon and intron (Figure 1B).

Differential expression of miRNAs during diapause

We began by looking at global patterns of miRNA expression in order to gain an overview of temporal changes across diapause. General expression patterns observed via multidimensional scaling (MDS) showed that direct developing individuals have a linear development pattern largely moving along dimension 2 in both head and abdomen tissue. Diapause individuals have a divergent development that moves along dimension 1 reaching a maximal divergence at day 114-144, after which miRNA expression converges with the developmental trajectory of the direct developing individuals at diapause day 155 and direct day 3 (Figure 2).

Continuing with our global perspective, differential expression analysis found that there were 117 DE miRNAs between all comparisons in the head tissue, and 49 in abdomen tissue (Table S4). In order to identify clusters of similarly expressed miRNA during the course of diapause development, differentially expressed miRNAs among timepoints of diapause were used as input for a cluster analysis. The optimal number of clusters for each of the analyses was 7 for the head tissue (Figure 3) and 5 for the abdomen tissue (Figure 4). When DE miRNAs were filtered to fit a membership of > 0.8 to one of the clusters, the number of DE miRNAs were reduced to 83 in the head, and 25 in the abdomen. Of these filtered DEGs, 19 were in both head and abdomen samples.

GO enrichment analysis of clusters

GSEA were run on genes targeted by multiple miRNAs in a given cluster. When run on all identified targets, regardless of the number of targets per gene, the head tissue had between 1332-2035 targets, with no overlap each cluster. Abdomen tissue had fewer GO terms enriched than the head tissue, with 607-1571 per cluster. In general, the GSEA from these analyses were large and diffuse, which prevented easy biological interpretation (Table S5).

We next filtered our miRNA targets for the GSEA, with the number of miRNAs targeting a given gene that met our threshold criteria of majority of enriched GO terms having more than one gene, varied between 3 to 8 miRNAs. Once the thresholds were applied, the number of genes was greatly reduced in both the head and abdomen tissue to between 10-79 gene targets per cluster (Table S6). In the head, clusters H2 and H4 had the most miRNAs targeting single genes. Importantly, these targets all had at least 6 miRNA targets each, and this high degree of multiple miRNAs targeting these genes in these clusters makes them interesting candidates for coordinated gene expression regulation. Cluster H2 was enriched for genes targeted by multiple miRNAs related to "regulation of lipid metabolic processes" and "lipid metabolic processes" (Figure 5A), which are highly expressed in the first days of pupation and remain low through the remainder of the samples (Figure 3). Cluster H4, which is characterized by upregulation around day 24 that plateaus between day 114 and 144, then is downregulated after (Figure 3), is enriched for genes that are involved in "regulation of Wnt signaling pathway" and "imaginal disc pattern formation" (Figure 5A). In the abdomen tissue, cluster A4 was also enriched for "Regulation of Wnt signaling pathway", A4, which has similar expression through diapause as cluster H4 (Figures 4 & 5B). Although there the amount of overlap in miRNAs targeting a given gene is lower than H2 and H4, H6 is enriched for genes involved in "glucocorticoid metabolic processes", which is not hormone class present in insects (Figure 5A).

Hypothesis testing of major transitions during diapause

In order to gain additional insights beyond our GSEA, we next identified differentially expressed miRNAs during the period when pupae transitioned from the deepest portion of diapause to post diapause quiescence (day 24-144). During this period butterflies were held at 2degC, allowing us to identify miRNAs associated with diapause progression independent of temperature change. There were eight DEGs identified in both the head and abdomen tissue, and of those, only two were differentially expressed in both tissues, miR-14-5p and miR-2a-3p. These two miRNAs have inverse normalized expression patterns, where miR-14-5p increases in expression early in diapause, and declines to a minimum expression after 114 days (Figure 6). miR-2a-3p has an expression pattern that starts off at the lowest expression, and increased through time until day 114, where it stays until 144, afterwards it starts to decline (Figure 6).

Expression of candidate miRNA from literature

To compare our results to previously published data sets, we examined expression of previously reported miRNAs involved in diapause regulation, of which only 50% could be identified in *P. napi*(Table S7). Of these eight identified miRNAs (miR-9c-5p, miR-13b-3p, miR-14-5p, miR-31-5p, miR-92-3p, miR-275-3p, miR-282-5p, and bantam-5p), all of them were found to be differentially expressed during diapause in the head tissue (Figure 7), with all falling in different clusters except miR-92-3p and bantam-5p both being in cluster H4. In the abdomen tissue, six were differentially expressed during diapause (miR-13b-3p, miR-14-5p, miR-31-5p, miR-92-3p, miR-282-5p, and bantam-5p) and only 2 (miR-31-5p and miR-282-5p) found in any identified clusters. Only two were differentially expressed during diapause termination, miR-14-5p and miR-92-3p, with miR-92-3p being only significant in the head, and miR-14-5p in both.

Discussion

Diapause is a complex, plastic phenotype that requires the sequential coordination of many physiological processes, making miRNAs interesting candidates due to their regulation of mRNA expression patterns by targeting many mRNAs at once. Here, we identified miRNAs across the genome of *P. napi,* and found that there are a diverse range of miRNA expression patterns that occur throughout diapause. Ultimately, we identified several miRNAs that may play an important role in diapause timing, notably miR-14-5p, which could have a key role in diapause termination by targeting genes in the ecdysone synthesis pathway.

We identified a number of miRNAs in *P. napi,* including 129 novel miRNAs that we describe for the first time here. These genes were located throughout the genome, with 32% being located within intron of genes, which is lower than the average of 52% across 80 species (Guerra-Assuncao & Enright, 2012), but higher than the 12% reported in zebrafish (Thatcher, Bond, Paydar, & Patton, 2008). miRNAs that are located in introns can act as functional miRNAs from the spliced-out intron, although these function differently

than canonical miRNAs and are hard to detect using normal miRNA detection tools (Rorbach, Unold, & Konopka, 2018). Nevertheless, finding these intronic miRNAs suggests that there is a substantial potential for the expressed gene harboring the intronic miRNA to be functionally related to the miRNA targets (Olena & Patton, 2010).

There is also a high concentration of miRNA on the Z chromosome compared to the autosomes, which is notable since this is the sex chromosome in Lepidoptera. The high concentration on the Z chromosome appears to be driven by several known clusters of miRNAs. These clusters include the miR-2 cluster, of which we identified 4 members of the miR-2 family (miR-2a, miR-13a, miR-13b, and miR-2c), along with miR-71, which flanks the miR-2 family in many other insects, and is heavily involved in neural development and maintenance (Marco, Hooks, & Griffiths-Jones, 2012). Finally, we note that our miRNA sequencing of two body parts composed of diverse tissues, across several developmental time points, provides a rather deep sampling of the miRNA, at least for these body parts. Nevertheless, the numbers of miRNAs we have identified are roughly comparable with similar efforts in other Lepidopteran species (Kozomara et al., 2019; Quah, Hui, & Holland, 2015; Reynolds et al., 2017), suggesting that the number of identified miRNAs is probably quite representative of detectable miRNAs in this butterfly.

We next quantified miRNA expression patterns across tissues and developmental trajectories, finding that their general expression patterns closely match expression patterns found in previous metabolomic and transcriptomic (mRNA and alternative splicing expression) assessments (Lehmann et al., 2018; Pruisscher et al., 2022; Steward et al., *in review*). Using multidimensional scaling (MDS) for a global assessment of the head tissues, there is a clear alternative trajectory mostly in dimension 1 away from the direct developing individuals until the latest stages of diapause. The main axis of variation in total expression accounts for diapause, and has a larger impact on miRNA expression than development, where 40% of variance is accounted for, likely due to the high activity of neuroendocrine regulation occurring throughout diapause, which is highly pleiotropic (Williams, 1952). Despite the overall lower expression of miRNAs patterns in abdomen tissue (there is a 7-fold higher overall expression in the head tissue), these patterns were nearly identical.

In an initial analysis containing all targets of a miRNA expression cluster, the largest GO terms were often the most enriched sets, but by reducing targets to include only genes targeted by multiple miRNAs, we were able to find much more specific and informative GO terms with potentially interesting biological roles in diapause. Given our interest in the termination of diapause, we were primarily interested in clusters with peaks between 6 and 114 days in diapause, which includes H4, H6, A3 and A4. Cluster H2 was enriched for GO terms related to lipid metabolism, since the expression of this cluster decreases in the first few days of pupation, this could indicate that there is a suppression of lipid metabolism upon entry in diapause, and a cessation of miRNA regulation as pupae shift to heavily relying on lipid metabolism in diapause. However, previous lipidomic studies of these timepoints in*P. napi* shows little decrease in storage lipids through diapause, and only displays a significant decrease after diapause has terminated (Lehmann et al., 2016). Even though there was no detectable lipid store depletion found, there is a decrease in the respiratory quotient to below 0.7 ten days after pupation, indicating they are primarily metabolizing lipid, matching the miRNA expression patterns of cluster H2 (Lehmann et al., 2018). Alternatively, it could be related to the extensive remodeling of both cell membrane lipid and lipid pool composition that can occur in response to cold in insects (Enriquez & Teets, 2023; V. r. Koštál, Berková, & Šimek, 2003).Cluster H4 is enriched for both regulation of Wnt signaling and imaginal disc pattern formation. Wnt signaling has been identified as an important regulator of diapause termination (V. Koštál et al., 2017; Ragland, Egan, Feder, Berlocher, & Hahn, 2011), and expression in cluster H4 peaks after diapause is terminated. Since Wnt signaling has been found to have higher expression just before diapause is terminated, perhaps the expression of cluster H4 after diapause is terminated is a signal related to cessation of the need for an upregulation of Wnt signaling as a signal to terminate diapause (V. Koštál et al., 2017). As for imaginal discs patterning, it is a critical component of pupal development, and in some species the lack of imaginal disc development is used as an indicator of diapause (V. Koštál, Šimůnková, Kobelková, & Shimada, 2009). The upregulation of cluster H4 after termination therefor may reflect a difference between diapause and post-diapause quiescence, in that

during quiescence the pupae is able to develop if conditions allow, but diapausing pupae cannot (V. Koštál, 2006; Lehmann et al., 2017; Süess et al., 2022). Cluster H6 is enriched for the GO term "glucocorticoid metabolic processes", which notably is a class of steroid hormones that are involved in mammalian stress response, and absent in insects. Insects do respond to cortisol, which is a glucocorticoid, but cannot produce it, instead insects use ecdysteroids in a similar manner as mammals use glucocorticoids (Gawienowski, Kessler, Tan, & Yin, 1987; Ivanovic, 2018). Regardless, the enrichment of a process that is absent in the sequenced taxa warrants further study, particularly due to the importance of hormonal signaling in the regulation of diapause timing.

In order to gain physiological informative insights, we further parsed these results to only focus upon miRNAs that were differentially expressed across termination time points and within the aforementioned clusters, in order to gain biological insight via first principals complexity reduction. This narrowing of focus identified interesting candidates playing a role in diapause regulation through their interaction with the PTTH/20E axis, a main candidate being miR-14-5p. The PTTH/20E axis is a critical regulator of determining the timing of transition from diapause to post-diapause quiescence in *P. napi* , where it is hypothesized that the presence of PTTH starts a cascade of ecdysone release, which leads to termination of diapause (Süess et al., 2022). miR-14, both 3p and 5p have been found to delay larval development by decreasing ecdysone signaling in *Bombyx mori* , by targeting both ecdysone receptors (Liu et al., 2018) and targeting genes in the ecdysone-signaling pathway (He et al., 2019). In the beetle*Galeruca daurica* there is further evidence that both miR-14-3p and miR-14-5p impact the production of ecdysone, with neither miRNA expressed in response to ecdysone, as when larvae were exposed to ecdysone these miR-14s were not among the miRNAs that were differentially expressed (Jin et al., 2020). Our target identification analysis identified a miR-14-5p binding site in the 3'-UTR of the gene*phantom,* which is in the ecdysone-synthesis pathway. The ecdysone-synthesis pathway converts cholesterol to 20E, meaning that downregulating a gene in the pathway would likely delay diapause termination. In *B. mori* miR-14-5p targets two ecdysone signaling genes, neither of which are *phantom* , in addition to several ecdysone response genes (He et al., 2019). While our results do not fully align with those in *B. mori* , it is encouraging to find the same miRNA targeting the same pathway. The expression of miR-14-5p is highest in the deepest phase of diapause, and is down-regulated as diapause termination occurs, matching the timing of our expectation of ecdysone presence. This provides a putative mechanism for a miRNA to have a large role in the timing of a critical transition in diapause.

In *P. napi,* the complementary sequence from miR-14-5p, miR-14-3p was found to target a receptor of ecdysone, but is only expressed early in diapause, which is a period when pupae are not sensitive to ecdysone to terminate diapause (Süess et al., 2022). In *B. mori* miR-14-3p targets the mRNA of two 20E receptors, *ecdysone receptor* and*ecdysone-induced protein 75 (E75)* (Liu et al., 2018), while we found that miR-14-3p in *P. napi* targets *E75* . However, miR-14-3p is not differentially expressed during diapause termination, instead it is downregulated in the first three days of pupation. That*P. napi* is insensitive to ecdysone during the first few weeks of diapause (Süess et al., 2022), could therefore partially be due to a lack of available ecdysone receptors (a dimer consisting of the two proteins ecdysone receptor (EcR) and ultraspiracle (USP) (Rinehart, Cikra-Ireland, Flannagan, & Denlinger, 2001)). In sum, it is notable that we are finding connections between the miRNAs, their phenotypically relevant expression, their identified targets in our focal species, and previous literature making similar connections.

The other interesting miRNA that is differentially expressed in diapause termination is miR-2a-3p, a miRNA that has been found to be responsive to ecdysone. miR-2a-3p is involved in metamorphosis by targeting the juvenile hormone response gene krüppel-homolog 1 (Lozano, Montañez, & Belles, 2015). In *B. mori* , miR-2a-3p is more highly expressed in embryogenic cells that are treated with 20E than in unexposed embryogenic cells (Jin et al., 2020). Being expressed in response to 20E in *P. napi* would both match our expectation based on the timing of ecdysone sensitivity in diapause, and support that miR-14-5p is involved in the regulation 20E synthesis, due to the expression patterns of miR-2a-3p being inverse of miR-14-5p. However, in a hemimetabolous insect, *Nilaparvata lugens* , miR-2a-3p is transcriptionally repressed by an early 20E response gene, broad-complex (*BR-C)*(Chen, Liang, Liang, Pang, & Zhang, 2013), which if that were the case in *P. napi* , would lead to an inverse expression pattern instead of the one we observed. We cannot

disentangle the role of miR-2a-3p in diapause ecdysone signaling in *P. napi* from this data, but it provides an interesting candidate for a miRNA involved in diapause regulation and a target for future research.

When we compare differentially expressed miRNAs identified in previous diapause studies, we find an overlap with several miRNA, but many more are missing. Of the 16 miRNA identified in the literature the only overlap with DE miRNAs we identified in diapause termination is miR-14-5p. miR-14-5p is found to be downregulated in diapausing pharate larvae of the mosquito *Aedes albopictus* (Batz et al., 2017), which would only match our findings as long as the diapause timepoint is taken after day114, when it is lower than day 6 of direct developing individuals. We also see patterns in miRNAs that match existing literature in expression direction, but only if taken at the right timepoints. miR-92-3p was found to be downregulated in diapausing pupae of the flesh fly *Sarcophaga bullata* , which match our findings, but when looking at the time course of expression there is a large upregulation at day 144. Additionally, the expression of miR-275-3p matches expression direction in flesh flies, but this depends on the time of sampling the direct developing individual (Figure 7). Overall, the extent that the results from this study overlap with existing work depends on the timepoint used for comparison. These discrepancies of expression patterns at specific times during diapause highlight the dynamic nature of diapause and the importance of sampling a time course to identify patterns that would be otherwise missed.

Conclusion

MicroRNAs are a post-transcriptional mechanism with potential to impact a large number of physiological systems, well suited for regulation of diapause. There are a large number of miRNAs that are differentially expressed during diapause in a dynamic manner through time, likely impacting many different systems that are vital for diapause. By searching at specific timepoints that are known to be critical for diapause regulation, we were able to find a candidate miRNA that potentially has a large impact on diapause progression by targeting genes involved in the ecdysone pathway. Our ability to connect findings to previous literature was limited by experimental design, as previous studies only used single timepoint sampling (in vs. out of diapause), and our results indicate that miRNA expression changes substantially across the different stages of diapause progression. We find that miRNAs likely both play a large and wide role in diapause phenotype by targeting many genes, potentially targeting a critical pathway at a critical time We find that miRNAs likely both play a large and wide role in diapause phenotype by targeting many genes, potentially targeting a critical pathway at a critical time. Future work should aim to link miRNAs in diapause to broader global gene expression patterns, and to mechanistically link hormonal regulation to miRNA expression.

Acknowledgements

We thank Bengt Sennblad and Payam Emami at the National Bioinformatics Infrastructure Sweden at SciLifeLab for fruitful discussions and help with the data analyses. We would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure and NGI for providing assistance in massive parallel sequencing. This work was supported by the Swedish Research Council (grants no. 2017-04159 and 2022-03343 to PL), the Research Council Formas (grant no. 2017-00965 to PL), the Knut and Alice Wallenberg Foundation (grant no. 2012.0058) and the Carl Tryggers Stiftelse (grant no. CTS20-242). The computation was enabled by resources in project SNIC 2022/22-87 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973BS. We would also like to thank Marc Friedlander and Emilio Sanchez for advice on data analysis.

Data availability

Raw sequence data is available through the NCBI Short Read Archive (SRR25775847- 25775919) under BioPrjoect PRJNA972990. All of the bioinformatic scripts other than default commands will be provided with publication.

References

Alexa, A., & Rahnenfuhrer, J. (2020). TOPGO: enrichment analysis for gene ontology 2.40. 0. *R package version, 2* (0), 2010.

Batz, Z. A., Goff, A. C., & Armbruster, P. A. (2017). MicroRNAs are differentially abundant during Aedes albopictus diapause maintenance but not diapause induction. *Insect molecular biology, 26* (6), 721-733.

Biggar, K. K., & Storey, K. B. (2018). Functional impact of microRNA regulation in models of extreme stress adaptation. *Journal of molecular cell biology, 10* (2), 93-101.

Chen, J., Liang, Z., Liang, Y., Pang, R., & Zhang, W. (2013). Conserved microRNAs miR-8-5p and miR-2a-3p modulate chitin biosynthesis in response to 20-hydroxyecdysone signaling in the brown planthopper, Nilaparvata lugens. *Insect biochemistry and molecular biology, 43* (9), 839-848.

Denlinger, D. (2002). Regulation of diapause. *Annual review of entomology, 47* (1), 93-122.

Denlinger, D. (2022). *Insect diapause* : Cambridge University Press.

Denlinger, D., Yocum, G., & Rinehart, J. (2012). Hormonal control of diapause. In *Insect endocrinology* (pp. 430-463): Elsevier.

Duan, T., Li, L., Tan, Y., Li, Y., & Pang, B. (2021). Identification and functional analysis of microRNAs in the regulation of summer diapause in Galeruca daurica. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 37* , 100786.

Duan, T. F., Li, L., Wang, H. C., & Pang, B. P. (2023). MicroRNA miR-2765-3p regulates reproductive diapause by targeting FoxO in Galeruca daurica. *Insect Science, 30* (2), 279-292.

Enriquez, T., & Teets, N. M. (2023). Lipid Metabolism in Response to Cold.

Forsberg, J., & Wiklund, C. (1988). Protandry in the green-veined white butterfly, Pieris napi L.(Lepidoptera; Pieridae). *Functional Ecology* , 81-88.

Friedlander, Chen, W., Adamidi, C., Maaskola, J., Einspanier, R., Knespel, S., & Rajewsky, N. (2008). Discovering microRNAs from deep sequencing data using miRDeep. *Nature biotechnology, 26* (4), 407-415.

Friedlander, Mackowiak, S. D., Li, N., Chen, W., & Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic acids research, 40* (1), 37-52.

Fruciano, C., Franchini, P., & Jones, J. C. (2021). Capturing the rapidly evolving study of adaptation. *Journal of Evolutionary Biology, 34* (6), 856-865.

Gawienowski, A. M., Kessler, L. J., Tan, B. S., & Yin, C.-M. (1987). Glucocorticoid action on the growth and development of insects.*Life sciences, 40* (17), 1725-1730.

Grossmann, S., Bauer, S., Robinson, P. N., & Vingron, M. (2007). Improved detection of overrepresentation of Gene-Ontology annotations with parent–child analysis. *Bioinformatics, 23* (22), 3024-3031.

Gudmunds, E., Wheat, C. W., Khila, A., & Husby, A. (2022). Functional genomic tools for emerging model species. *Trends in Ecology & Evolution* .

Guerra-Assuncao, J. A., & Enright, A. J. (2012). Large-scale analysis of microRNA evolution. *BMC genomics, 13* , 1-12.

He, K., Xiao, H., Sun, Y., Situ, G., Xi, Y., & Li, F. (2019). microRNA-14 as an efficient suppressor to switch off ecdysone production after ecdysis in insects. *RNA biology, 16* (9), 1313-1325.

Ivanovic, J. (2018). *Hormones and metabolism in insect stress* : CRC Press.

Jain, S., Rana, V., Tridibes, A., Sunil, S., & Bhatnagar, R. K. (2015). Dynamic expression of miRNAs across immature and adult stages of the malaria mosquito Anopheles stephensi. *Parasites & vectors, 8* (1), 1-20.

Jin, X., Wu, X., Zhou, L., He, T., Yin, Q., & Liu, S. (2020). 20-Hydroxyecdysone-responsive microRNAs of insects. *RNA biology, 17* (10), 1454-1471.

Kang, W., Eldfjell, Y., Fromm, B., Estivill, X., Biryukova, I., & Friedlander, M. R. (2018). miRTrace reveals the organismal origins of microRNA sequencing data. *Genome biology, 19* (1), 1-15.

Kerr, G., Ruskin, H. J., Crane, M., & Doolan, P. (2008). Techniques for clustering gene expression data. *Computers in biology and medicine, 38* (3), 283-293.

Koštál, V. (2006). Eco-physiological phases of insect diapause.*Journal of insect physiology, 52* (2), 113-127.

Koštál, V., Šimůnková, P., Kobelková, A., & Shimada, K. (2009). Cell cycle arrest as a hallmark of insect diapause: changes in gene transcription during diapause induction in the drosophilid fly, Chymomyza costata. *Insect biochemistry and molecular biology, 39* (12), 875-883.

Koštál, V., Štětina, T., Poupardin, R., Korbelová, J., & Bruce, A. W. (2017). Conceptual framework of the eco-physiological phases of insect diapause development justified by transcriptomic profiling.*Proceedings of the National Academy of Sciences, 114* (32), 8532-8537.

Koštál, V. r., Berková, P., & Šimek, P. (2003). Remodelling of membrane phospholipids during transition to diapause and cold-acclimation in the larvae of Chymomyza costata (Drosophilidae). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 135* (3), 407-419.

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucleic acids research, 47* (D1), D155-D162.

Kumar, L., & Futschik, M. E. (2007). Mfuzz: a software package for soft clustering of microarray data. *Bioinformation, 2* (1), 5.

Lees, E., & Archer, D. (1980). Diapause in various populations of Pieris napi L. from different parts of the British Isles [Ecology, genetics, photoperiodism]. *Journal of Research on the Lepidoptera* .

Lehmann, P., Pruisscher, P., Koštál, V., Moos, M., Šimek, P., Nylin, S., . . . Wheat, C. W. (2018). Metabolome dynamics of diapause in the butterfly Pieris napi: distinguishing maintenance, termination and post-diapause phases. *Journal of Experimental Biology, 221* (2), jeb169508.

Lehmann, P., Pruisscher, P., Posledovich, D., Carlsson, M., Käkelä, R., Tang, P., . . . Gotthard, K. (2016). Energy and lipid metabolism during direct and diapause development in a pierid butterfly. *Journal of Experimental Biology, 219* (19), 3049-3060.

Lehmann, P., Van Der Bijl, W., Nylin, S., Wheat, C. W., & Gotthard, K. (2017). Timing of diapause termination in relation to variation in winter climate. *Physiological Entomology, 42* (3), 232-238.

Liu, Z., Ling, L., Xu, J., Zeng, B., Huang, Y., Shang, P., & Tan, A. (2018). MicroRNA-14 regulates larval development time in Bombyx mori.*Insect biochemistry and molecular biology, 93* , 57-65.

Lohse, K., Hayward, A., Ebdon, S., of Life, W. S. I. T., & Consortium, D. T. o. L. (2021). The genome sequences of the male and female green-veined white, Pieris napi (Linnaeus, 1758). *Wellcome Open Research, 6* .

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology, 15* (12), 1-21.

Lozano, J., Montañez, R., & Belles, X. (2015). MiR-2 family regulates insect metamorphosis by controlling the juvenile hormone signaling pathway. *Proceedings of the National Academy of Sciences, 112* (12), 3740-3745.

Lucas, K. J., Zhao, B., Liu, S., & Raikhel, A. S. (2015). Regulation of physiological processes by microRNAs in insects. *Current opinion in insect science, 11* , 1-7.

Marco, A., Hooks, K., & Griffiths-Jones, S. (2012). Evolution and function of the extended miR-2 microRNA family. *RNA biology, 9* (3), 242-248.

Meuti, M. E., Bautista-Jimenez, R., & Reynolds, J. A. (2018). Evidence that microRNAs are part of the molecular toolkit regulating adult reproductive diapause in the mosquito, Culex pipiens. *PLoS One, 13* (11), e0203015.

Olena, A. F., & Patton, J. G. (2010). Genomic organization of microRNAs. *Journal of cellular physiology, 222* (3), 540-545.

Pruisscher, P., Lehmann, P., Nylin, S., Gotthard, K., & Wheat, C. W. (2022). Extensive transcriptomic profiling of pupal diapause in a butterfly reveals a dynamic phenotype. *Molecular ecology, 31* (4), 1269-1280.

Pruisscher, P., Nylin, S., Wheat, C. W., & Gotthard, K. (2021). A region of the sex chromosome associated with population differences in diapause induction contains highly divergent alleles at clock genes. *Evolution, 75* (2), 490-500.

Quah, S., Hui, J. H., & Holland, P. W. (2015). A burst of miRNA innovation in the early evolution of butterflies and moths. *Molecular biology and evolution, 32* (5), 1161-1174.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* (26), 841-842.

R Core Team, R. (2021). R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2020. In.

Ragland, G. J., Egan, S. P., Feder, J. L., Berlocher, S. H., & Hahn, D. A. (2011). Developmental trajectories of gene expression reveal candidates for diapause termination: a key life-history transition in the apple maggot fly Rhagoletis pomonella. *Journal of Experimental Biology, 214* (23), 3948-3960.

Reynolds, J. A. (2019). Noncoding RNA regulation of dormant states in evolutionarily diverse animals. *The Biological Bulletin, 237* (2), 192-209.

Reynolds, J. A., Nachman, R. J., & Denlinger, D. L. (2019). Distinct microRNA and mRNA responses elicited by ecdysone, diapause hormone and a diapause hormone analog at diapause termination in pupae of the corn earworm, Helicoverpa zea. *General and comparative endocrinology, 278* , 68-78.

Reynolds, J. A., Peyton, J. T., & Denlinger, D. L. (2017). Changes in microRNA abundance may regulate diapause in the flesh fly, Sarcophaga bullata. *Insect biochemistry and molecular biology, 84* , 1-14.

Rinehart, J. P., Cikra-Ireland, R. A., Flannagan, R. D., & Denlinger, D. L. (2001). Expression of ecdysone receptor is unaffected by pupal diapause in the flesh fly, Sarcophaga crassipalpis, while its dimerization partner, USP, is downregulated. *Journal of insect physiology, 47* (8), 915-921.

Ritchie, W., Flamant, S., & Rasko, J. E. (2009). Predicting microRNA targets and functions: traps for the unwary. *Nature methods, 6* (6), 397-398.

Robinson, M., McCarthy, D., & Bioinformatics, G. (2020). undefined. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. academic. In: oup. com. https://academic. oup. com/bioinformatics/article-abstract/26/1 ....

Rodríguez del Río, Á., Giner-Lamia, J., Cantalapiedra, C. P., Botas, J., Deng, Z., Hernández-Plaza, A., . . . Bork, P. (2022). Functional and evolutionary significance of unknown genes from uncultivated taxa. *bioRxiv* , 2022.2001. 2026.477801.

Rorbach, G., Unold, O., & Konopka, B. M. (2018). Distinguishing mirtrons from canonical miRNAs with data exploration and machine learning methods. *Scientific reports, 8* (1), 1-13.

Schnall-Levin, M., Zhao, Y., Perrimon, N., & Berger, B. (2010). Conserved microRNA targeting in Drosophila is as widespread in coding regions as in 3' UTRs. *Proceedings of the National Academy of Sciences, 107* (36), 15751-15756.

Steward, R. A., Pruisscher, P., Roberts, K. T., & Wheat, C. W. (*in review*). Genetic constraints in genes exhibiting splicing plasticity in facultative diapause

Su, Z., Wilson, B., Kumar, P., & Dutta, A. (2020). Noncanonical roles of tRNAs: tRNA fragments and beyond. *Annual review of genetics, 54* , 47-69.

Suess, P., Dircksen, H., Roberts, K. T., Gotthard, K., Nassel, D. R., Wheat, C. W., . . . Lehmann, P. (2022). Time-and temperature-dependent dynamics of prothoracicotropic hormone and ecdysone sensitivity co-regulate pupal diapause in the green-veined white butterfly Pieris napi. *Insect biochemistry and molecular biology, 149* , 103833.

Thatcher, E. J., Bond, J., Paydar, I., & Patton, J. G. (2008). Genomic organization of zebrafish microRNAs. *BMC genomics, 9* , 1-9.

Wheat, C. W., Steward, R. A., Okamura, Y., Vogel, H., Lehmann, P., & Roberts, K. T. (*in review*). Functional coherence among miRNA targets: a potential metric for assessing biological signal among target prediction methods in non-model species

Wienholds, E., & Plasterk, R. H. (2005). MicroRNA function in animal development. *FEBS letters, 579* (26), 5911-5922.

Williams, C. M. (1952). Physiology of insect diapause. IV. The brain and prothoracic glands as an endocrine system in the Cecropia silkworm. *The Biological Bulletin, 103* (1), 120-138.

Wilsterman, K., Ballinger, M. A., & Williams, C. M. (2021). A unifying, eco-physiological framework for animal dormancy. *Functional Ecology, 35* (1), 11-31.

Yates, L. A., Norbury, C. J., & Gilbert, R. J. (2013). The long and short of microRNA. *Cell, 153* (3), 516-519.

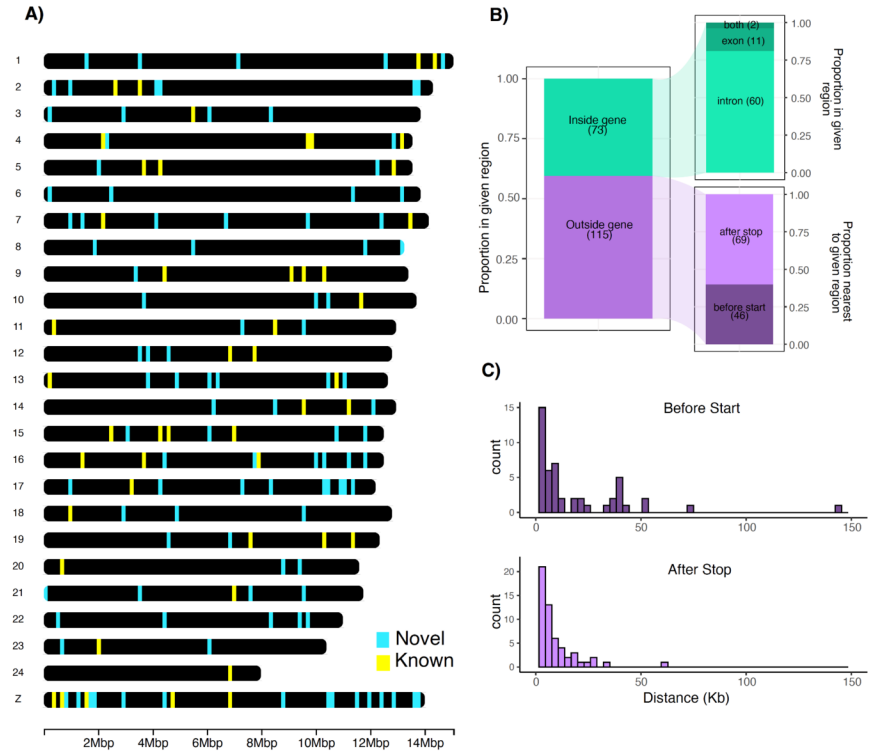

Figure 1) Locations of identified miRNA in *Pieris napi*.A) Chromosomal positions of all identified miRNAs in the *P. napi*genome, identified as novel (cyan) and known (yellow). B) Location of each miRNA relative to the nearest coding sequence (CDS). C) Distance from the nearest CDS for each miRNA that is located outside of a CDS, for miRNAs located before the start codon (top) and after the stop codon (bottom).

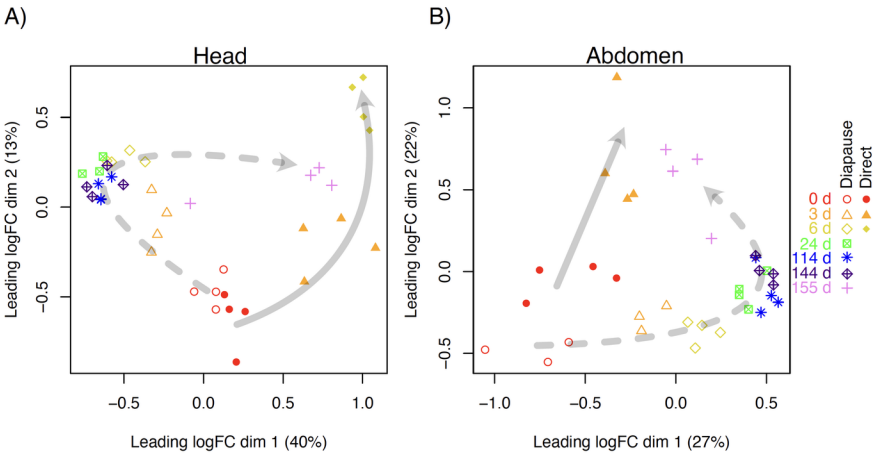

Figure 2) Global expression patterns of miRNA in both head and abdomen tissues throughout direct development and diapause. Multidimensional scaling vales for global miRNA expression for all samples used in this study, for head (left) and abdomen (right) tissues. Grey arrows show general direction of direct developing pupae (solid) and diapausing pupae (dashed).

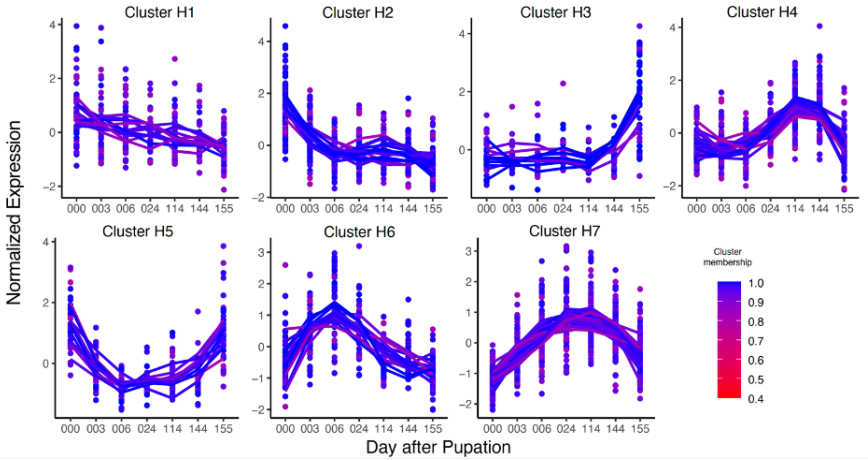

Figure 3) Clustered expression of differentially expressed miRNAs in head tissue diapausing of *P. napi* pupae. Clusters identified by fuzzy clustering with a minimum cluster membership of 0.8. Lines show median expression for each miRNA, and color corresponds to cluster membership.

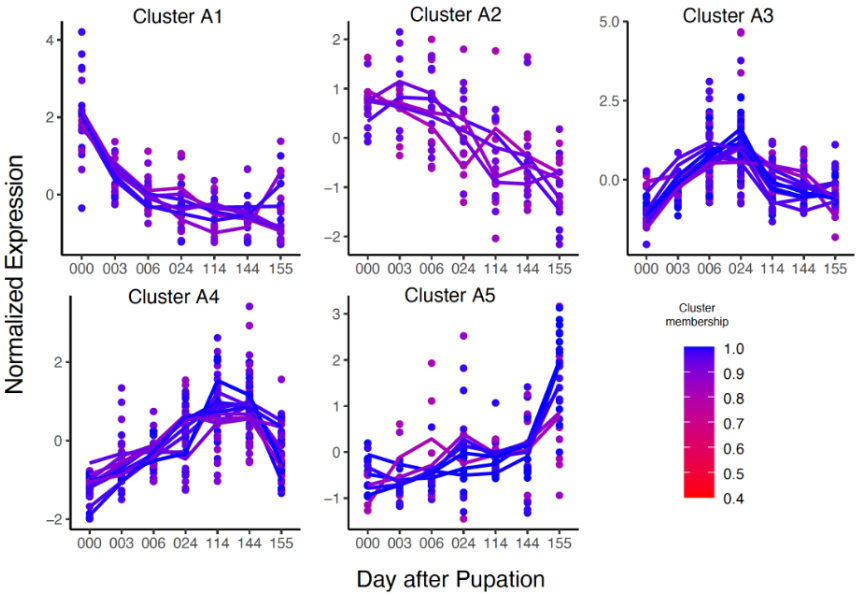

Figure 4) Clustered expression of differentially expressed miRNAs in abdomen tissue diapausing of *P. napi* pupae. Clusters identified by fuzzy clustering with a minimum cluster membership of 0.8. Lines represent median normalized expression for each miRNA, and color corresponds to cluster membership.

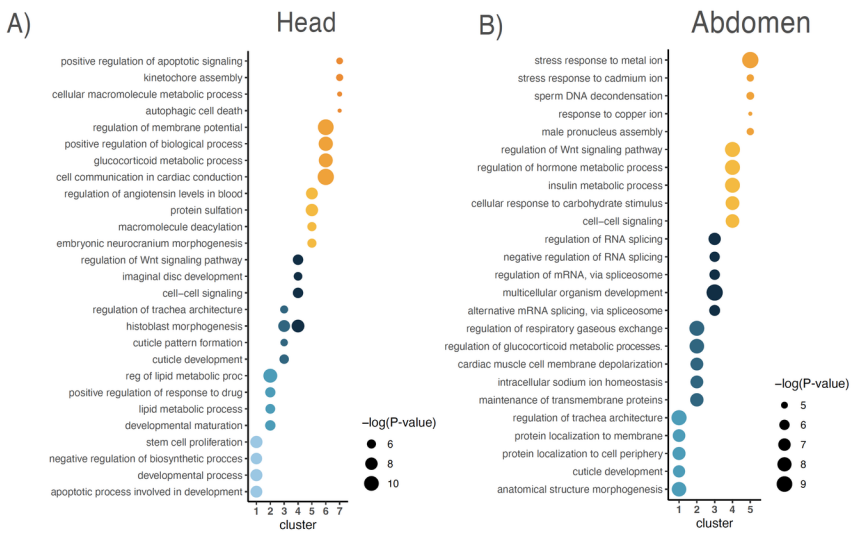

Figure 5) Gene set enrichment of genes targeted by multiple miRNAs in each cluster. A) The top 4 most enriched GO terms from each cluster in the head tissue. B) The top 5 most enriched GO terms from each cluster in the abdomen tissue. Color corresponds to cluster, and size of circle is related to the negative log P-value. To see top ten most enriched GO terms for each cluster available in Table S6.

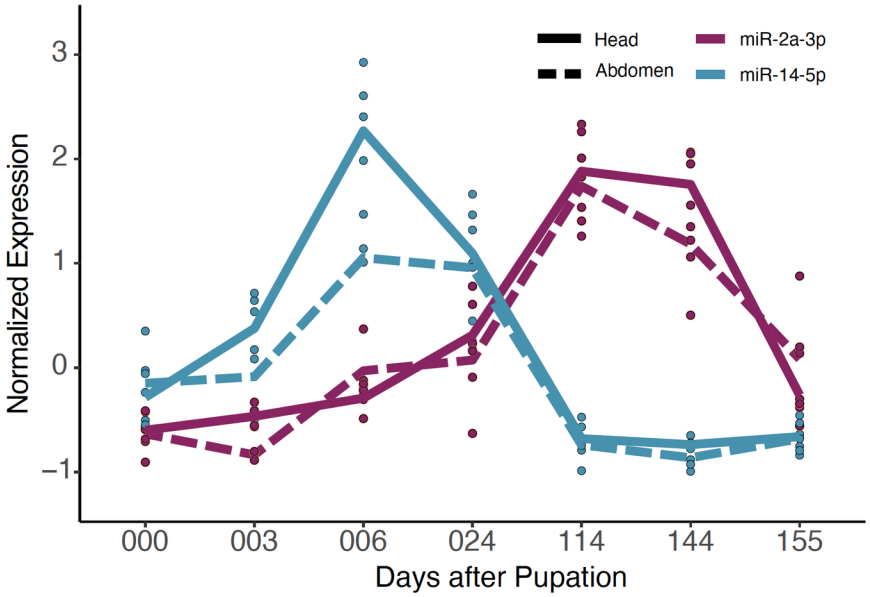

Figure 6) Expression patterns of miRNA that were differently expressed during diapause termination in both the abdomen and head. Lines represent median normalized expression of miR-2a-3p and miR-14-5p in the head (solid) and abdomen (dashed) tissues of diapausing*Pieris napi* pupae.

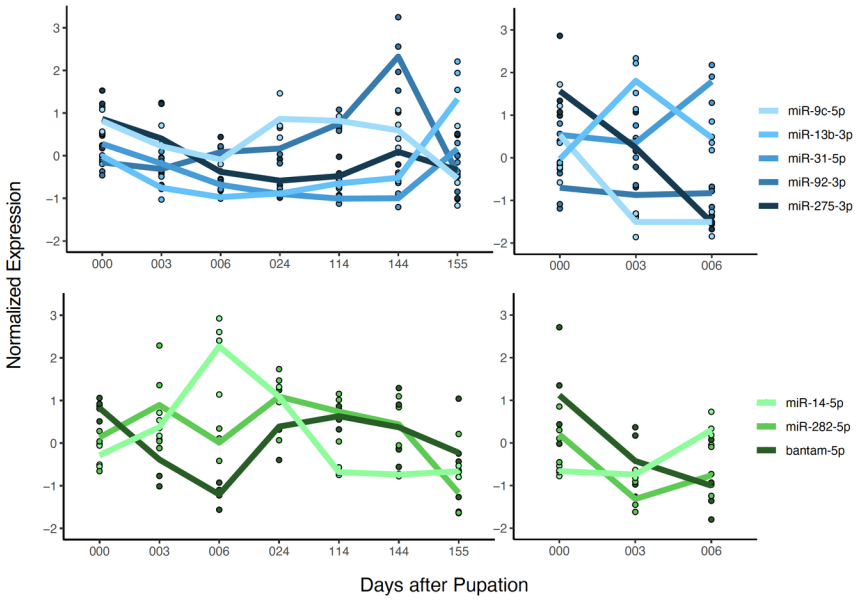

Figure 7) Expression of differentially expressed miRNAs identified in previous studies as important in diapause. Lines represent median normalized expression of differentially expressed miRNAs in diapause and direct *Pieris napi* pupae that overlapped with Reynolds et al. 2017 (top, blue) and Batz et al. 2017 (bottom, green).