Journal of Insect Physiology 137 (2022) 104361

Contents lists available at ScienceDirect

Journal of Insect Physiology

journal homepage: www.elsevier.com/locate/jinsphys

Life stage and the environment as effectors of transposable element activity in two bee species

Sarah Signor [a,*], George Yocum [b], Julia Bowsher [a]

[a] North Dakota State University, Department of Biological Sciences, United States
[b] Bioscience Research Laboratory, U.S. Department of Agriculture/Agricultural Research Service, United States

ARTICLE INFO

Keywords:
Diapause
Transposable elements
Stress
Osmia
Megachile

ABSTRACT

Diapause is a complex physiological phenomenon that allows insects to weather stressful environmental conditions. The regulation of diapause is accordingly complex, including signaling pathways that involve both small RNA and mRNA and affect the cell cycle, stress resistance, and developmental timing. Transposable elements, mobile genetic elements that replicate within the genome, are also thought to be stress responsive and regulated by the small RNA pathway. Therefore, we asked what the relationship was between environmental stress, diapause status, and transposable element expression in two species of agriculturally important bees, *Megachile rotundata* and *Osmia lignaria*. We characterized the TE content of the genomes of both species, then evaluated the expression of TE families during temperature stress, general environmental stress, and diapause stage. We found that the genomic TE content of the two species was very different, and *M. rotundata* has a larger number of annotated TEs compared to *O. lignaria*. We also found that both diapause stage and temperature stress had large effects on TE expression. The fold change of TE famlies tended to be larger in those expressed during diapause, however there was only a small majority that were upregulated during diapause. This suggests that stress and diapause do not break down to a simple up-regulation or down-regulation of TEs, but rather that the TE family, the genomic position of its insertions, and the exact heterochromatin formation of the organism at any given environmental state or life stage may affect overall expression of TEs.

1. Introduction

Insects have colonized nearly every available terrestrial habitat from the tropics to the polar regions, including regions with low temperatures that could negatively impact insect survival and reproduction. Diapause is a key physiological adaptation that enables insects to survive harsh environmental conditions and to synchronize with biotic and abiotic cues required for development and reproduction. During diapause, metabolism and developmental progression decrease while stress tolerance increases (Tauber et al., 1986; Danks, 1987). Diapause is not merely a stopping or slowing of development in response to current harsh conditions, but rather a complex physiological phenomenon involving three phases: anticipation and preparation for dormancy, the dormant phase, and a final phase that is primed for a return to developmental progression (Koštál, 2006; Koštál et al., 2017). In the pre-diapause phase, the insect perceives environmental cues that foretell winter and trigger behavioral and physiological mechanisms preparing the insect to enter diapause. The second stage is the "true diapause" phase that is characterized by endogenously blocked morphogenesis. The final phase of diapause development is post-diapause quiescence, in which metabolism is still repressed but the insect will exit dormancy in response to environmental cues. Each of the phases of diapause requires a complex regulation of gene expression to bring about the required changes in physiology.

Our understanding of the complexity of the molecular regulation of diapause has expanded greatly since the development of RNA-seq technology (Koštál et al., 2017; Emerson et al., 2010; Ragland et al., 2010; Ragland et al., 2011; Poelchau et al., 2011; Poelchau et al., 2013a; Poelchau et al., 2013b; Gong et al., 2013; Dong et al., 2014; Huang et al., 2015; Qi et al., 2015; Yocum et al., 2015; Yocum et al., 2018; Hao et al., 2016; Kang et al., 2016; Meyers et al., 2016). Recently, all three small RNA pathways have also been implicated in different aspects of diapause regulation (Reynolds et al., 2017) (microRNAs, siRNAs, and piRNAs). For example, in *M. rotundata* proteins involved in the siRNA pathway are differentially regulated between diapause stages (Yocum et al., 2015). The proteins in the piRNA and siRNA pathway are both differentially

* Corresponding author.
E-mail address: sarah.signor@ndsu.edu (S. Signor).

https://doi.org/10.1016/j.jinsphys.2022.104361
Received 21 April 2021; Received in revised form 12 January 2022; Accepted 13 January 2022
Available online 19 January 2022

regulated in prediapause stages of *Sarcophaga bullata* and *Chymomyza costata*, potentially because diapause calls for changes or increases in heterochromatin (Reynolds et al., 2017; Poupardin et al., 2015; Reynolds et al., 2013; Taliaferro et al., 2013). TEs in turn make up large fractions of eukaryotic genomes and are able to replicate by transposing to new locations of the hosts' genome. These two major classes of genomic parasite, retrotransposons (which transpose with an RNA intermediate) and DNA transposons (which transpose with a DNA intermediate), are therefore likely also linked through the piRNA and siRNA pathways to the epigenetic changes in the genome such as diapause that are also influenced by the environment.

TE activity has also long been associated with stress, for example TEs interact with stress-specific transcription factors and change the expression of genes in response to stress (Horváth et al., 2017; Villanueva-Cañas et al., 2019) . TEs are also induced by various environmental stressors. TEs can carry regulatory elements that alter the transcription and splicing of nearby genes, and in at least one example, a TE carried a stress responsive regulatory element which induced expression of neighboring genes following environmental stress (Horváth et al., 2017; Villanueva-Cañas et al., 2019; Bourque et al., 2008; Leem et al., 2008; Faulkner et al., 2009; Feng et al., 2013). In other cases TEs are simply de-repressed in response to stress because the regulatory apparatus that normally silences them is re-routed for other more pressing tasks (Rinehart et al., 2007). There is considerable evidence in the literature for TE mobilization under stress, in some cases including the mechanistic basis, but it is still a complicated area of research with many open questions (Van Meter et al., 2014; Voronova et al., 2014; Romero-Soriano et al., 2016; Ryan et al., 2016; Zovoilis et al., 2016; Hummel et al., 2017).

Considering this information in aggregate – that diapause is a stress resistant phenotype, its regulation involves small RNA pathways, small RNAs regulate stress resistance, and TEs are both stress responsive and silenced by the small RNA pathway – it suggests a potential interaction between stress, diapause, and TE activity with potentially profound effects on organismal health. In addition, buried in the many massive RNA-seq datasets that exist on diapause, there is evidence for the differential regulation of TEs during the course of diapause, which allows us to approach these questions (Yocum et al., 2015; Yocum et al., 2018; Rinehart et al., 2007). We wanted to ask the following – how does TE expression differ between diapausing and non-diapausing individuals, how does it change over the course of diapause as an organism transitions between diapause stages, and how does temperature stress affect TE expression during diapause?

To address these questions, we characterized the expression of TE families during diapause development in two solitary bees, the alfalfa leafcutting bee, *M. rotundata* and the blue orchard bee, *O. lignaria* (Yocum et al., 2015; Yocum et al., 2018). These bees are both agriculturally important pollinators. Both species undergo diapause, but they do so in different life stages. *M. rotundata* facultatively diapauses in the pre-pupal stage (Stephen and Osgood, 1965; Klostermeyer and Rank, 1982; Sgolastra et al., 2012) . *O. lignaria* has an obligate diapause as an adult, and has an additional pre-pupal dormancy that is very similar to diapause (Sgolastra et al., 2012; Torchio, 1989). The diapause stage is important for agricultural management of these species (Yocum et al., 2015; Yocum et al., 2018; Torson et al., 2015; Torson et al., 2017). We compared the following to address our questions about the regulation of TEs, stress, and diapause: 1) TE expression in *M. rotundata* during each of the three phases of diapause: maintenance, termination, and post-diapause quiescence 2) lab and field raised *M. rotundata* to determine if the presumably more stressful field conditions resulted in differences in TE expression 3) TE expression in *M. rotundata* during post-diapause quiescence under cold stress and with reduced cold stress 4) TE expression between diapausing adults and dormant prepupae in *O. lignaria* to compare patterns of TE expression in different types of dormancy 5) diapausing and non-diapausing adults of *O. lignaria* to compare patterns of TE expression relative to diapause in more than one species.

2. Methods

2.1. Genome assemblies and TE annotation

The *M. rotundata* genome assembly is available from NCBI and was used for genomic mapping and TE identification (NCBI:txid143995). The *O. lignaria* genome was acquired from the USDA in September of 2020 (https://i5k.nal.usda.gov/content/osmia-lignaria-genome-assembly-usdaolig10-gcf0122742951). The TE library for both genomes was constructed using the Extensive de-novo TE Annotator pipeline (EDTA) (Ou et al., 2019). These annotations are available at (redacted for blind peer review). This pipeline is intended to create a non-redundant TE library based off of a reference genome. We chose only to differentiate between general DNA transposons, MITEs, *Helitrons*, and three groups of LTR transposons – *copia*, *gypsy* and unknown– because previous benchmarking suggests EDTA sub-designations between types of DNA transposons are unreliable (Tiedeman, 2020). The TEs were compared to existing annotations using the Dfam database (Hubley et al., 2016) and hmmer (hmmer.org).

2.2. Identifying TE insertions

We ran RepeatMasker on the *M. rotundata* and *O. lignaria* genomes using the TE libraries from EDTA as input (Tarailo-Graovac and Chen, 2009). Simple repeats were excluded, along with any hit with a Smith-Waterman alignment score of less than 225 (Smith and Waterman, 1981). Repeatmasker may annotate single copies of TEs more than once if they are fragmented or contain regions of significant divergence, therefore this is an approximation of potential insertions.

RNA-seq data

2.3. Diapause phase and laboratory conditions in M. rotundata

The protocols for collecting bees were described previously, and will be briefly summarized (Yocum et al., 2018). The M. rotundata originated from broods reared in commercial alfalfa fields in Wyoming, USA; the bees were then purchased by a Utah alfalfa seed grower for pollination in summer (41°47′37.04″N; 112°8′18.35″W). Bees were maintained outdoors from September 2010 until 22 October 2010. On that date, bees were divided into two overwintering management groups: 1) lab bees, overwintered at constant 6 °C and 2), and field bees, overwintered outside at ambient temperatures. Prepupal bees were sampled in November for diapause maintenance, January for diapause termination, and either March or May for post-diapause quiescence. 48 samples were sequenced: 12 for each period of diapause, and 24 each lab or field raised. Illumina sequencing was carried out by the University of Georgia Genomic Facility. Stranded Illumina libraries were constructed using the TruSeq RNA kit (Illumina Inc., San Diego, CA, USA) and sequencing was performed using an Illumina HiSeq 2000 sequencer.

2.4. M. rotundata temperatures during post-diapause quiescence

The protocol for the fluctuating temperature regime and RNA-seq collection are described elsewhere, and will be briefly summarized here(Torchio and Bosch, 1992). Bees were purchased from JWM Leafcutter Inc. and were of Canadian origin. Experiments were conducted at the USDA-ARS in Fargo, ND in 2010. Diapausing bees were stored for seven months at constant temperature (6 °C ± 0.5 °C with a 15 h:9h (L: D) photoperiod). After seven months they were split into two treatments, static temperature (STR) and fluctuating temperature (FTR). Bees reared under FTR were exposed to 6 °C with a daily warm pulse of 20 °C for an additional 5–7 months. Total RNA was collected from six post-diapause quiescent prepupae reared under either STR or FTR using the Invitrogen TRIzol protocol (Carlsbad, CA, USA). Individuals for both

treatments were harvested during the cold phase (6 °C). Based on overall mortality rates those in the STR treatment regime are subject to more stressful conditions and are more susceptible to chill injury than those with daily warm pulses (Torson et al., 2015; Torson et al., 2017). 6 RNA-seq libraries for each treatment were prepared using Illumina TruSeq mRNA standard protocol at the University of Georgia Genomics Facility and sequencing was performed by the University of Missouri Columbia DNA Sequencing Core on a HiSeq2000.

2.5. *O. lignaria* diapausing adults and prepupae

In the spring of 2003, an active nesting population of *O. lignaria* was maintained in an apple orchard (Hanson's farm, Logan, UT). Nesting blocks were inspected daily and capped nests were collected and transferred to the USDA-ARS, Logan, UT. The nests were dissected and individual larvae were placed into clay blocks (Torchio and Bosch, 1992). The blocks were kept in a plastic box with extra clay blocks filled with water to maintain humidity. The boxes with the developing bees were kept in an unheated barn through the autumn–winter-spring of 2003/2004. Once the larvae completed cocooning they were placed into gel capsules and arranged onto sticky boards. The bees were x-rayed every 3 days to monitor for pupation and adult eclosion (Kemp et al., 2004). The dates of pupation and adult eclosion were recorded for each bee. Samples were collected starting with fifth instar larvae, through pupal development, adult diapause (autumn–winter), and finally adults emerging in the spring of 2004 (Bosch et al., 2008). 11 non-diapausing prepupal samples were collected for sequencing, 15 diapausing adult samples, and 3 non-diapausing adults.

2.6. RNA-seq mapping and differential expression of TEs

For each data set reads were mapped to the TE library using bwa mem version 0.7.15 (Li, 2013) and processed with samtools v.1.9 (Li, 2013; Li et al., 2009). Read counts for each TE family were obtained, and TEs were designated as not expressed according to the following criteria: No more than 20% of the population have no more than 2 reads mapping to the TE. DEseq2 was used to evaluate differential expression between treatments (Love et al., 2014). *p*-values for all tests were adjusted for multiple testing (Benjamini and Hochberg, 1995). For a comparison to another method please see Supplementary File 1.

3. Results

3.1. TEs in M. rotundata

We identified 851 distinct TE families in *M. rotundata*. The vast majority of these are DNA transposons: 529 DNA transposons, 172 *MITE*s, and 47 *Helitrons*. The remainder consist of 45 *copia* elements, 51 *gypsy* elements, and 7 unknown LTRs (Fig. 1A). 96 of these TEs could be associated with a known element in the Dfam database, predominantly LTRs, suggesting that they have spread more recently within the population than DNA transposons – 12% of annotated TE families in the species and 53% of those with a match in Dfam (Supp. File 2). LTRs are thought to be a younger group of TEs compared to DNA transposons, though a formal analysis of TE age in these species would be a valuable contribution, as similarity to Dfam could depend only on the particular species previously analyzed in the database (Tiedeman, 2020; Bowen and McDonald, 2001; Signor, 2020) . Notably, 103 LTR elements are annotated overall, and 24 of these are related to *copia* superfamily elements from *D. melanogaster*. This suggests specifically that the *copia* superfamily has spread more recently into *M. rotundata* from dipterans or their relatives.

3.2. TEs in O. lignaria

We identified 738 TE families in *O. lignaria*. This includes 251 DNA transposons, 149 *MITE*s and 63 *Helitrons*. The remainder consists of 67 *copia* elements, 200 *gypsy* elements, and 7 unknown LTR elements (Fig. 1A). 165 of these TEs could be tied to existing Dfam annotations, likely due to the larger presence of recent *gypsy* and *copia* superfamily invasions, as only 15% of those with a putative related TE in Dfam are DNA transposons. Of these LTR transposons, elements related to *Bel*, *Max-Element*, *roo*, and *Quasimodo* are private to *O. lignaria* compared to *M. rotundata*. Both species contain *ZAM*-related elements, but in *O. lignaria* this is a particularly abundant subset of TEs (16 elements) (Supp. File 3).

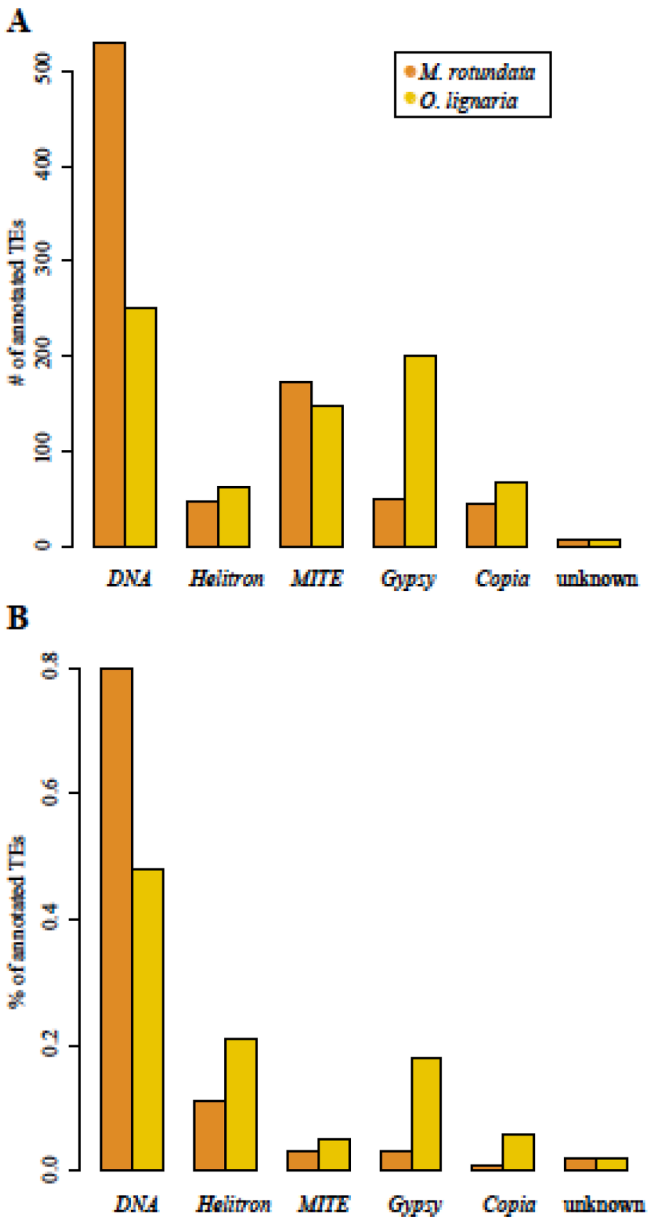

Fig. 1. A. The number of TEs in each category - DNA transposons, *Helitrons*, *Copia*, *Gypsy*, and unknown LTR – that were identified in *Megachile rotundata* and *Osmia lignaria*. Many more *gypsy* elements were identified in *O. lignaria* compared to *M. rotundata*. B. The percentage of identified insertions that belong to each category of TEs.

3.2.1. TE content of the M. rotundata and O. lignaria genomes

Repeatmasker identified 198,434 fragments or potential TE insertions in TEs in *M. rotundata* (Supp. File 4). This is likely an underestimate as the *M. rotundata* genome is somewhat fragmented. In *O. lignaria,* we identified 87,347 TE fragments or potential insertions

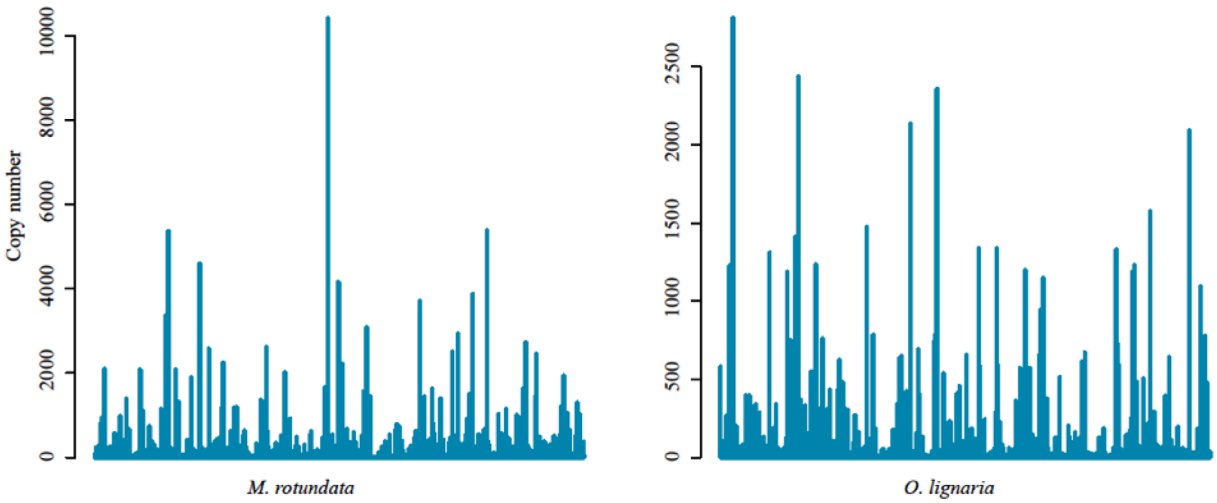

Fig. 2. TE fragments or potential TE insertions as determined by Repeatmasker. While in *O. lignaria* (right) all TEs have 2500 copies or fewer, in *M. rotundata* 9 TEs have >3000 insertions. One TE has 2x the number of insertions of its next closest TE (5388 insertions versus 10433). The TEs are in order by number along the X axis.

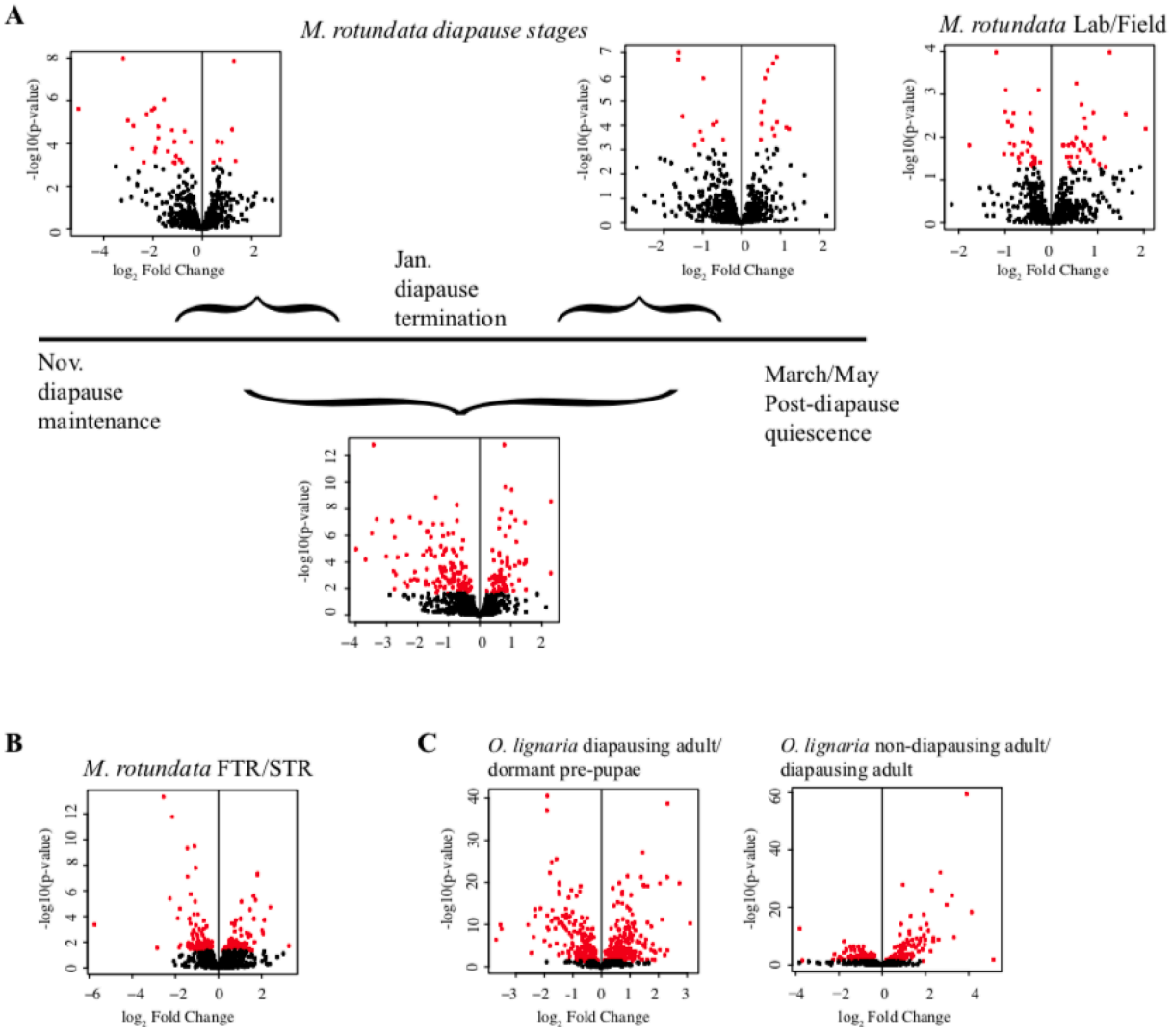

Fig. 3. TE expression was plotted by fold change (log_2) in expression and the adjusted *p*-value for each comparison. Red points indicate transcripts that have significantly different expression between samples. Along the x-axis, a negative value indicates higher expression in the first treatment listed. A. The distribution of *p*-values for pairwise comparisons of *M. rotundata* diapause stages as well as the Lab/Field comparison. B. The distribution of *p*-values for differences in TE expression in *M. rotundanta* under fluctuating (FTR) and static (STR) temperature regimes. C. The distribution of *p*-values in *O. lignaria* for TE expression differences between diapausing adults and dormant prepupae, followed by *O. lignaria* diapausing adults versus non-diapausing adults. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Supp. File 4). 16.98% of the *M. rotundata* genome is estimated to be TEs, while only 11.61% in *O. lignaria*. The *M. rotundata* genome is also larger overall (272 MB in *M. rotundata*, versus 177.2 MB in *O. lignaria*), thus this represents a larger number of total base pairs as well. Therefore *M. rotundata* bears a considerably larger number of potential TE insertions within its genome than *O. lignaria*.

3.2.2. Copy number of TEs in the genome of M. rotundata and O. lignaria

Copy number, or number of potential insertions/fragments, is variable between TEs in *M. rotundata*, with 80 elements having as few as 1 copy and 1 element having as many as 10,000 insertions (Fig. 2, TE 403). It is possible that older TEs are more fragmented and appear as more 'copies'. In contrast, in *O. lignaria* no TE has a copy number greater than 2500. This is in line with our previous results which found that *M. rotundata* contained more TE insertions than *O. lignaria*. Of the 1% of TEs with the highest copy number, in *M. rotundata* this includes 9 TEs with >3000 insertions that are all DNA transposons, only one of which can be tied to an existing Dfam element (TE 680, *CMC-EnSpm* DNA transposon). In *O. lignaria* this includes 6 TEs with >2000 insertions, 2 DNA transposons, 2 *Helitrons*, and a *copia* and a *gypsy* element.

3.2.3. Diapause phase and laboratory conditions in M. rotundata

Between diapause maintenance (Nov.) and diapause termination (Jan.) 31 TE families were differentially expressed (Supp. File 5). Between diapause termination (Jan.) and post-diapause quiescence (March/May) 24 TE families were differentially expressed (Supp. File 6). Between diapause maintenance (Nov.) and post-diapause quiescence (March/May) 202 TE families were differentially expressed (Supp. File 7; See Supp. File 6 for a PCA). Overlap between the sets of significant TE families is interesting – in the diapause maintenance/diapause termination comparison 29 of the 31 TE families were also differentially expressed between diapause maintenance/post-diapause quiescence (a significantly higher overlap than expected by chance, hypergeometric probability $p < 4.46 \times 10^{-17}$). In addition, between diapause termination/post-diapause quiescence 14 of the 24 TE families are also differentially expressed between diapause maintenance/post-diapause quiescence ($p < 0.006$), but there is 0 overlap between the two respective datasets. Therefore, transcribed TE families are distinct to particular periods of diapause, and also form part of a larger set of potentially transcribed TEs between the diapause maintenance and post-diapause quiescence stages. In addition, a majority of TE families are expressed at higher levels during diapause maintenance (Nov.) compared to post diapause quiescence (March/May, 62%), and they represent the largest fold changes. 59 TE families were differentially expressed between lab raised and field raised pre-pupae (Figs. 3-4, Supp. File 5; See Supp. File 6 for a PCA). They are equally split between increases and decreases in expression. If these differences were the result of a general deregulation of TEs between conditions, an increased copy

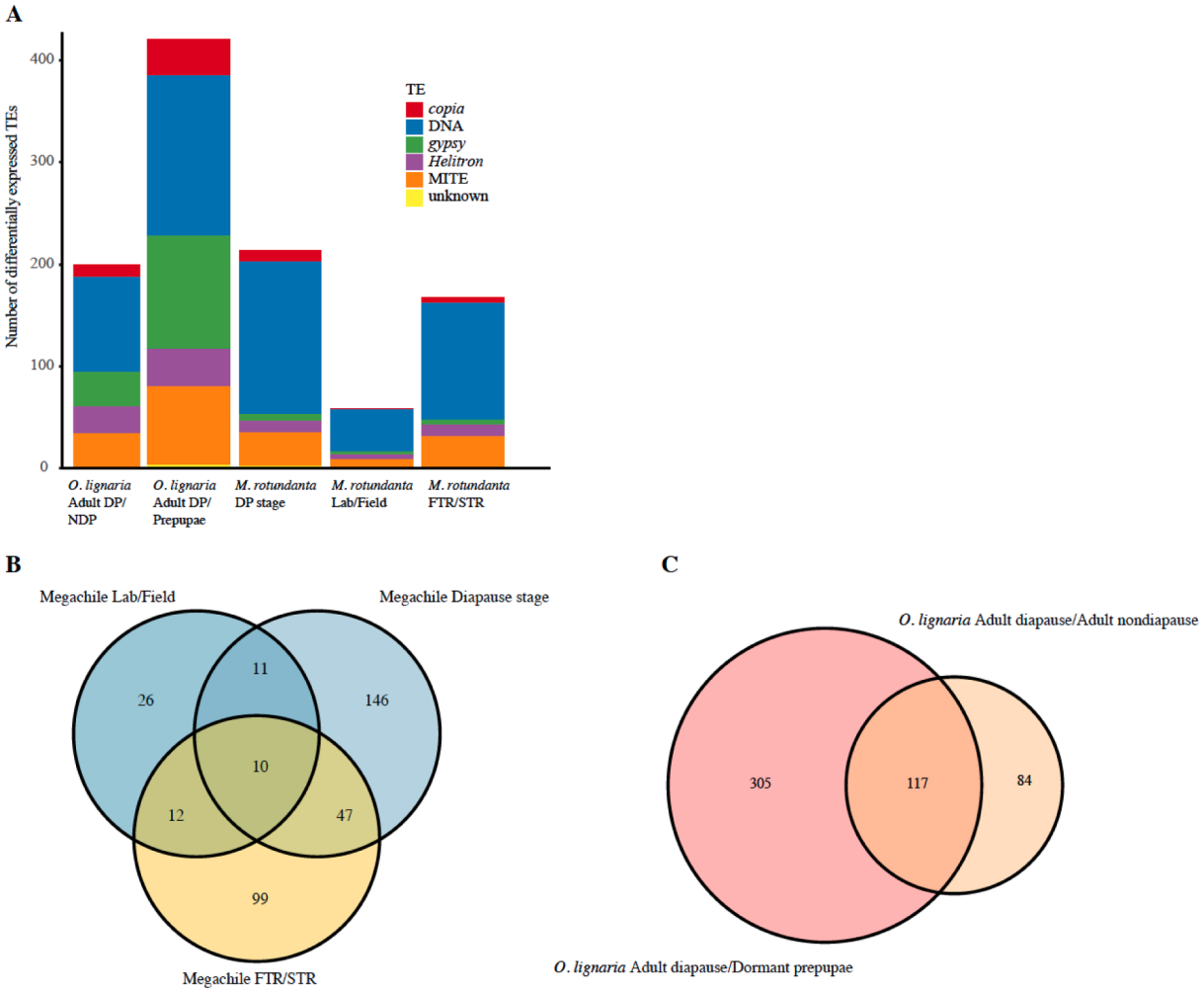

Fig. 4. A. The number of TEs from different taxonomic categories that were differentially expressed in each comparison between *O. lignaria* and *M. rotundata*. B. Overlapping and unique TEs from each dataset in *M. rotundata*. Because the Nov./Jan. and Jan./March/May diapause stages had largely overlapping significant TEs with the Nov./March/May comparison they were grouped into a single comparison. C. Overlapping TEs in *O. lignaria*, adult diapausing/prepupae dormant and adult diapausing/non-diapausing.

number of TEs could result in an increased likelihood of significant differential expression, however we do not see this relationship in this or the following analysis of differential TE expression (Supp. File 7).

3.3. *M. rotundata* temperatures during post-diapause quiescence

168 TE families responded significantly to different temperature regimes (Figs. 3-4, Supp. File 5; See Supp. File 6 for a PCA). The direction of expression change is not preferential to either treatment (52% increase in FTR, 48% increase in STR). The largest fold change in any treatment or species examined in this study is an increase in FTR (~6 fold change) in a DNA transposon (TE 621). This dataset shares 10 TEs with the diapause and lab/field datasets, suggesting that these TEs are particularly able to exploit changes in the environment and/or life stage of the organism (Fig. 4).

3.3.1. *Diapausing adult and dormant prepupae in O. lignaria*

422 TE families are differentially expressed between stages in *O. lignaria* (Supp. File 8; Figs. 3-4; See Supp. File 9 for a PCA). The largest fold change, and more frequently larger fold changes, are in those TE families that are upregulated in diapausing individuals compared to dormant individuals (the largest is ~ 3.7 fold change), though the direction of TE expression change between diapause and dormant individuals is approximately evenly split.

3.3.2. *Diapausing adult and non-diapausing adults in O. lignaria*

201 TE families are differentially expressed between diapausing adults and non-diapausing adults (Fig. 4, Supp. File 8; See Supp. File 9 for a PCA). This includes TE 17, a DNA transposon, which is a large outlier in *p*-value and whose log 2 fold change is 4 times higher in diapausing adults compared to non-diapausing adults. Overall, the direction of change is relatively evenly split with 89 TE families more highly expressed in non-diapausing adults and 111 TE families more highly expressed in diapausing adults (44/56%).

4. Discussion

In this study we found abundant differences in TE expression as a result of life stage, environment, and diapause status. There was no large overall bias in the direction of change of TE expression, though the largest fold changes were consistently upregulation in diapause. *M. rotundata* diapause stages – maintenance versus termination and termination versus post-diapause quiescence – share very few transcribed TEs. However, these comparisons include a subset of the significant TEs between diapause maintenance and post diapause quiescence. The most significant difference in terms of number of TE families was between *O. lignaria* dormant prepupae and diapausing adults. *O. lignaria* prepupae are thought to be in a summer dormancy stage that resembles diapause (Sgolastra et al., 2012). Our results show TE activity is very different during these two dormancies, and indeed finds differential expression of 57% of TE families. The magnitude of differential TE activity in *O. lignaria* is also interesting because a smaller percentage of its genome is occupied by TEs compared to *M. rotundata*, and the constituent TE families are present in much lower copy numbers.

Very few studies have investigated the role of different life stages in TE activity. In one study in *Drosophila montana*, 15 TEs were found to be differentially expressed between the diapause stages, 12 of which were upregulated during diapause (Kankare et al., 2016). In *Culex pipiens*, two TEs were upregulated during diapause (Robich et al., 2007). Diapause upregulation of two TEs in *Bombyx mori*, and one in *Sarcophaga crassipalpis*, has also been noted (Reynolds et al., 2017; Yamashita et al., 2001). Typically a small majority of TEs were upregulated during diapause, and the degree of increase was largest for TEs that were upregulated during diapause. Our study also finds many more TEs affected by diapause than these previous investigations, and recapitulates the tendency for TEs to be upregulated during diapause with larger fold changes.

TE transposition is suppressed by the activity of the piRNA and siRNA small RNA pathways. The siRNA pathway is well understood and involves degradation of mRNA transcripts through the interaction of double stranded RNA with the Dicer2 and Ago2 proteins. This is the predominant mode of somatic suppression of TEs. In the piRNA pathway TEs are transcribed and spliced into short RNAs termed piwi-RNAs (piRNA) (Brennecke et al., 2007). PIWI-clade proteins bind these piRNAs and suppress transposon activity transcriptionally and post-transcriptionally (Brennecke et al., 2007). Suppression can occur through many mechanisms, such as the formation of heterochromatin or cleaving of the transcripts of TEs. However, recent research has also found that piRNAs are important in somatic cells. For example, there is a twofold increase in transcript abundance of genes encoding components of the piRNA pathway in larva of *Sarcophaga bullata* that are destined for diapause, potentially because the piRNA pathway can direct the formation of additional heterochromatin prior to diapause (Reynolds et al., 2017; Reynolds et al., 2013). These two genes are also more abundant in *Dianemobius nigrofasciatus* that are programmed to enter diapause compared to embryos programmed for continuous development (Shimizu et al., 2018). In the previous RNA-seq analysis of the *M. rotundata* heat and diapause experiments no differences in expression were found in *Dicer2*; *Ago2*; *PIWI like protein*, or *Aubergine like protein*, though there are likely other as yet unannotated argonaute proteins in this species (Yocum et al., 2018; Yocum et al., 2015). Differences in the activity of TEs between different diapause stages could be due to either the side effects from the activity of the piRNA pathway in different contexts, changes in the location of piRNA induced heterochromatin through different life stages, or changes in heterochromatin induced by other genes or environmental factors in both of these spcies.

In diapause *M. rotundata* gene expression over the course of diapause development is highly dependent upon the environmental history of the individual bee (Yocum et al., 2018). The differential regulation of TEs between the field and laboratory samples strengthen the argument put forward by Yocum et al. (2018) that these two groups are physiologically distinct from each other. This could be due to differences in stress in the two environments, or stress independent differences in the effect of the environment. In all cases of stress activated TEs, a proportion of TEs are activated rather than general de-repression, suggesting that there is a specific mechanism involved (Dubin et al., 2018; Liu et al., 2021). For example, in *Arabidopsis*, loss of DNA methylation in response to heat enhances TE activation, and the effect is position specific (Liu et al., 2021; Kumar and Wigge, 2010; Tittel-Elmer et al., 2010; Cavrak et al., 2014). This could be explained by two mechanisms - the aforementioned changes in chromatin structure (Liu et al., 2021; Feschotte, 2008; Drongitis et al., 2019; Diehl et al., 2020) , or stress-specific transcription factors. TEs have been found previously to interact with stress-specific transcription factors, including providing binding sites (Villanueva-Cañas et al., 2019). Following stress, TEs are also associated with up or down regulation of neighboring genes (Horváth et al., 2017).

Two patterns are clear from this study 1) The largest fold changes in TE expression occur during diapause as an increase in expression. 2) The populations of TEs that respond to different environmental or developmental cues are relatively distinct, suggesting it may not be the TE itself that determines differences in expression, but rather genomic position and chromatin state or some other local determinant of TE expression likelihood.

5. Availability of data and materials

The data that support the findings of this study are openly available in the NCBI SRA at PRJNA528472, PRJNA400265, and PRJNA553801.

Funding

This work was supported by the National Science Foundation

Established Program to Stimulate Competitive Research NSF-EPSCoR-1826834 and NSF-EPSCoR-2032756.

CRediT authorship contribution statement

Sarah Signor: Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **George Yocum:** Resources, Supervision, Data curation, Writing – original draft, Writing – review & editing. **Julia Bowsher:** Project administration, Funding acquisition, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

S.S. would like to thank C. & S. Emery for insightful commentary on the manuscript. G.Y. and J. B. would like to thank B. Kemp for sample collection of *O. lignaria*.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jinsphys.2022.104361.

References

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc B 57 (1), 289–300.
Bosch, J., Sgolastra, F., Kemp, W.P., 2008. Life cycle ecophysiology of Osmia mason bees used as crop pollinators. In: James, R.R., Pitts-Singer, T.L. (Eds.), Bee Pollination in Agricultural Ecosystems. Oxford University Press, Oxford UK, pp. 83–104.
Bourque, G., Leong, B., Vega, V.B., Chen, X.i., Lee, Y.L., Srinivasan, K.G., Chew, J.-L., Ruan, Y., Wei, C.-L., Ng, H.H., Liu, E.T., 2008. Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res. 18 (11), 1752–1762.
Bowen, N.J., McDonald, J.F., 2001. Drosophila euchromatic LTR retrotransposons are much younger than the host species in which they reside. Genome Res. 11 (9), 1527–1540.
Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., Hannon, G. J., 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128 (6), 1089–1103.
Cavrak, V.V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L.M., Mittelsten Scheid, O., Kakutani, T., 2014. How a retrotransposon exploits the plant's heat stress response for its activation. PLoS Genet. 10 (1), e1004115.
Danks, H.V., 1987. Insect dormancy: An ecological perspective. Biological Survey of Canada monograph series. Biological Survey of Canada (Terrestrial Arthropods), Ottawa, Canada.
Diehl, A.G., Ouyang, N., Boyle, A.P., 2020. Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes. Nature Comm. 11, 1796.
Dong, Y., Desneux, N., Lei, C., Niu, C., 2014. Transcriptome characterization analysis of Bactrocera minax and new insight into its pupal diapause development with gene expression analysis. Int. J. Biol. Sci. 10, 1051–1063.
Drongitis, D., Aniello, F., Fucci, L., Donizetti, A., 2019. Roles of transposable elements in the different layers of gene expression regulation. Int. J. Mol. Sci. 20 (22), 5755.
Dubin, M.J., Mittelsten Scheid, O., Becker, C., 2018. Transposons: a blessing curse. Curr. Opin. Plant Biol. 42, 23–29.
Emerson, K.J., Bradshaw, W.E., Holzapfel, C.M., 2010. Microarrays reveal early transcriptional events during the termination of larval diapause in natural populations of the mosquito, Wyeomyia smithii. PLoS ONE 5, e9574.
Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T., Waki, K., Hornig, N., Arakawa, T., Takahashi, H., Kawai, J., Forrest, A.R.R., Suzuki, H., Hayashizaki, Y., Hume, D.A., Orlando, V., Grimmond, S.M., Carninci, P., 2009. The regulated retrotransposon transcriptome of mammalian cells. Nature Genet. 41 (5), 563–571.
Feng, G., Leem, E.Y., Levin, H.L., 2013. Transposon integration enhances expression of stress response genes. Nucl. Acids Res. 41 (2), 1775–1789.
Feschotte, C., 2008. Transposable elements and the evolution of regulatory networks. Nature Rev. Genet. 9 (5), 397–405.
Gong, Z.-J., Wu, Y.-Q., Miao, J., Duan, Y., Jiang, Y.-L., Li, T., 2013. Global transcriptome analysis of orange wheat blossom midge, Sitodiplosis moseliana (Gehin) (Diptera: Cecidomiidae) to identify candidate transcripts regulating diapause. PLoS ONE 8, e71564.
Hao, Y-J, Zhang, Y-J. Si, F-L, Fu, D-Y, He, Z-B and Chen, B., 2016. Insight into the possible mechanism of the summer diapause of Delia antiqua (Diptera: Anthomyiidae) through digital gene expression analysis. Insect Sci. 23: 438-451.
Horváth, V., Merenciano, M., González, J., 2017. Revisiting the relationship between transposable elements and the eukaryotic stress response. Trends Genet. 33 (11), 832–841.
Huang, X., Poelchau, M.F., Armbruster, P.A., 2015. Global transcriptional dynamics of diapause induction in non-blood-fed and blood-fed Aedes albopictus. PLoS Neg. Trop. Dis. Dis. 9, e0003724.
Hubley, R., Finn, R.D., Clements, J., Eddy, S.R., Jones, T.A., Bao, W., Smit, A.F.A., Wheeler, T.J., 2016. The Dfam database of repetitive DNA families. Nucleic Acids Res. 44 (D1), D81–D89.
Hummel, B., Hansen, E.C., Yoveva, A., Aprile-Garcia, F., Hussong, R., Sawarkar, R., 2017. The evolutionary capacitor HSP90 buffers the regulatory effects of mammalian endogenous retroviruses. Nature Struct. Mol. Biol. 24 (3), 234–242.
Kang, D.S., Cotton, M.A., Denlinger, D.L., Sim, C., 2016. Comparative transcriptomics reveals key gene expression differences between diapausing and non-diapausing adults of Culex pipiens. PLoS ONE 11, e0154892.
Kankare, M., Parker, D.J., Merisalo, M., Salminen, T.S., Hoikkala, A., 2016. Transcriptional differences between diapausing and non-diapausing D. montana females reared under the same photoperiod and temperature. PLoS ONE 11, e0161852.
Kemp, W.P., Bosch, J., Dennis, B., 2004. Oxygen consumption during the life cycle of the prepupa-wintering bee Megachile rotundata (F.) and the adult-wintering bee Osmia lignaria (Hymenoptera: Megachilidae). Ann. Entomol. Soc. Am. 97, 161–170.
Klostermeyer, E.C. Biology of the alfalfa leafcutting bee. 1982. G.H. Rank (Ed.), Proceedings of the First International Symposium on Alfalfa Leafcutting Bee Management, University of Saskatchewan, Saskatoon, Saskatchewan, pp. 10-19.
Kostál, V., 2006. Eco-physiological phases of insect diapause. J. Insect Physiol. 52 (2), 113–127.
Kostál, 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. Proc. Nat. Acad. Sci. USA 114 (32), 8532–8537.
Kumar, S.V., Wigge, P.A., 2010. H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell 140 (1), 136–147.
Leem, Y.-E., Ripmaster, T.L., Kelly, F.D., Ebina, H., Heincelman, M.E., Zhang, K.e., Grewal, S.I.S., Hoffman, C.S., Levin, H.L., 2008. Retrotransposon Tf1 is targeted to Pol II promoters by transcription activators. Mol. Cell 30 (1), 98–107.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25 (16), 2078–2079.
Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM arXiv.
Liu, S., Jonge, J., Trejo-Arellano, M.S., Santos-González, J., Köhler, C., Hennig, L., 2021. Role of H1 and DNA methylation in selective regulation of transposable elements during heat stress. New Phyt. 229 (4), 2238–2250.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 31–121.
Meyers, P.J., Powell, T.H.Q., Walden, K.K.O., Schieferecke, A.J., Feder, J.L., Hahn, D.H., Robertson, H.M., Berlcher, S.H., Ragland, G.J., 2016. Divergence of the diapause transcriptome in apple maggot flies: winter regulation and post-winter transcriptional repression. J. Exp. Biol. 219, 2613–2622.
Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J.R.A., Hellinga, A.J., Lugo, C.S.B., Elliott, T. A., Ware, D., Peterson, T., Jiang, N., Hirsch, C.N., Hufford, M.B., 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. Genome Biol. 20, 275.
Poelchau, M.F., Reynolds, J.A., Denlinger, D.L., Elsik, C.G., Armbruster, P.A., 2011. A de novo transcriptome of the Asian tiger mosquito, Aedes albopictus, to identify candidate transcripts for diapause preparation. BMC Genomics 12, 619.
Poelchau, M.F., Reynolds, J.A., Elsik, C.G., Denlinger, D.L., Armbruster, P.A., 2013a. Deep sequencing reveals complex mechanisms of diapause preparation in the invasive mosquito, Aedes albopictus. Proc. R. Soc. B. 280, 20130143.
Poelchau, M.F., Reynolds, J.A., Elsik, C.G., Denlinger, D.L., Armbruster, P.A., 2013b. RNA-Seq revels early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, Aedes albopictus. J. Exp. Biol. 216, 4082–4090.
Poupardin, R., Schottner, K., Korbelova, J., Provaznik, J., Zolexel, D., Pavlinic, D., Benes, V., Kostal, V., 2015. Early transcriptional events linked to induction of diapause revelaed by RNAseq in larvae of drosophilid fly, Chymomyza costata. BMC Genomics 16, 720.
Qi, X., Zhang, L., Han, Y., Ren, X., Huang, J., Chen, H., 2015. De novo transcriptome sequencing and analysis of Coccinella septempunctata L. in non-diapause, diapause and diapause-terminated states to identify diapause-associated genes. BMC Genomics 16, 1086.
Ragland, G.J., Denlinger, D.L., Hahn, D.A., 2010. Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. Proc. Nat. Acad. Sci. U.S.A. 107, 14909–14914.
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. J. Exp. Biol. 214, 3948–3960.
Reynolds, J.A., Clark, J., Diakoff, S.J., Denlinger, D.L., 2013. Transcriptional evidence for small RNA regulation of pupal diapause in the flesh fly, Sarcophaga bullata. Insect Bioch. Mol. Biol. 43, 982–989.
Reynolds, J.A., 2017. Epigenetic Influences on Diapause. in: Advances in Insect Physiology v. 53. Verlinden H and Jurenka R ed. Academic Press, London, England.

Rinehart, J.P., Li, A., Yocum, G.D., Robich, R.M., Hayward, S.A., Denlinger, D.L., 2007. Up-regulation of heat shock proteins is essential for cold survival during insect diapause. Proc. Nat. Acad. Sci. USA 104 (27), 11130–11137.

Robich, R.M., Rinehart, J.P., Kitchen, L.J., Denlinger, D.L., 2007. Diapause-specific gene expression in the northern house mosquito, Culex pipiens L., identified by suppressive subtractive hybridization. J. Insect Physiol. 53, 235–245.

Romero-Soriano, V., Garcia Guerreiro, M.P., Michalak, P., 2016. Expression of the retrotransposon Helena reveals a complex pattern of TE deregulation in Drosophila hybrids. PLoS ONE 11 (1), e0147903.

Ryan, C.P., Brownlie, J.C., Whyard, S., 2016. Hsp90 and physiological stress are linked to autonomous transposon mobility and heritable genetic change in nematodes. Genome Biol. Evol. 8 (12), 3794–3805.

Sgolastra, F., Kemp, W.P., Maini, S., Bosch, J., 2012. Duration of prepupal summer dormancy regulates synchronization of adult diapause with winter temperatures in bees of the genus Osmia. J. Insect Physiol. 58 (7), 924–933.

Shimizu, Y., Tamai, T., Goto, S.G., 2018. Cell cycle regulator, small silencing RNA, and segmentation patterning gene expression in relation to embryonic diapause in the band-legged ground cricket. Insect Biochem. Mol. Biol. 102, 75–83.

Signor, S., 2020. Transposable elements in individual genotypes of Drosophila simulans. Ecol. Evol. 130, 499–511.

Smith, T.F., Waterman, M.S., 1981. Identification of common molecular subsequences. J. Mol. Biol. 147 (1), 195–197.

Stephen, W.P., Osgood, C.E., 1965. The induction of emergence in the leaf-cutter bee Megachile rotundata, an important pollinator of alfalfa. J. Economic Entomol. 58, 284–286.

Taliaferro, J.M., Aspden, J.L., Bradley, T., Marwha, D., Blanchette, M., Rio, D.C., 2013. Two new and distinct roles for Drosophila Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression. Genes Dev. 27 (4), 378–389.

Tarailo-Graovac, M., Chen, N., 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr. Prot. Bioinf. Chapter 25 (1), 4–10.

Tauber, M.J., Tauber, C.A., Masaki, S., 1986. Seasonal Adaptation of Insects. Oxford University Press, Oxford, England.

Tiedeman, Z. and Signor, S. 2020. The transposable elements of the Drosophila serrata reference panel. bioRxiv. doi:10.1101/2020.06.11.146431v1.full.

Tittel-Elmer, M., Bucher, E., Broger, L., Mathieu, O., Paszkowski, J., Vaillant, I., Barsh, G. S., 2010. Stress-induced activation of heterochromatic transcription. PLoS Genetics 6 (10), e1001175.

Torchio, P.F., 1989. In-nest biologies and development of immature stages of three Osmia species (Hymenoptera: Megachilidae). Ann. Entomol. Soc. Am. 82 (5), 599–615.

Torchio, P.F., Bosch, J., 1992. Biology of Tricrania stansburyi, a meloid beetle cleptoparasite of the bee Osmia lignaria propinqua (Hymenoptera: Megachilidae). Ann. Entomol. Soc. Am. 85, 713–721.

Torson, A.S., Yocum, G.D., Rinehart, J.P., Kemp, W.P., Bowsher, J.H., 2015. Transcriptional responses to fluctuating thermal regimes underpinning differences in survival in the solitary bee Megachile rotundata. J. Exp. Biol. 218, 1060–1068.

Torson, A.S., Yocum, G.D., Rinehart, J.P., Nash, S.A., Kvidera, K.M., Bowsher, J.H., 2017. Physiological responses to fluctuating temperatures are characterized by distinct transcriptional profiles in a solitary bee. J. Exp. Biol. 220, 3372.

Van Meter, M., Kashyap, M., Rezazadeh, S., Geneva, A.J., Morello, T.D., Seluanov, A., Gorbunova, V., 2014. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. Nature Comm. 5 (1), 5011.

Villanueva-Cañas, J.L., Horvath, V., Aguilera, L., González, J., 2019. Diverse families of transposable elements affect the transcriptional regulation of stress-response genes in Drosophila melanogaster. Nucl. Acids Res. 47 (13), 6842–6857.

Voronova, A., Belevich, V., Jansons, A., Rungis, D., 2014. Stress-induced transcriptional activation of retrotransposon-like sequences in the Scots pine (Pinus sylvestris L.) genome. Tree Genet. Genomes 10 (4), 937–951.

Yamashita O, Shiomi K, Ishida Y, Katagiri N, Niimi T. 2001. Insights for future studies on embryonic diapause promoted by molecular analyses of diapause hormone and its action in Bombyx mori. In: Denlinger DL, Giebultowicz J, Saunders DS, editors. Insect timing: circadian rhythmicity to seasonality. Elsevier Science; Amsterdam: 2001. pp. 145–152.

Yocum, G.D., Rinehart, J.P., Horvath, D.P., Kemp, W.P., Bosch, J., Alroobi, R., Salem, S., 2015. Key molecular processes of the diapause to post-diapause quiescence transition in the alfalfa leafcutting bee Megachile rotundata identified by comparative transcriptome analysis. Physiol. Ent. 40 (2), 103–112.

Yocum, G.D., Childers, A.K., Rinehart, J.P., Rajamohan, A., Pitts-Singer, T.L., Greenlee, K.J., et al., 2018. Environmental history impacts gene expression during diapause development in the alfalfa leafcutting bee, Megachile rotundata. J. Exp. Biol. 221 (jeb173443).

Zovoilis, A., Cifuentes-Rojas, C., Chu, H.-P., Hernandez, A.J., Lee, J.T., 2016. Destabilization of B2 RNA by EZH2 activates the stress response. Cell 167 (7), 1788–1802.e13.