Divergence of the diapause transcriptome in apple maggot flies: winter regulation and post-winter transcriptional repression

Peter J. Meyers[1], Thomas H. Q. Powell[2], Kimberly K. O. Walden[3], Adam Shieferecke[4], Jeffrey L. Feder[1,5], Daniel A. Hahn[2], Hugh M. Robertson[3], Stewart H. Berlocher[3], and Gregory J. Ragland [1,4,5,6*]

[1] Department of Biological Sciences, University of Notre Dame

[2] Entomology and Nematology Department, University of Florida

[3] Department of Entomology, University of Illinois

[4] Department of Entomology, Kansas State University

[5] Environmental Change Initiative, University of Notre Dame

[6] Current Address: Department of Integrative Biology, University of Colorado, Denver

[*] Correspondence:

Gregory Ragland, Department of Integrative Biology, University of Colorado, Denver, 1151 Arapahoe, SI 2071, Denver, CO 80204, USA. email: GREGORY.RAGLAND@UCDENVER.EDU

Keywords: diapause, phenology, overwintering, transcriptome, *Rhagoletis*

Summary Statement

Transcriptomics reveals that summer emergence timing of flies is regulated during winter, while initial transcriptional responses to post-winter warming are largely suppressed.

Abstract

Duration of dormancy regulates seasonal timing in many organisms and may be modulated by day length and temperature. Though photoperiodic modulation has been well studied, temperature modulation of dormancy has received less attention. Here, we leverage genetic variation in diapause in the apple maggot fly, *Rhagoletis pomonella*, to test whether gene expression during winter or following spring warming regulates diapause duration. We used RNAseq to compare transcript abundance during and after simulated winter between an apple-infesting population and a hawthorn-infesting population where the apple population ends pupal diapause earlier than the hawthorn-infesting population. Marked differences in transcription between the two populations during winter suggests that the 'early' apple population is developmentally advanced compared to the 'late' hawthorn population prior to spring warming, with transcripts participating in growth and developmental processes relatively up-regulated in apple pupae during the winter cold period. Thus, regulatory differences during winter ultimately drive phenological differences that manifest themselves in the following summer. Expression and polymorphism analysis identify candidate genes in the Wnt and insulin signaling pathways that contribute to population differences in seasonality. Both populations remained in diapause and displayed a pattern of up- and then down-regulation (or vice versa) of growth-related transcripts following warming, consistent with transcriptional repression. The ability to repress growth stimulated by permissive temperatures is likely critical to avoid mismatched phenology and excessive metabolic demand. Compared to diapause studies in other insects, our results suggest some overlap in candidate genes/pathways, though the timing and direction of changes in transcription are likely species-specific.

Journal of Experimental Biology • Advance article

Introduction

The waxing and waning of climatic factors and resources over predictable time periods dictates nearly every aspect of an organism's life history in seasonal environments. Accordingly, understanding the synchronization of growth and reproduction with permissive conditions and the physiological mechanisms that determine this synchrony has been an important focus of research on adaptive evolution. Through such studies, a general understanding has emerged of how life histories are shaped by environmental uncertainty (Childs et al., 2010; Cohen, 1970; Kingsolver, 1979), of environmental effects on growth and development (Amano et al., 2014; Powell et al., 2000; Taylor, 1981), and of the genetic architecture underlying seasonal timing (Bradshaw et al., 2012; Lair et al., 1997; Li et al., 2010; Schmidt et al., 2008). Similarly, the physiological mechanisms that regulate seasonal timing in animals and plants have been investigated, particularly endocrine mechanisms regulating dormant, overwintering life stages (diapause) in insects (Denlinger, 2002; Hahn and Denlinger, 2011).

Diapause in insects is typically induced and ended by environmental cues such as temperature and photoperiod. These cues are important even for univoltine species that enter an obligate diapause stage, because termination of obligate diapause is often sensitive to temperature and photoperiod (Tauber et al., 1986). Different physiological mechanisms likely transduce temperature, photoperiod, and other important cues such as diet quality or quantity. However, the same major neuroendocrine systems regulate entrance into and exit from developmental arrest and metabolic suppression broadly across taxa, although regulatory mechanisms upstream of hormonal signals appear to be taxon-specific (Denlinger et al., 2005). These neuroendocrine signals control the

physiologically dynamic progression of insects through initiation, maintenance, and termination phases collectively termed 'diapause development' (Kostal, 2006).

Because diapause involves a developmental progression, processes affecting the rate of diapause development will affect diapause duration. Diapause duration, in turn, determines the seasonal timing of exit from diapause and resumption of active growth and reproduction. Thus, regulatory mechanisms acting relatively early in diapause development may influence the timing of the end of diapause occurring weeks, months, or even years later. These mechanisms may be less important for species that complete diapause in response to a specific photoperiodic cue. For example, diapausing pitcher plant mosquito larvae rapidly end diapause and resume active growth when switched from relatively short to long day lengths, constituting a reliable cue for permissive environmental conditions in the field (Bradshaw and Lounibos, 1977). However, temperature appears to be the primary factor determining the duration of diapause in many species, even for some insects that initiate diapause based on photoperiodic cues (Tauber et al., 1986). Thermal responses over the course of diapause in such species may thus regulate diapause development rate, and consequently the total duration of diapause.

While physiological responses to photoperiod have been well established in a number of diapausing species, little is known about when and how temperature modifies diapause development. Photoperiodic responses can be readily manipulated in the lab, and numerous physiological and genetic studies have identified candidate genes and mechanisms that may transduce photoperiodic signals upstream and downstream of major hormonal cues (Emerson et al., 2010; Poelchau et al., 2013; Poupardin et al.,

2015; Schmidt et al., 2008; Sim and Denlinger, 2008; Tauber et al., 2007; Wadsworth and Dopman, 2015; Williams et al., 2006). In contrast, it has been more difficult to experimentally manipulate temperature effects because the sensitivity of development to thermal conditions often changes over the course of diapause (Hodek and Hodkova, 1988). For example, many, though not all diapause-overwintering insects have relatively low thermal thresholds above which diapause development does not proceed or progresses very slowly, similar to vernalization in plants. In the literature, such a threshold is often referred to as a 'chilling requirement', which is something of a misnomer because typically exposure to low temperature is not necessarily required, but rather affects the rate at which insects progress through phases of diapause development (Hodek and Hodkova, 1988). In addition, thermal thresholds can change over the course of diapause. As a result, thermal conditions during the fall, winter, and spring may have varied effects on the timing of diapause termination.

Here, we test for transcriptomic signatures of regulatory events that may determine the duration of diapause in the apple maggot fly (*Rhagoletis pomonella*; Diptera: Tephritidae). Specifically, we test for transcriptomic differences between apple and hawthorn-infesting host races of *R. pomonella* that recently evolved seasonal differences in the timing of their diapause termination. Pupal diapause in *R. pomonella* is functionally obligate in the field (populations are univoltine) and is mainly influenced by overwinter and post-winter temperature (Neilson, 1962). The termination of diapause synchronizes adult flies with the availability of host fruits where they court, mate, oviposit their eggs, and larvae develop. Recently derived populations of *R. pomonella* attacking apple (*Malus domestica*) have evolved to terminate their diapause earlier than flies attacking the ancestral hawthorn (*Crataegus sp.*) host to track the

earlier fruiting time of apples compared to hawthorns (Bush, 1969; Feder and Filchak, 1999; Feder et al., 1993; Smith, 1988). Though hawthorn- and apple-infesting populations eat different fruits as larvae, apples and hawthorns are nutritionally similar, with larger apples generally yielding larger-bodied adults with greater lipid reserves because of greater resource quantity (Ragland et al., 2012). However, common garden experiments rearing both populations through apple clearly demonstrate that differences in diet do not contribute to the pronounced population difference in eclosion time observed in the field or the laboratory (Dambroski and Feder, 2007; Feder and Filchak, 1999; Smith, 1988).

We applied an RNA sequencing (RNAseq) approach to test two hypotheses concerning the eclosion time difference between apple and hawthorn flies. We examine whether earlier diapause termination and adult eclosion in apple flies is associated with: 1) differences in gene expression during winter ("*during winter*" hypothesis); or 2) changes in gene expression in response to warming temperatures following winter ("*after winter*" hypothesis). If the *during winter* hypothesis is true, then by the end of winter many genes associated with development will be differentially expressed in apple compared to hawthorn flies (Fig. 1A). In contrast, if the *after winter* hypothesis is true, then developmentally related genes will be differentially expressed between the host races only following post-winter warming (Fig. 1B). Both hypotheses predict that apple flies up- or down-regulate development-related genes earlier than hawthorn flies, in accord with apple flies terminating diapause and eclosing earlier as adults than hawthorn flies. The distinction concerns when the predicted change in gene expression occurs between the two host races, during winter or during spring/summer. We note that dietary differences between populations could contribute to observed gene

Journal of Experimental Biology • Advance article

expression differences, but are unlikely to account for population differences in the non-feeding, diapausing pupal stage given that diet does not account for population differences in *R. pomonella* seasonality (Dambroski and Feder, 2007; Smith, 1988). Finally, we test whether single nucleotide polymorphisms (SNPs) in differentially expressed genes between apple and hawthorn flies display significant allele frequency differences between the host races in nature. Such a finding would imply that the differentially expressed loci could be the actual targets of divergent selection on diapause timing.

Materials and Methods

Study system and fly rearing

Rhagoletis pomonella is a frugivorous fly native to North America that infests fruits of various hawthorn species in Eastern North America (Berlocher and McPheron, 1996). After apples were introduced from Eurasia (~400 years ago), a population of these flies evolved to specialize to attack apple fruit during the mid-nineteenth century (Feder et al., 1988; Walsh, 1867). The derived apple- and ancestral hawthorn-infesting populations or 'host races' are hypothesized to be in the early stages of speciation-with-gene-flow, with natural selection maintaining genetic divergence despite ongoing gene flow (up to 4% migrants per year; (Feder et al., 1994)). Seasonal timing is a primary target of divergent natural selection, driven by differences in fruiting time between apples and hawthorns. At a typical site in the Midwestern US with sympatric apple and hawthorn fly host races, apple trees fruit on average three weeks earlier than hawthorn trees (Feder and Filchak, 1999). The fly has one generation per year, wherein adults rendezvous and mate on host fruit, females oviposit into fruit, larvae consume the fruit then exit, burrowing into the soil and entering a pupal diapause that lasts until the following growing season. Natural selection to synchronize adults with host fruit availability is very strong because fruits are only available for a month or less, and adult flies typically only live a few weeks (Feder and Filchak, 1999). Individuals overwinter as diapausing pupae. Thus, there is strong selection for apple flies to terminate pupal diapause earlier compared to hawthorn flies, because the timing of the end of diapause determines the timing of adult emergence. Genetic association studies reveal that loci most strongly associated with eclosion timing are also the most genetically divergent loci between haw- and apple-infesting populations (Feder et al., 1993) (Michel et al., 2010).

The goal of our experiment was to identify transcriptional differences between the apple and hawthorn host races during diapause development that may underlie their observed difference in seasonality (Fig. 1). We collected larval-infested apple fruit from East Lansing, Michigan on September 15, 2013, and hawthorn fruits from the University of Notre Dame campus (Notre Dame, Indiana) on October 15, 2013. Although apple and hawthorn flies from these two collecting sites do not represent a co-occurring sympatric population pair, they are located in the same ecogeographic region at similar latitudes and the ~ 1 month fruiting time difference between E. Lansing and Notre Dame mirrors differences observed at sympatric sites. Temperatures prior to diapause can affect diapause incidence and duration, but thermal exposure after larval wandering is of primary importance (Feder et al., 1997a; Neilson, 1962), and this exposure was controlled and equivalent for both populations in our experimental design.

Infested fruit were placed on wire mesh trays held over plastic collecting bins in a greenhouse maintained at 23°C (natural light conditions). Wandering third instar larvae emerging from fruit were collected daily, placed in petri dishes maintained at 85% RH, and allowed to complete pupal development at 23°C for 10 days. The difference in the collection times of apple and hawthorn flies from E. Lansing and Notre Dame resulted in a difference in photoperiod exposure between the host races in the greenhouse reflecting natural conditions. However, photoperiod during pre-winter development affects diapause development of *R. pomonella* minimally and does not affect pupae or account for host race differences in diapause development (Filchak et al., 2001). After 10 days, pupae were moved to a refrigerator and held at 4°C in constant darkness for 23 weeks (±3 days) to simulate winter. Previous studies have shown that when pupae are

subsequently removed from the simulated winter and exposed to spring-like temperatures (20-25°C), apple flies terminate pupal diapause earlier than hawthorn flies (Feder and Filchak, 1999; Feder et al., 1997a; Smith, 1988). We note that while a shift from cold to warm temperatures is necessary to terminate diapause, diapause does not end immediately. Pupae remain developmentally arrested and metabolically depressed for days to weeks before initiating pupal-to-adult apolysis (molting) and beginning adult morphogenesis (Ragland et al., 2011; Ragland et al., 2009).

Experimental Design

The experiment was designed to test whether the apple and hawthorn host races diverge in growth-related transcription levels during winter (*during winter* hypothesis) or only after a shift from cold to warm temperatures simulating spring (*after winter* hypothesis; Fig. 1). To answer this question, we compared apple versus hawthorn fly transcription profiles at three sampling time points: 1) directly out of the 4°C, 23 week winter treatment, 2) 24 hours after removal from the cold and transfer to 23°C, and 3) 48 hours after moving from 4°C to 23°C, hereafter '0hr', '24h', and '48h' time points (Fig. 1C). Thus, transcriptional differences between the host races at the 0h time point reflect regulatory differences induced during overwintering, while differences after the thermal shift reflect effects of post-winter warming.

Diapause developmental progression is under neuroendrocrine control, so we sampled the transcriptome of only fly pupal heads containing the brain, ring gland, and suboesophageal ganglion. For the 0h sample, pupae were flash-frozen in liquid nitrogen and rapidly decapitated with a sterilized razor blade, then transferred to Ambion TRI Reagent (Life Technologies, NY). To obtain sufficient amounts of RNA and to account for

intra-population variation in developmental progression, we pooled the heads of ten flies from the same race in a single extraction. Heads were pestle-homogenized and stored at -80°C for no more than 4 weeks before RNA was extracted using Ambion RiboPure kits following the manufacturers recommendations. To confirm that pupae sampled after the shift to 23°C were still in diapause, we conducted stop-flow respirometry on each individual following (Ragland et al., 2009), using a LI-COR 6252 CO_2 analyzer (LI-COR Biosciences, Lincoln, NB) coupled to Sable Systems pumping and metering components (Sable Systems International, Las Vegas, NV) to measure metabolic rates prior to sampling at 24 and 48 hours after the temperature shift. All pupae measured at 24 and 48 hours exhibited metabolic rates indicative of diapause (Supplemental table S1-1). After respirometry, pupae sampled at 24 and 48 hours were flash frozen, decapitated, pooled, and extracted as described above for the initial 0h time point. We generated three replicate, pooled samples (10 heads per pool) for each population (apple and haw) at each time point, yielding 18 total samples.

Library preparation, sequencing, and informatics

Libraries were prepared for RNA sequencing at the Notre Dame Genomics and Bioinformatics core facility using the TruSeq RNA Sample Preparation v2 Kit (Illumina, Inc., San Diego, CA). Quality control identified one library (an Apple 24h sample) of poor quality that was excluded from further analysis, leaving 17 total samples. Libraries were sequenced (100bp Paired End) across two lanes (9 and 8 samples multiplexed per lane) on an Illumina HiSeq 2000 at the Beijing Genomics Institute. After de-multiplexing, we excluded one additional sample (a Haw 0h sample) with very low read counts from further analysis. All statistics thus reflect an analysis of 16 total samples, three replicate pools per population per treatment except for Apple 24h and Haw 0h, each of which had

only two replicate pools. The generalized linear models that we describe below are valid with imbalanced replicates, though imbalance does reduce statistical power. Excluding the poor samples, we were left with 380,831,464 total paired-end reads.

Though a genome is not currently completed for *R. pomonella*, there is a published transcriptome based on 454 sequence data (Schwarz et al., 2009). We also generated additional 454 transcriptome sequence data in the current study based on pools of 100 adult apple and hawthorn fly heads (50 male, 50 female for each race) sampled directly from host trees at a sympatric site in Urbana, Il. Live adult flies were individually aspirated from unsprayed apple trees at the University of Illinois Pomology area and hawthorn trees near the University of Illinois married student housing. Heads were removed, pooled, and DNA was extracted using a phenol-chloroform procedure. Libraries were generated using a Roche Lib-L emPCR kit and XL+ sequencing kit (Roche, Indianapolis, USA) and sequenced on a single Roche/454 GS FLX Titanium plate (one quarter plate per sex/host race combination). The resulting data were used to estimate allele frequency differences for SNPs between apple and hawthorn fly populations in nature, and were combined with the current Illumina RNAseq data and the previously published 454 data to produce an updated transcriptome assembly. In order to use the Trinity assembler (Haas et al., 2013), which is optimized for short reads, we first simulated 100bp paired end reads from the 454 data using a custom script (see DRYAD archive). These simulated reads were pooled with the Illumina reads, and Trinity was run on the combined pool using default parameters. The assembly yielded 212,600 isogroups or clusters of transcripts likely to represent single transcripts with alternate splicing. We used TGICL (Pertea et al., 2003) to identify redundant sequences, but this yielded negligible improvement (only 4,368 redundant sequences). Thus, we started all

of our analyses with 212,600 isogroups prior to filtering (see below). Compared to the previously published assembly that contained only 27% percent of a conserved, Benchmarking Universal Single-Copy Ortholog set of genes compiled for arthropoda (BUSCO set, see (Simão et al., 2015)), the new assembly contains 79% of arthropod BUSCOs, a substantial improvement. All sequences were annotated using blastx searches against Flybase and Uniprot databases (expect <= $1x10^{-4)}$. Analyses requiring annotations used only the Flybase annotations, though all annotations are provided in Supplement S1-2 (see Supplement S1-3 for all assembly statistics).

For differential expression analysis (see below), all retained Illumina reads were mapped to the new reference transcriptome using RNA star (Dobin et al., 2013) with default parameters. We then used RSEM (Li and Dewey, 2011) to count reads per isogroup (total across all possible alternative splice variants within a isogroup). Transcripts not represented by at least one count in 50% of the samples were filtered out, leaving a total of 66,235 of the initial 212,600 transcripts for downstream analysis.

Differential expression analysis

We used the edgeR package to apply a generalized linear model to the read count data assuming a negative binomial distribution (Robinson et al., 2010). Scale factors calculated using the weighted trimmed mean of M-values (TMM) method were incorporated into the model, correcting for differences in library composition (Robinson and Oshlack, 2010). We started by fitting a full model to each transcript that included an interaction term:

$$y_{ijk} = \mu + H_i + T_j + (HT)_{ij} + e_{ijk}$$

Where host (*H*) and time point (*T*) are factors, and (*HT*) is the interaction term. Likelihood ratio tests were applied to test the statistical significance of each term, applying a Benjamini and Hochberg false discovery rate (FDR) threshold of 0.05 as a significance cutoff. For all transcripts with a non-significant interaction term, we applied a reduced model excluding the interaction effect. We then used linear contrasts to test the following null hypotheses: 1) no difference between the host races, 2) no difference between the 0h and 24h treatments, and 3) no difference between the 24h and 48h treatments. Overlaps of these sets were visualized in an area-proportional Venn diagram using eulerAPE (Micallef and Rodgers, 2014) and tested for significance using Fisher's exact tests.

Our test to distinguish the *during* vs. *after* winter hypotheses centered on the host and host-by-time point interaction terms in the linear model. Specifically, models for transcripts with a significant main host term but a non-significant interaction term are consistent with the *during winter* hypothesis (Fig. 1A). Moreover, we predict that such genes, if they are involved in diapause termination, should be significantly up or down regulated in apple vs. hawthorn pupae at the 0h treatment and that this host-associated difference should remain constant (or show the same directionality) in the subsequent 24h and 48h samples following heating. We tested for significant differences between the host races in the 0, 24, and 48h sampling time points using linear contrasts. In comparison, models for transcripts that have a significant interaction term are consistent with the *after winter* hypothesis for a host race difference occurring after the transfer to warmer conditions. Moreover, for the *after winter* hypothesis to be correct, not only should there be a significant interaction term, but transcript levels should also

not differ significantly between the host races at the 0h point. Rather, they should vary between apple and hawthorn flies 24 and 48 hours after the temperature shift, and in the same direction for both comparisons (Fig. 1B).

We also hypothesized that if differentially expressed sets of transcripts exhibiting host race differences during winter or after heating were involved in diapause development, then they should demonstrate enrichment or over-representation of functional categories related to growth and development. To test for such enrichment, we submitted annotated lists of genes identified in the tests above to DAVID (Huang et al., 2008) for functional category (e.g., GO, KEGG, INTERPRO) enrichment analysis using the statistical procedure described at (https://david.ncifcrf.gov/helps/functional_annotation.html#fisher).

454 SNP analysis

The 454 sequencing of separate pools of 100 heads each from apple- and hawthorn-infesting populations at the Urbana site allowed testing for SNP allele frequency differences between the host races, using read counts to estimate population allele frequencies (Futschik and Schlötterer, 2010). We identified SNPs using the Genome Analysis Toolkit (GATK) Unified Genotyper (version 3.3; (McKenna et al., 2010)), removing duplicate reads and filtering for a minimum phred-scaled SNP probability of 21 for bi-allelic SNPs with at least 10x coverage, and a minimum alternate allele frequency of 0.05. We then applied a Fisher's exact test (with Benjamini and Hochberg FDR correction) to the read counts to test for allele frequency differences. Given our filtering criterion, we applied the test to 65,793 total SNPs across 4,262 transcripts. We

then submitted the list of transcripts containing SNPs with significant FDR values to the DAVID tool to assess functional enrichment.

Results

Differential expression between the host races

Expression patterns supported the *during winter* hypothesis that the host races differ in diapause development during winter, with overwintering apple pupae appearing to be developmentally advanced relative to hawthorn pupae. Out of 66,235 total analyzed transcripts, 3,002 (4.5%) showed a significant main host effect (FDR < 0.05) and were differentially expressed in the same direction between apple and hawthorn pupae across all three measured time points (0h, 24h, and 48h) (Fig 2, S3). These 3,002 significant transcripts represent a maximum of 2,823 different genes (= total transcripts minus transcripts with the same annotation). In contrast, only 371 transcripts demonstrated a significant host race by time point interaction, as predicted by the *after winter* hypothesis. Moreover, of these 371 transcripts, 142 displayed a significant difference between the host races in the 0h treatment, suggesting that they were already differentially expressed in apple versus hawthorn flies prior to pupal heating, consistent with the *during winter* hypothesis. Only 166 of the transcripts having significant interaction terms were differentially expressed between the host races only in the 24 and 48h samples, as expected under the *after winter* hypothesis (S1-4). Thus, a total of 166 transcripts displayed a pattern that was consistent with the *after winter* hypothesis compared to 3,144 (n = 3,002 + 142) that supported the *during winter* hypothesis.

The 3,144 DE transcripts displaying a main host effect (n = 3,002) or host x time point interaction accompanied by a race difference in the 0h sample (n = 142) were significantly enriched for growth-related functional categories (e.g., neuron development, cell motion, gland development), suggesting that they represent differences directly involved in diapause progression (Table 1). As discussed in the Methods, all flies in the 24h and 48h remain in diapause. Hence, the transcription differences we observed during winter likely represented preparation for or the initiation of cell differentiation and proliferation in anticipation of diapause termination, rather than overt post-diapause development. The directionality of expression of the 3,144 differentially expressed transcripts further supported the *during winter* hypothesis of greater preparatory or developmental activity in apple flies. Only the subset of transcripts up-regulated in apple relative to hawthorn flies (n = 1,809) was enriched for the developmentally-related categories of loci; there was no enrichment for the transcripts down-regulated in apple fly pupae (n = 1,335; Table 1).

SNP analysis

From the 454 data set comparing pools of adult apple and hawthorn flies, we identified a total of 42 (FDR < 0.05) and 79 transcripts (FDR < 0.1) that contained at least one SNP displaying significant frequency differences between the host races at the Urbana, IL site (S1-5). These transcripts were significantly enriched for genes related to oxidation-reduction functions (S1-6). However, there was a potential detection bias for the SNPs displaying host differences, which had an average of 3x greater coverage compared to non-significant SNPs. The set of transcripts differentially expressed between apple and hawthorn flies at FDR < 0.05 was marginally enriched for SNPs showing significantly allele frequency differences between the host races at Urbana (p = 0.05, Fisher's exact

test). All told, these transcripts annotated to seven unique Flybase genes (*hui*, *CG3902*, *Mgstl*, *Ddx1*, *CG13639*, *Non1*, *CG9917*), including two loci involved in embryonic and imaginal disc development (*hui*, *Ddx1*) and one unnamed gene (*CG3902*) exhibiting physical interactions with PI3K, a major mediator of insulin-regulated events (S1-7).

Post-winter transcriptional responses

Following warming, both races undergo marked changes in expression for many transcripts that do not exhibit a host race by time interaction. Out of the total of 66,235 transcripts analyzed in the study, 22,803 (34%) were differentially expressed between the 0h and 24h samples. For the subset of 7,557 of these 22,803 transcripts showing the most pronounced expression differences (≥ 2-fold), particularly those up-regulated at the 24h time point (n = 5,633), there was no evidence for any enrichment for a functional category directly related to stress (Table 1), which is known to elicit massive transcriptional responses (Gasch et al., 2000; Sorensen et al., 2005). Rather, these transcripts were highly enriched in categories related to developmental progression and cell cycling that are typically observed in *Rhagoletis* when pupal diapause ends and morphogenesis resumes (Table 1; (Ragland et al., 2011)). However, metabolic rate measurements taken for each fly included in the 24 and 48h samples indicated that all pupae remained in diapause. Thus, diapause does not end immediately in either host race upon exposure to warmer temperature conditions, despite the pronounced transcriptomic response.

Discussion

Winter diapause development

Transcriptome-wide gene expression data support our *during winter* hypothesis (Fig. 1) that regulatory differences ultimately dictating seasonal timing during the summer occur during the winter, long before diapause is terminated. Moreover, signals of relative up-regulation of growth and development in apple compared to hawthorn flies suggests that diapause development is generally more advanced in apple compared to hawthorn fly pupae during winter, consistent with the apple race terminating diapause earlier than the hawthorn race. Though some transcriptional differences may be the result of geographic and temporal differences between the E. Lansing and Notre Dame collecting sites unrelated to diapause, we consider geographic variation to be an unlikely explanation for the marked enrichment in growth and developmental transcripts observed for differentially expressed genes between the host races. Likewise, differences in diet may influence transcription. For example, insulin signaling (discussed below) is clearly altered by diet. However, given that: 1) diet does not contribute substantially to host race differences in diapause duration (Dambroski and Feder, 2007; Smith, 1988), 2) genetic variation accounts for a large proportion of variation in eclosion timing in *R. pomonella* (Feder et al., 1993; Michel et al., 2010), and 3) a myriad of developmental processes related to morphogenesis are clearly up-regulated in apple relative to hawthorn flies (Table1, and see following two paragraphs), we suggest that many of the transcriptional differences that we observed influence diapause duration and are underlain by genetic differences between host races. However, future time series experiments will be necessary to test whether hawthorn pupae differentially regulate the same transcripts as apple flies during the winter, but at later time points. Similarly, higher resolution genotype-to-phenotype

associations will be required to bolster evidence for evolutionary divergence. Here we have identified several candidate SNPs in development-related transcripts, but none were also significantly differentially regulated between host races. Non-transcribed, regulatory regions are crucial for regulating transcript levels, however, and we expect that ongoing, full genome scans of multiple population pairs in *R. pomonella* will yield much greater statistical power and the ability to identify regulatory sequence variation.

Although many studies have documented the physiologically dynamic nature of diapause (Denlinger, 2002), it remains unclear whether diapause progression and embryogenic, morphogenic, or oogenic development are completely uncoupled. Several flow cytometry studies in flies provide clear evidence of cell cycle arrest during dormancy (Kostal, 2006; Tammariello and Denlinger, 1998), but does some level of cell proliferation or differentiation occur despite a background signal of cell cycle arrest? The host race differences in expression of developmental genes that we observed prior to diapause termination could be associated with preparatory steps or with actual cell proliferation and morphological differentiation. We do not know at precisely which developmental stage *R. pomonella* pupae arrest development during diapause, but it is very close to and before the onset of pupal-adult apolysis (Dean and Chapman, 1973), or stage 28 of *Drosophila melanogaster* metamorphosis in Bainbridge and Bownes (Bainbridge and Bownes, 1981). There are similar descriptions of pupal diapause arrest after head evagination but before the pupal-adult molt in other dipterans including two flesh flies (Fraenkel and Hsiao, 1968a; Fraenkel and Hsiao, 1968b) and another *Rhagoletis* fly, *Rhagoletis cerasi* (Papanastasiou and Papadopoulos, 2014). Further, the morphological progression observed immediately upon the cessation of metabolic arrest in *R. pomonella* is consistent with arrest before the pupal-adult molt (Ragland et

al., 2011). Every sampling point in our current study occurs prior to the overt developmental changes at and following pupal-adult apolysis, but it is possible that some low level of morphological change occurs between the onset of diapause and the end of diapause, and that the rate of this change is accelerated in earlier-eclosing apple flies. This would be consistent with a study on diapausing aphid embryos suggesting that cell proliferation continues at a decelerated but detectable rate during diapause (Shingleton et al., 2003).

Histological studies will be necessary to definitively test for morphogenesis during diapause, but the clear up-regulation of a number of genes in the Wnt signaling pathway in apple flies (independent of sampling time point) seems to support some amount of morphogenesis (Fig. 4). The *wnt* gene is a major hub for several *wnt*-mediated pathways, regulating diverse aspects of embryogenesis and morphogenesis (Reya and Clevers, 2005; Wodarz and Nusse, 1998). Wnt signaling is also connected to several other interacting developmental pathways, such as insulin signaling. We illustrate some of these interactions in Figure 4 by mapping our expression data to an interaction network including all transcripts significantly differentially expressed between host races that are connected to any gene in the Wnt or insulin signaling pathways (as determined by GO category and from the Interactive fly; http://www.sdbonline.org/sites/fly/aignfam/sgmtplty.htm#Wingless) by at least one protein-protein interaction (as determined from the DroID databse; http://www.droidb.org). Insulin signaling influences many aspects of growth and metabolism, and has repeatedly been connected to diapause developmental transitions (Sim and Denlinger, 2013). Note, however, that insulin signaling is sensitive to diet, so differences in insulin-related transcripts in our study may reflect feeding history rather

than evolved differences between the host races. Up-regulation of transcripts related to cuticle synthesis in apple pupae (Table 1) may also reflect early progression of adult cuticular development that will later culminate in pupal-adult apolysis.

To further explore the potential importance of transcripts that are differentially expressed between the host races during the cold period for developmental progression, we compared our lists of differentially expressed transcripts with transcripts that are differentially expressed between 0 and 24 hours post pupal formation in *D. melanogaster* (i.e., during early morphogensis; (Lebo et al., 2009). We find that 68 unique transcripts significantly differentially expressed in apple relative to hawthorn flies during our simulated winter treatments are also at least 2-fold differentially expressed (in the same direction) at 24 hours relative to 0 hours post pupal formation in *D. melanogaster* (see R script and gene list deposited in Dryad; *D. melanogaster* data archived in NCBI GEO GSE11313). This transcript set is highly enriched for cuticle synthesis proteins, suggesting that upregulation of cuticle synthesis processes in apple-origin pupae reflect developmental progression. This list also contains genes with specific GO annotations to neuron formation and morphogenesis, providing additional evidence that transcriptional differences observed during the simulated winter period reflect actual diapause developmental differences between the host races.

Post-winter transcriptional repression

Analysis of the direction of differential expression between the 0h and 24h versus the 24h and 48h time points revealed a robust pattern suggesting that diapausing pupae may actively suppress growth and developmental processes stimulated by permissive

temperatures. Of 1,394 transcripts differentially expressed across all time points (no host race by time interaction), 75% reversed the direction of their expression between the 0h to 24 h vs. 24h to 48h samples ($p << 0.001$ deviation from 50% expectation, binomial test). In other words, transcripts that increased in abundance between the 0h and 24 h time points tended to change direction and become down regulated between the 24 h and 48 h time points, and vice-versa. Genes displaying such reversals were enriched for several functional categories related to transcription and translation (Table 1). For example, every representative member of the GO category 'Ribosome Biogenesis' was up-regulated 24 hours after heating, but by 48 hours displayed expression levels approaching those of 0h (Fig. 3A). Members of the GO category 'Nucleotide Binding' that includes several important transcription factors involved in insulin signaling and cell cycle regulation connected to growth and development, showed a similar pattern (Fig. 3B). The group of transcripts that decreased in abundance 24 h after the shift to warm temperatures and then rebound after 48 h also contains *pepck*, a key enzyme in gluconeogenesis that is almost universally up-regulated in the diapause responses of insects, and subsequently down-regulated following diapause termination (Poelchau et al., 2013; Ragland et al., 2010). Here, *pepck* is initially down-regulated, as would be expected at the end of diapause, but by 48h *pepck* transcript abundance has returned to the same levels observed during diapause at 4°C. Overall, the strong overarching pattern of 'bounce back' in gene expression agrees well with observed initial increases followed by gradual decreases in metabolic rate following 4°C to 23°C shifts in other *R. pomonella* diapause experiments (Powell, Hahn, and Ragland, unpublished data).

Active repression of development and metabolism has significant implications for energy expenditure over the course of winter. Even during metabolic depression,

diapausing insects may exhaust fuel stores vital for successful completion of development following winter (Hahn and Denlinger, 2011). Events that increase winter temperatures, such as transient warm fronts that melt insulating snow layers may be particularly important, causing short periods of intense metabolic demand (Irwin and Lee, 2000; Williams et al., 2012a; Williams et al., 2012b). However, there is evidence for active suppression of temperature-elevated metabolism during winter in insects (Williams et al., 2015). In addition to increasing metabolic demand, transient warming exposes diapausing pupae to conditions permissive for diapause termination. Thus, the most drastic fitness outcome from winter warming would likely be premature termination of diapause during a transient event that will be followed by additional days or weeks of cold, unfavorable temperatures. Suppression of growth and development despite transiently permissive conditions is therefore critical to overwinter survival and may reflect the reversal in expression patterns for growth related transcripts we observed in the current study.

Comparison to other diapause studies

A rich literature on the comparative physiology of insect diapause establishes a clear connection between three major hormones, ecdysteroids, juvenile hormone (JH), and diapause hormone and the regulation of diapause (Denlinger et al., 2005). Diapause is a physiologically dynamic process that proceeds through stages often categorized as initiation, maintenance, and termination (Kostal, 2006). Hormones appear to regulate transitions between all of these stages, although the particular developmental roles and molecular interactions mediated by these hormones are often lineage-specific (Denlinger, 2002). Among this clear diversity of mechanisms, however, are some

commonalities in the functionally related sets of genes that have repeatedly been associated with diapause regulation across taxa.

Insulin is an important component of diapause regulation in several fly and mosquito species, where it regulates developmental arrest and nutrient provisioning. In flies (*D. melanogaster*) and mosquitoes (*Culex pipiens*), insulin signaling is an upstream regulator of juvenile hormone production, where knockdown of insulin signaling suppresses JH activity (Sim and Denlinger, 2013). A previous study in *R. pomonella* did not detect differential regulation of insulin-signaling genes during the rapid transcriptional changes associated with the end of diapause (Ragland et al., 2011). However, the regulation of diapause termination, which culminates in the end of diapause, likely involves a distinct set of genes that may or may not also change after the transition to post-diapause development. Here, we detected host race differences in the abundance of the insulin receptor *InR* and in *Tsc1*, which mediates cross-talk between the insulin and Tor pathways and has known effects on cell growth (Wullschleger et al., 2006). After 23 weeks at 4°C both Midwestern apple- and hawthorn-infesting populations remain in diapause (Feder et al., 1997b). Therefore the observed differences in expression between the apple and hawthorn host races occur during the maintenance and/or termination phases of diapause, and thus may reflect regulatory differences that influence when the end of diapause occurs and adults emerge. We observed a *decrease* in the expression of these insulin-signaling related genes in the apple host race, which by all other transcriptional indicators seems to be developmentally advanced relative to the hawthorn host race. In *C. pipiens* mosquitoes, an *increase* in insulin signaling regulates the timing of diapause termination (Sim and Denlinger, 2008). But, *C. pipiens* undergoes adult ovarian arrest, illustrating that

apparently conserved diapause-related genes may act in a species-specific and stage-specific manner.

Many developmental events are coordinated by the Wnt pathway, a conserved, master regulator of tissue proliferation and patterning across animals (Reya and Clevers, 2005; Wodarz and Nusse, 1998) that has also been previously connected to diapause development. In *D. melanogaster*, *wg*, a protein in the multi-member *wnt* family, sits at the head of the so-called 'canonical' and calcium-dependent Wnt signaling pathways that play a central role throughout embryogenesis and morphogenesis (Wodarz and Nusse, 1998). Various studies of diapause developmental arrest implicate down-regulation of the *wnt*-related pathways as either a regulator or result of the cessation of cell differentiation and proliferation (Wadsworth and Dopman, 2015; Wodarz and Nusse, 1998). In *R. pomonella*, calcium-dependent Wnt signaling is up-regulated one week after transfer from 4 to 23°C, suggesting a regulatory role upstream of pupal-adult apolysis and adult morphogenesis (Ragland et al., 2011). Here, we have detected increased transcript abundance mainly of the hub gene *wg* and downstream genes of the canonical Wnt pathways (the Beta-catenin-dependent pathway) in apple compared to hawthorn pupae. This is consistent with developmental advancement of the apple host race, which completes diapause and emerges earlier than the hawthorn host race in the field, and mirrors the positive relationship between developmental progression and Wnt signaling in other diapause studies (Chen and Xu, 2014; Lin et al., 2009).

Conclusions

Clear patterns of differential regulation of genes and pathways implicated in growth and development suggest that: 1) regulatory differences that lead to differences in summer emergence timing across host races of *R. pomonella* likely begin during winter, in either the diapause maintenance or early termination phases, and 2) upon transfer from cold, winter-like to warm, summer-like conditions, diapausing pupae appear to repress temperature-driven increases in growth and development, remaining firmly in diapause despite the permissive conditions. Key developmental genes involved in insulin and *wnt* signaling may play a role in the microevolution of seasonal timing in *R. pomonella* and in other insects. While whole transcriptome analyses are excellent at identifying large-scale physiological patterns, they can only nominate, but not confirm individual candidate genes, Further, many transcriptional differences between populations may reflect downstream responses to upstream, unobserved regulatory events because of the difficulty in sampling at precise time points at which environmental signals are transduced. Thus, additional functional testing combined with genotype-phenotype association tests will be necessary to confirm the roles of candidate genes in diapause development and termination.

Acknowledgements

We thank the Notre Dame genomics core for library preparations and the Kansas State University Beocat computing staff for computational support. This work was supported by funds from the Notre Dame Environmental Change initiative, USDA, and NSF IOS 1451274 to GJR and JLF, NSF IOS 1257298, the Florida Agricultural Experiment Station, and the joint FAO/IAEA CRP Dormancy Management to Enable Mass-rearing to DAH, and USDA AG 2007-35604-17886 to HR and SB.

Competing Interests

The authors declare no competing interests.

Author's Contributions

PJM, GJR, THQP, DAH, and JLF conceived the gene expression study, PJM and GJR designed the gene expression experiments, SHB, HMR, and GJR conceived the 454 SNP study, PJM performed the gene expression experiments, SHB, HMR, and KOW performed the 454 SNP study, PJM, GJR, SHB, HMR, and KOW analyzed the data, and all authors contributed to manuscript drafting.

Data Avalilability

All raw reads will be deposited (in progress) in the NCBI short read archive (SRA).

References

Amano, T., Freckleton, R. P., Queenborough, S. A., Doxford, S. W., Smithers, R. J., Sparks, T. H. and Sutherland, W. J. (2014). Links between plant species' spatial and temporal responses to a warming climate. *Proceedings of the Royal Society B-Biological Sciences* **281**.

Bainbridge, S. P. and Bownes, M. (1981). Staging the metamorphosis of Drosophila melanogaster. *Journal of embryology and experimental morphology* **66**, 57-80.

Berlocher, S. H. and McPheron, B. A. (1996). Population structure of Rhagoletis pomonella, the apple maggot fly. *Heredity* **77**, 83-99.

Bradshaw, W. E., Emerson, K. J., Catchen, J. M., Cresko, W. A. and Holzapfel, C. M. (2012). Footprints in time: comparative quantitative trait loci mapping of the pitcher-plant mosquito, Wyeomyia smithii. *Proceedings of the Royal Society B-Biological Sciences* **279**, 4551-4558.

Bradshaw, W. E. and Lounibos, L. P. (1977). Evolution of dormancy and its photoperiodic control in pitcher plant mosquitos. *Evolution* **31**, 546-567.

Bush, G. L. (1969). Sympatric host race formation and speciation in frugivorous flies of genus Rhagoletis (Diptera, Tephritidae). *Evolution* **23**, 237-251.

Chen, W. and Xu, W.-H. (2014). Wnt/β-catenin signaling regulates Helicoverpa armigera pupal development by up-regulating c-Myc and AP-4. *Insect biochemistry and molecular biology* **53**, 44-53.

Childs, D. Z., Metcalf, C. J. E. and Rees, M. (2010). Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. *Proceedings of the Royal Society B-Biological Sciences* **277**, 3055-3064.

Cohen, D. (1970). A theoretical model for optimal timing of diapause. *American Naturalist* **104**, 389-400.

Dambroski, H. R. and Feder, J. L. (2007). Host plant and latitude-related diapause variation in Rhagoletis pomonella: a test for multifaceted life history adaptation on different stages of diapause development. *Journal of Evolutionary Biology* **20**, 2101-2112.

Dean, R. W. and Chapman, P. J. (1973). Bionomics of the apple maggot in Eastern New York. In *Search Agric. Entomol. Geneva No. 3*. Geneva, N.Y.

Denlinger, D. L. (2002). Regulation of diapause. *Annual Review of Entomology* **47**, 93-122.

Denlinger, D. L., Yocum, G. D. and Rinehart, J. P. (2005). Hormonal control of diapause. In *Comprehensive Molecular Insect Science: Endocrinology*, vol. 3 eds. L. I. Gilbert K. Iatrou and S. S. Gill). Oxford: Elsevier.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.

Dopman, E. B., Bogdanowicz, S. M. and Harrison, R. G. (2004). Genetic mapping of sexual isolation between E and Z pheromone strains of the European corn borer (Ostrinia nubilalis). *Genetics* **167**, 301-309.

Emerson, K. J., Bradshaw, W. E. and 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**.

Feder, J. L., Chilcote, C. A. and Bush, G. L. (1988). Genetic differentiation between sympatric host races of the apple maggot fly Rhagoletis pomonella. *Nature* **336**, 61-64.

Feder, J. L. and Filchak, K. E. (1999). It's about time: the evidence for host plant-mediated selection in the apple maggot fly, Rhagoletis pomonella, and its implications for fitness trade-offs in phytophagous insects. *Entomologia Experimentalis Et Applicata* **91**, 211-225.

Feder, J. L., Hunt, T. A. and Bush, G. L. (1993). The effects of climate, host-plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of Rhagoletis pomonella. *Entomologia Experimentalis Et Applicata* **69**, 117-135.

Feder, J. L., Opp, S. B., Wlazlo, B., Reynolds, K., Go, W. and Spisak, S. (1994). Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 7990-7994.

Feder, J. L., Roethele, J. B., Wlazlo, B. and Berlocher, S. H. (1997a). Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 11417-11421.

Feder, J. L., Stolz, U., Lewis, K. M., Perry, W., Roethele, J. B. and Rogers, A. (1997b). The effects of winter length on the genetics of apple and hawthorn races of Rhagoletis pomonella (Diptera: Tephritidae). *Evolution* **51**, 1862-1876.

Filchak, K. E., Roethele, J. B. and Feder, J. L. (2001). Effects of photoperiod and light intensity on the genetics of diapause in the apple maggot (Diptera: Tephritidae). *Annals of the Entomological Society of America* **94**, 902-908.

Fraenkel, G. and Hsiao, C. (1968a). Manifestations of a pupal diapause in two species of flies, Sarcophaga argyrostoma and S. bullata. *Journal of Insect Physiology* **14**, 689-705.

Fraenkel, G. and Hsiao, C. (1968b). Morphological and endocrinological aspects of pupal diapause in a fleshfly, Sarcophaga agryrostoma. *Journal of Insect Physiology* **14**, 707-718.

Futschik, A. and Schlötterer, C. (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics* **186**, 207-218.

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D. and Brown, P. O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell* **11**, 4241-4257.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B. and Lieber, M. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* **8**, 1494-1512.

Hahn, D. A. and Denlinger, D. L. (2011). Energetics of Insect Diapause. *Annual Review of Entomology* **56**, 103-121.

Hodek, I. and Hodkova, M. (1988). Multiple role of temperature during insect diapause - a review. *Entomologia Experimentalis Et Applicata* **49**, 153-165.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57.

Irwin, J. T. and Lee, R. E. (2000). Mild winter temperatures reduce survival and potential fecundity of the goldenrod gall fly, Eurosta solidaginis (Diptera : Tephritidae). *Journal of Insect Physiology* **46**, 655-661.

Kingsolver, J. G. (1979). Thermal and hydric aspects of environmental heterogeneity in the pitcher plant mosquito. *Ecological Monographs* **49**, 357-376.

Kostal, V. (2006). Eco-physiological phases of insect diapause. *Journal of Insect Physiology* **52**, 113-127.

Lair, K. P., Bradshaw, W. E. and Holzapfel, C. M. (1997). Evolutionary divergence of the genetic architecture underlying photoperiodism in the pitcher-plant mosquito, Wyeomyia smithii. *Genetics* **147**, 1873-1883.

Lebo, M. S., Sanders, L. E., Sun, F. and Arbeitman, M. N. (2009). Somatic, germline and sex hierarchy regulated gene expression during Drosophila metamorphosis. *Bmc Genomics* **10**, 1.

Li, B. and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* **12**.

Li, Y., Huang, Y., Bergelson, J., Nordborg, M. and Borevitz, J. O. (2010). Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21199-21204.

Lin, J.-L., Lin, P.-L. and Gu, S.-H. (2009). Phosphorylation of glycogen synthase kinase-3β in relation to diapause processing in the silkworm, Bombyx mori. *Journal of Insect Physiology* **55**, 593-598.

Mathias, D., Jacky, L., Bradshaw, W. E. and Holzapfel, C. M. (2007). Quantitative trait loci associated with photoperiodic response and stage of diapause in the pitcher-plant mosquito, Wyeomyia smithii. *Genetics* **176**, 391-402.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S. and Daly, M. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303.

Micallef, L. and Rodgers, P. (2014). eulerAPE: Drawing Area-Proportional 3-Venn Diagrams Using Ellipses. *Plos One* **9**.

Michel, A. P., Sim, S., Powell, T. H., Taylor, M. S., Nosil, P. and Feder, J. L. (2010). Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9724-9.

Neilson, W. T. A. (1962). Effects of temperature on development of overwintering pupae of the apple maggot, Rhagoletis pomonella (Walsh). *Canadian Entomologist* **94**, 924-928.

Papanastasiou, S. A. and Papadopoulos, N. T. (2014). Description of Rhagoletis cerasi (Diptera: Tephritidae) Pupal Developmental Stages: Indications of Prolonged Diapause. *Journal of Insect Science* **14**, 156.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F. and Parvizi, B. (2003). TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* **19**, 651-652.

Poelchau, M. F., Reynolds, J. A., Elsik, C. G., Denlinger, D. L. and Armbruster, P. A. (2013). Deep sequencing reveals complex mechanisms of diapause preparation in the invasive mosquito, Aedes albopictus. *Proceedings of the Royal Society B-Biological Sciences* **280**.

Poupardin, R., Schoettner, K., Korbelova, J., Provaznik, J., Dolezel, D., Pavlinic, D., Benes, V. and Kostal, V. (2015). Early transcriptional events linked to induction of diapause revealed by RNAseq in larvae of drosophilid fly, Chymomyza costata. *Bmc Genomics* **16**.

Powell, J. A., Jenkins, J. L., Logan, J. A. and Bentz, B. J. (2000). Seasonal temperature alone can synchronize life cycles. *Bulletin of Mathematical Biology* **62**, 977-998.

Ragland, G. J., Denlinger, D. L. and Hahn, D. A. (2010). Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14909-14914.

Ragland, G. J., Egan, S. P., Feder, J. L., Berlocher, S. H. and 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**, 3948-3959.

Ragland, G. J., Fuller, J., Feder, J. L. and Hahn, D. A. (2009). Biphasic metabolic rate trajectory of pupal diapause termination and post-diapause development in a tephritid fly. *Journal of Insect Physiology* **55**, 344-50.

Ragland, G. J., Sim, S. B., Goudarzi, S., Feder, J. L. and Hahn, D. A. (2012). Environmental interactions during host race formation: host fruit environment moderates a seasonal shift in phenology in host races of Rhagoletis pomonella. *Functional Ecology* **26**, 921-931.

Reya, T. and Clevers, H. (2005). Wnt signalling in stem cells and cancer. *Nature* **434**, 843-850.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.

Robinson, M. D. and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25.

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

Schwarz, D., Robertson, H. M., Feder, J. L., Varala, K., Hudson, M. E., Ragland, G. J., Hahn, D. A. and Berlocher, S. H. (2009). Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot Rhagoletis pomonella. *Bmc Genomics* **10**.

Shingleton, A. W., Sisk, G. C. and Stern, D. L. (2003). Diapause in the pea aphid (Acyrthosiphon pisum) is a slowing but not a cessation of development. *BMC developmental biology* **3**, 1.

Sim, C. and Denlinger, D. L. (2008). Insulin signaling and FOXO regulate the overwintering diapause of the mosquito Culex pipiens. *Proceedings of the National Academy of Sciences* **105**, 6777-6781.

Sim, C. and Denlinger, D. L. (2013). Insulin signaling and the regulation of insect diapause. *Frontiers in Physiology* **4**.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212.

Smith, D. C. (1988). Heritable divergence of Rhagoletis pomonella host races by seasonal asynchrony. *Nature* **336**, 66-67.

Sorensen, J. G., Nielsen, M. M., Kruhoffer, M., Justesen, J. and Loeschcke, V. (2005). Full genome gene expression analysis of the heat stress response, in Drosophila melanogaster. *Cell Stress & Chaperones* **10**, 312-328.

Tammariello, S. P. and Denlinger, D. L. (1998). G0/G1 cell cycle arrest in the brain of Sarcophaga crassipalpis during pupal diapause and the expression pattern of the cell cycle regulator, proliferating cell nuclear antigen. *Insect biochemistry and molecular biology* **28**, 83-89.

Tauber, E., Zordan, M., Sandrelli, F., Pegoraro, M., Osterwalder, N., Breda, C., Daga, A., Selmin, A., Monger, K., Benna, C. et al. (2007). Natural selection favors a newly derived timeless allele in Drosophila melanogaster. *Science* **316**, 1895-1898.

Tauber, M. J., Tauber, C. A. and Masaki, S. (1986). Seasonal adaptations of insects. New York: Oxford University Press.

Taylor, F. (1981). Ecology and evolution of physiological time in insects. *American Naturalist* **117**, 1-23.

Wadsworth, C. B. and Dopman, E. B. (2015). Transcriptome profiling reveals mechanisms for the evolution of insect seasonality. *The Journal of experimental biology* **jeb-126136**.

Walsh, B. J. (1867). The apple-worm and the apple maggot. *Journal of Horticulture* **2**, 338-3434.

Williams, C. M., Chick, W. D. and Sinclair, B. J. (2015). A cross-seasonal perspective on local adaptation: metabolic plasticity mediates responses to winter in a thermal-generalist moth. *Functional Ecology* **29**, 549-561.

Williams, C. M., Hellmann, J. and Sinclair, B. J. (2012a). Lepidopteran species differ in susceptibility to winter warming. *Climate Research* **53**, 119-130.

Williams, C. M., Marshall, K. E., MacMillan, H. A., Dzurisin, J. D. K., Hellmann, J. J. and Sinclair, B. J. (2012b). Thermal Variability Increases the Impact of Autumnal Warming and Drives Metabolic Depression in an Overwintering Butterfly. *Plos One* **7**.

Williams, K. D., Busto, M., Suster, M. L., So, A. K. C., Ben-Shahar, Y., Leevers, S. J. and Sokolowski, M. B. (2006). Natural variation in Drosophila melanogaster diapause due to the insulin-regulated PI3-kinase. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15911-15915.

Wodarz, A. and Nusse, R. (1998). Mechanisms of Wnt signaling in development. *Annual review of cell and developmental biology* **14**, 59-88.

Wullschleger, S., Loewith, R. and Hall, M. N. (2006). TOR signaling in growth and metabolism. *Cell* **124**, 471-484.

Tables

Table 1. Functional categories enriched in sets of transcripts significantly differentially expressed (DE; FDR < 0.05) in single comparisons (log(Group1/Group2)), or pairs of comparisons. For single comparisons, the direction column indicates whether the test was performed on all DE transcripts ('both') or just the subsets that were up- or down-regulated. 'Same' and 'opposite' refer to tests performed on subsets of DE transcripts with the same or opposite signs for the indicated comparisons. Due to the very large number of DE transcripts in the 24h/0h comparison, enrichment tests were only performed on the significant subset having absolute values of log fold changes greater than 1 (2-fold DE).

Comparison	Direction	Category	FDR	N
Apple/Haw	both	neuron devel.	0.02	53
Apple/Haw	both	cell motion	0.06	43
Apple/Haw	both	gland devel.	0.02	29
Apple/Haw	both	ribosome biogenesis	0.02	16
Apple/Haw	up	neuron devel.	0.01	36
Apple/Haw	up	gland devel.	0.01	22
Apple/Haw	up	insect cuticle protein	0.08	17
Apple/Haw	up	respiratory syst. devel.	0.02	20
Apple/Haw	up	ectodermal gut devel.	0.02	10
Apple/Haw	up	imaginal disc devel.	0.02	38
Apple/Haw	down	**none**	NA	NA
24h/0h (LFC > 1)	both	cell morphogenesis	< 0.01	98
24h/0h (LFC > 1)	both	DNA binding	< 0.01	175
24h/0h (LFC > 1)	both	transcription	< 0.01	93
24h/0h (LFC > 1)	both	imaginal disc devel.	< 0.01	83
24h/0h (LFC > 1)	both	Zinc finger, C2H2-like	0.02	76
24h/0h (LFC > 1)	both	regulation of cell cycle	0.01	40
24h/0h (LFC > 1)	both	gland devel.	< 0.01	41
24h/0h (LFC > 1)	up	DNA binding	< 0.01	167
24h/0h (LFC > 1)	up	imaginal disc devel.	0.22	69
24h/0h (LFC > 1)	up	cell morphogenesis	< 0.01	98
24h/0h (LFC > 1)	up	transcription	< 0.01	91
24h/0h (LFC > 1)	up	Zinc finger, C2H2-like	0.15	72
24h/0h (LFC > 1)	up	regulation of cell cycle	0.01	40
24h/0h (LFC > 1)	up	gland devel.	0.01	40

24h/0h (LFC > 1)	up	cell-cell signaling	0.01	48
24h/0h (LFC > 1)	down	endopeptidase inhibitor activity	< 0.01	17
24h/0h (LFC > 1)	down	aging	0.06	15
24h/0h (LFC > 1)	down	MADF domain	0.02	13
Apple/Haw and 24h/0h	both	neuron devel.	< 0.01	39
Apple/Haw and 24h/0h	both	respiratory system devel.	0.06	18
Apple/Haw and 24h/0h	both	stem cell division	0.04	10
Apple/Haw and 24h/0h	both	gland devel.	0.01	23
Apple/Haw and 24h/0h	both	ribosome biogenesis	0.02	12
24h/0h and 48h/24h	both	ribosome biogenesis	< 0.01	32
24h/0h and 48h/24h	both	mRNA binding	< 0.01	27
24h/0h and 48h/24h	both	nucleotide binding	< 0.01	83
24h/0h and 48h/24h	both	helicase activity	< 0.01	17
24h/0h and 48h/24h	both	RNA methyltransferase activity	< 0.01	8
24h/0h and 48h/24h	both	devel.al growth	0.02	13
24h/0h and 48h/24h	opposite	ribosome biogenesis	< 0.01	32
24h/0h and 48h/24h	opposite	mRNA binding	< 0.01	24
24h/0h and 48h/24h	opposite	helicase activity	< 0.01	16
24h/0h and 48h/24h	opposite	nucleotide binding	0.01	64
24h/0h and 48h/24h	opposite	RNA methyltransferase activity	< 0.01	7
24h/0h and 48h/24h	same	**none**	NA	NA

Figures

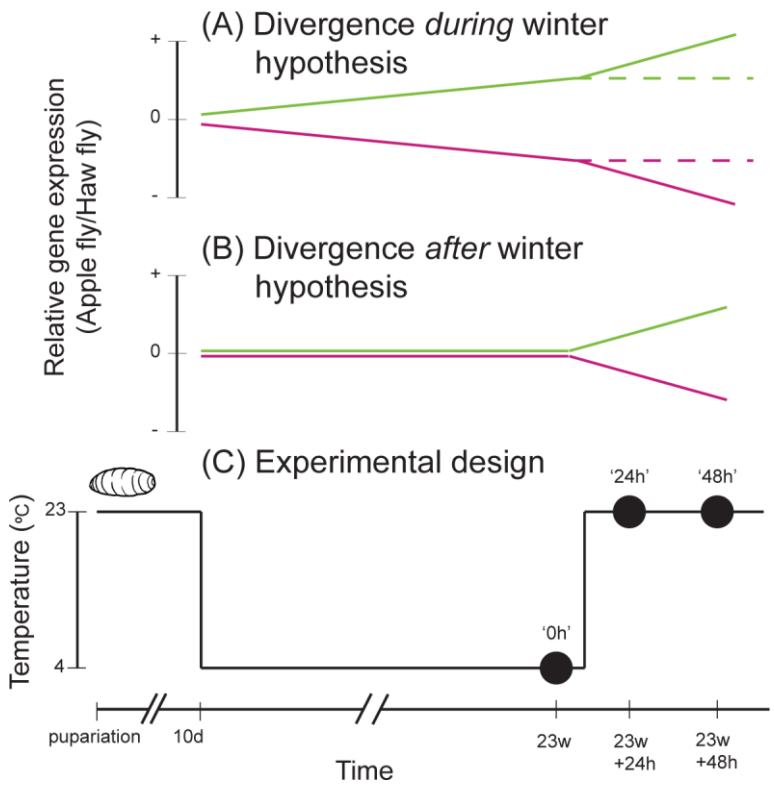

Figure 1. Conceptual illustration of the *during* and *after* winter hypotheses for apple vs. hawthorn fly race divergence in gene expression and the experimental design. Transcripts regulating earlier diapause termination in apple relative to hawthorn flies could be differentially regulated between populations during winter (A), or only after exposure to warmer temperatures necessary to begin post-diapause development (B). Green and pink represent up- and down-regulation transcripts in apple *relative to* hawthorn flies and not absolute values, so the key feature is when the shifts in transcription levels occur between the host races, i.e., the fork in up and down regulation can be seen prior to 0h in the *during winter* hypothesis and extends afterward, whereas it is initiated only after the 0h time point in the *after winter*

hypothesis. Note that both positive and negative values of this relative measure may reflect up- *or* down-regulation of transcripts in one population compared to the other. Solid and dashed lines in the upper panel denote accelerating and constant differences between populations following the temperature shift. Experimental design - upon pupariation, apple and haw fly pupae were exposed to common, simulated overwinter conditions, then sampled for RNAseq at three time points, one at the end of winter (0h), one at 24h after the end of winter, and one 48h after transfer to warmer conditions (C).

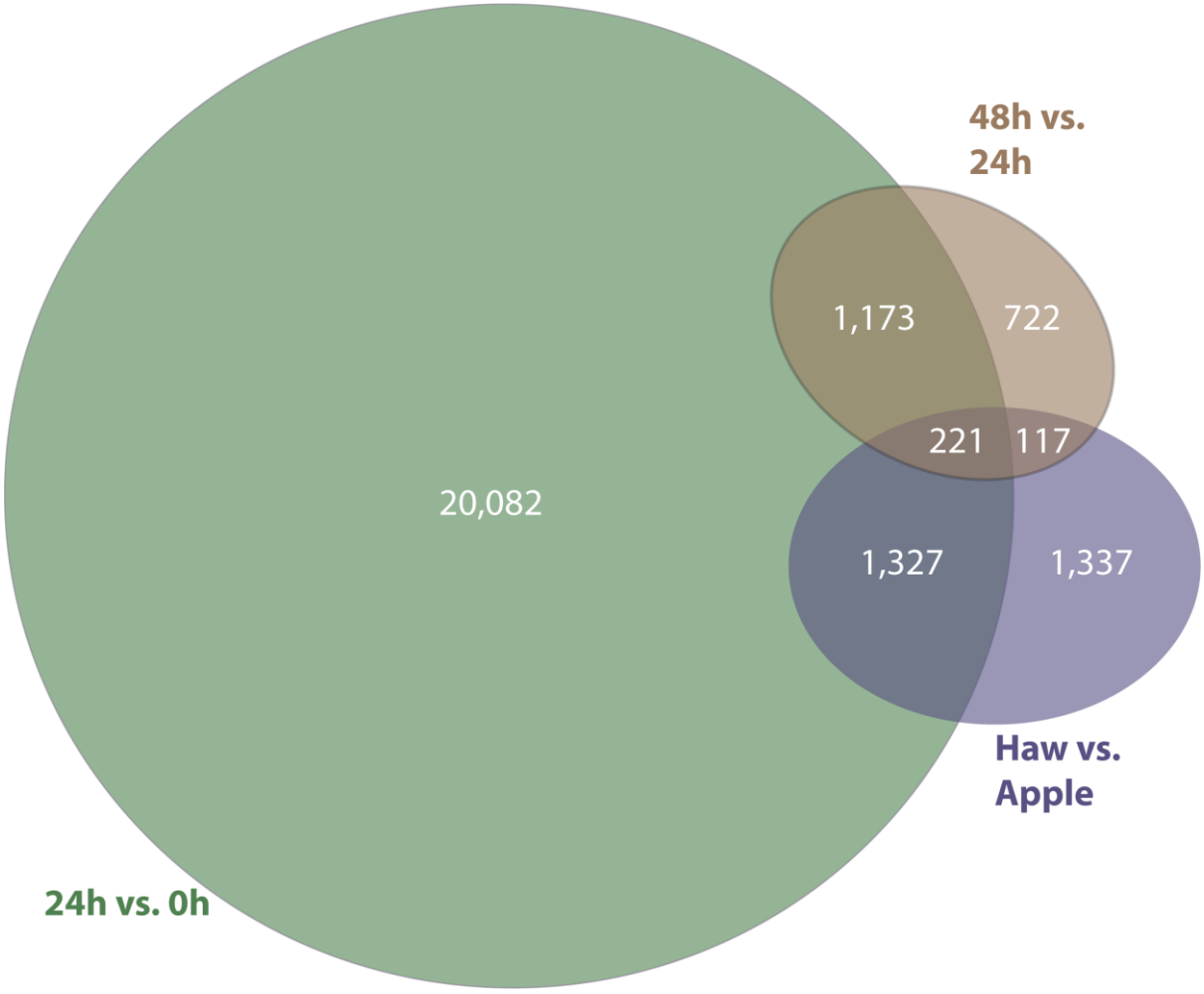

Figure 2. Venn diagram depicting numbers of significantly differentially expressed transcripts (FDR < 0.05 as determined from linear models) identified in three pairwise comparisons as inferred by generalized linear models of transcript counts including parameters for sampling time point and host race but no interaction; two comparisons between sampling time points (0h vs. 24h and 24h vs. 48h; Fig. 1) and one comparison between the apple and hawthorn host races (populations). Overlapping regions represent sets of transcripts that were differentially expressed in two or more comparisons.

Journal of Experimental Biology • Advance article

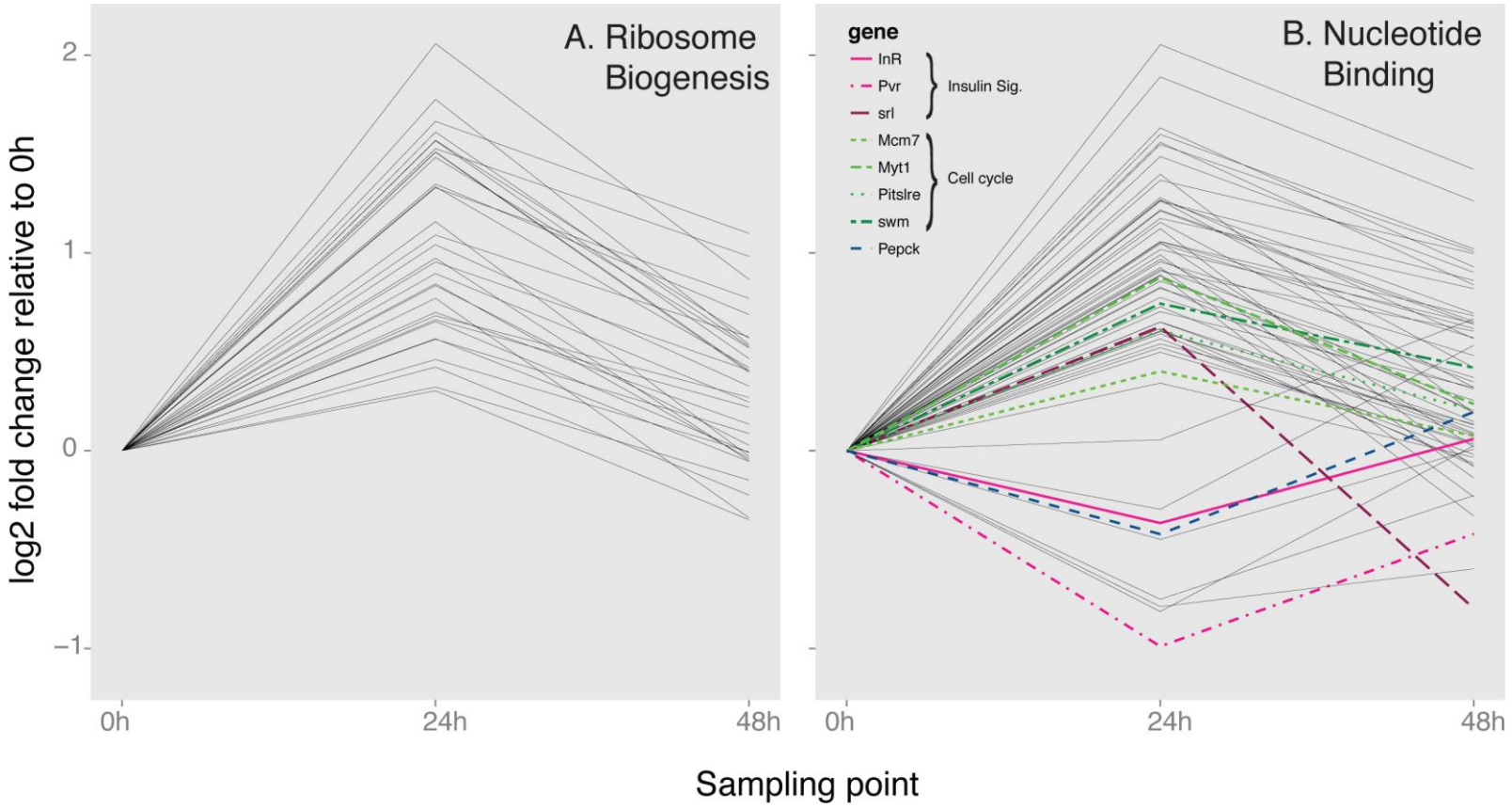
A. Ribosome Biogenesis
B. Nucleotide Binding
log2 fold change relative to 0h
2
1
0
−1
0h
24h
48h
gene
InR
Pvr
srl
Insulin Sig.
Mcm7
Myt1
Pitslre
swm
Cell cycle
Pepck
Sampling point

Figure 3. Network diagram of all transcripts significantly differentially expressed between host races (populations; significant host effect in the GLM, FDR < 0.05, and non-significant host x time point interaction) connected to core *wnt* and insulin signaling pathway genes by at least one edge. Edges represent the existence of one or more gene-gene interactions in the Drosophila Interactions Database (DroID). Gene names are from the *Drosophila melanogaster* annotation v6.10 referenced in supplement S1.

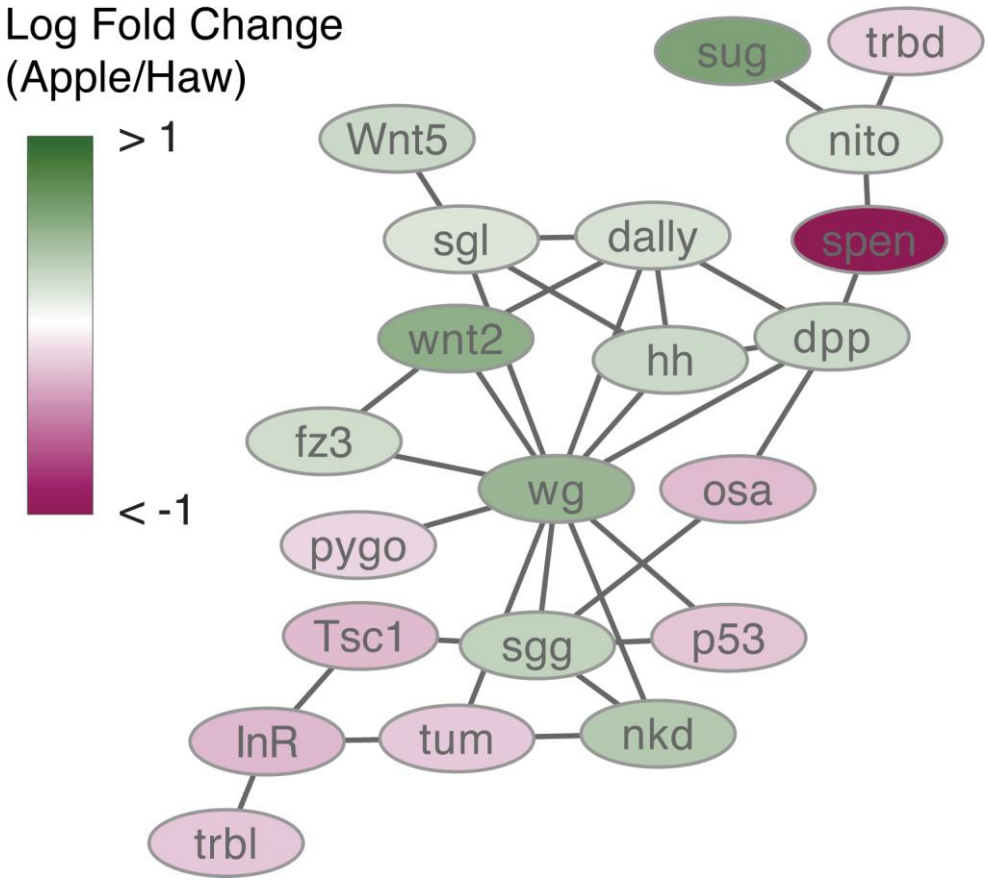

Figure 4. Time series of transcript abundance relative to expression at the first sampling point ('0h') for transcripts consistent with a pattern of post-winter suppression. Values are estimates of Log_2(expression at time point/expression at 0h) obtained via post-hoc contrasts of GLM parameters. These include transcripts from two functional categories that were enriched in the set of all transcripts that were oppositely differentially expressed (up-regulated at 24h but down-regulated at 48h, or vice versa) between 1) '0h' and '24h', and 2) '24h' and '48h': A, Ribosome Biogenesis, and B, Nucleotide Binding.

Journal of Experimental Biology • Advance article