Physiological Entomology (2017) **42**, 246–256 DOI: 10.1111/phen.12193

Comparative transcriptomics support evolutionary convergence of diapause responses across Insecta

GREGORY J. RAGLAND[1,2] and EVAN KEEP[1,2]

[1]Department of Integrative Biology, University of Colorado – Denver, Denver, Colorado, U.S.A. and [2]Department of Entomology, Kansas State University, Manhattan, Kansas, U.S.A.

Abstract. Diapause is a common phenotype that is broadly phylogenetically dispersed across Insecta and appears to have multiple evolutionary origins. Nevertheless, there are clear commonalities in diapause regulation across insect taxa. In the present study, we report a meta-analysis of diapause whole transcriptomic data sets from 11 different insect species that addresses three questions: (i) how similar are whole-transcriptome diapause responses across species within and across different diapause life-cycle stages; (ii) do the most closely-related species demonstrate the most similar diapause responses within and across diapause life-cycle stages (the existence of phylogenetic signal); and (iii) is there a core set of regulatory genes that universally associate with insect diapause at the transcript level? The included species are mostly Dipterans ($n = 9$), plus one species each from Lepidoptera and Hymenoptera. The group includes multiple species that enter diapause as either larvae, pupae or adults. We establish a set of 4791 orthologous transcript sequences with expression data acquired from published studies of diapause transcriptomes. We find no support for phylogenetic signal. Transcriptomic responses of nondipterans clustered within rather than outside of Dipteran responses. However, expression profiles do tend to cluster with the diapause stage of developmental arrest, although this pattern is only moderately supported. We identify a statistically significant set of 542 orthologues (11% of all orthologues) that are commonly differentially regulated during diapause across all included species. From this core set, we identify candidate genes participating in circadian rhythmicity, insulin signalling and Wnt signalling, which are pathways previously associated with insect diapause development. Clustering relationships among species are most consistent with evolutionary convergence of the shared transcriptomic response, although we are unable to determine whether this reflects convergence of the diapause initiation, maintenance or termination phases.

Key words. Comparative transcriptomics, convergence, diapause, dormancy, evolution.

Introduction

Insect diapause, a developmentally regulated dormant life stage, is a pervasive, yet highly diverse phenotype that occurs in members of every major taxonomic group. Diapause evolves as a response to seasonal variability in climatic factors or resources and occurs in both temperate (mainly overwinter diapause) and tropical (mainly dry season diapause) species. Although most species enter diapause in a single, stereotypical life-cycle stage, different species diapause in all possible life-cycle stages, with limited phylogenetic conservation (Tauber *et al.*, 1986). Diapause represents an alternative developmental pathway characterized by the initiation, maintenance and termination phases (Kostal, 2006). The maintenance phase is always characterized by a deceleration of the normal developmental programme, although this may be achieved via different strategies (e.g. arrest during different stages of the cell cycle) (Nakagaki *et al.*, 1991; Tammariello & Denlinger, 1998). Diapausing insects also typically remodel metabolism, although diapause in different

Correspondence: Gregory J. Ragland, Department of Integrative Biology, University of Colorado – Denver, 1115 Arapahoe St., Denver, Colorado 80217, U.S.A. Tel.: +1 303 556 5970; e-mail: gregory.ragland@ucdenver.edu

Table 1. Select metadata for each included transcriptomic study (full details are available in the Supporting information, Table S1).

Order	Species	Diapause stage	Reference	Data type	Replication
Diptera	*Aedes albopictus*	Larva (pharate)	Poelchau *et al.* (2013)	RNAseq	3 pools
Diptera	*Bactrocera minax*	Pupa	Dong *et al.* (2014)	RNAseq	1 pool
Diptera	*Delia antiqua*	Pupa	Hao *et al.* (2016)	RNAseq	1 pool
Diptera	*Rhagoletis pomonella*	Pupa	Ragland *et al.* (2011)	Microarray	4 ind.
Diptera	*Sarcophaga crassipalpis*	Pupa	Ragland *et al.* (2010)	Microarray	4 pools
Diptera	*Drosophila montana*	Adult	Parker *et al.* (2016)	RNAseq	3 pools
Diptera	*Culex pipiens*	Adult	Kang *et al.* (2016)	RNAseq	3 pools
Diptera	*Chymomyza costata*	Larva	Poupardin *et al.* (2015)	RNAseq	3 pools
Diptera	*Drosophila melanogaster*	Adult	Baker and Russell (2009)	Microarray	4 pools
Lepidoptera	*Ostrinia Nubilalis*	Larva	Wadsworth & Dopman (2015)	RNAseq	5 pools
Hymenoptera	*Megachile Rotundata*	Pupa	Yocum *et al.* (2015)	RNAseq	3 ind.

For each replicate sample, studies used homogenates either from multiple pooled (pools) of individuals or single individuals (ind.).

species is associated with different strategies, including a general depression in energy consumption, increased nutrient storage, a shift to anaerobic catabolism and the production of metabolites that lower supercooling points (Hahn & Denlinger, 2011). Even within species, there is often wide variation in aspects of the diapause phenotype within and among populations (Danilevsky *et al.*, 1970; Bradshaw & Lounibos, 1977) that can quickly evolve in response to changing environmental conditions (Gomi & Takeda, 1996; Lounibos *et al.*, 2003).

Based on this broad variation in diapause development and its potential for rapid evolution, it is generally accepted (although not explicitly tested) that diapause has evolved many times during the evolutionary diversification of insects (Danks, 1987). Yet, there are many physiological commonalities in diapause responses across insects. For example, the abundance of transcripts and protein products of genes in important regulatory pathways such as insulin and Wnt signalling is associated with diapause development in a number of studies (Sim & Denlinger, 2013; Chen & Xu, 2014). Genes in the canonical circadian rhythm feedback loops also exhibit expression differences and segregating natural variation in diapause development (Goto, 2013). Likewise, there are commonalities in the hormonal control of diapause, particularly among species that diapause in the same life-cycle stage (Denlinger, 2002).

Although all of the above examples provide evidence for select commonalities in diapause responses across species, there are no published studies applying a statistical comparison of many species simultaneously. Targeted studies of small sets of genes or hormones that are necessarily specific for revealing mechanisms within species are not well suited to make broad, cross-species comparisons. However, with the advent of transcriptomics, we can now rapidly acquire data on the mRNA expression levels of many thousands of genes. Transcriptomics certainly has limitations, principally that transcript levels do not perfectly correlate with protein levels, nor do they perfectly predict levels of metabolites. However, transcriptomes currently represent the most tractable resource available with respect to the simultaneously tracking of changes in many physiological processes. Conveniently, diapause is one of the most common phenotypic targets of transcriptomics in insects and other invertebrates.

Ragland *et al.* (2010) report a comparison of diapause transcriptomic responses among two Dipterans and the dormant (dauer) stage of the worm *Caenorhabditis elegans*. This three-species comparison suggests largely species-specific responses with a handful of genes commonly differentially expressed between dormant and nondormant individuals sampled from comparable life-cycle stages. However, all three of these species undergo dormancy in different stages (pupal, adult reproductive and larval diapause) and *C. elegans* is a very distant evolutionary relative of insects (approximately 500–600 million years). With the rapid development of transcriptomic resources in non-model organisms through RNA sequencing, many more insect data sets characterizing diapause transcriptomes are now available. In the present study, we apply a comparative meta-analysis to a subset of these data sets that includes multiple representative species for each of three diapause life-cycle stages (larval, pupal and adult reproductive). This comparison includes mostly Dipteran species, but also one Lepidopteran and one Hymenopteran. We address three primary questions: (i) how similar are whole-transcriptome diapause responses across species within and across different diapause life-cycle stages; (ii) do the most closely-related species demonstrate the most similar diapause responses within and across diapause life-cycle stages (i.e. the existence of phylogenetic signal); and (iii) is there a core set of regulatory genes that universally associate with insect diapause at the transcript level?

Materials and methods

Included data sets

We retrieved transcriptome-wide gene expression data from studies published on 11 insect species, with most (eight of 11) performed using whole-body RNA extractions (Table 1; see also Supporting information, Table S1). Approximate phylogenetic relationships among the included species are depicted in Fig. 1; the most closely-related species diverge by 35–40 million years (Vieira *et al.*, 1997; Ding *et al.*, 2015), whereas the most recent common ancestor of Diptera and Lepidoptera and Hymenoptera occurred an estimated 300 and 320 years ago, respectively (Misof *et al.*, 2014). Our goal was to use a standardized measurement that could be compared across all taxa in the analysis.

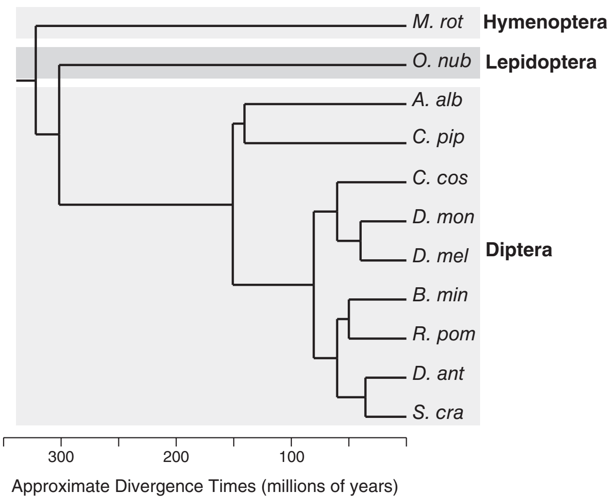

Fig. 1. Approximate phylogenetic relationships among the 11 insect species included in the analysis. Topology and branch lengths constructed based on Ayala (1997), Peterson *et al.* (2009), Reidenbach *et al.* (2009), Misof *et al.* (2014) and Ding *et al.* (2015).

Based on commonalities in experimental design across studies, we settled on the log (base 2) fold changes in gene expression comparing diapausing with nondiapausing individuals, both in approximately comparable life-cycle stages. Different studies achieve this comparison using different methods, such as comparing diapause with post-diapause or diapause with pre-diapause, and the methods of diapause induction varied because some species initiate diapause in response to photoperiod, some in response to temperature, and some in response to the combination of the two. Moreover, there was broad variation among studies in the statistical methods used to estimate log-fold changes. For all but one study, we used the estimates from the original study and, in that one case (Parker *et al.*, 2016), we applied a general linear model implemented in the edgeR package in R (Robinson *et al.*, 2010) to obtain a log-fold change estimate comparable with those from other studies (see Supporting information, Script S1). All of these attributes and other relevant metadata for all included studies are also summarized in the Supporting information (Table S1). We also note that some of these studies include more extensive experimental designs incorporating time series and species/strain comparisons that were not comparable across studies and thus are not included in this analysis. To establish orthology of transcripts across species, we retrieved sequence data from either the transcriptome assembly or the genome assembly (all studies relied on *de novo* transcriptome assemblies except *Drosphila melanogster*) to which RNAseq data were mapped or from which microarray probes were designed.

Additional diapause transcriptomic studies are available, although we excluded a number of studies because of (i) a lack of publically available data or (ii) difficulty in assessing the level of biological replication. We also excluded several studies because it was difficult to obtain a clear diapause versus nondiapause comparison conforming with the criteria above or because a study used tissue-specific homogenates when another study using whole-body homogenates was available. Tissue-specific studies are preferable for a number of reasons, although almost all available studies use whole body homogenates; thus whole-body studies were most appropriate for cross-species comparisons.

Orthology assignment

Ideally, orthology assignments would be made using the orthoDB method that considers the evolutionary relationships among all included species simultaneously (Waterhouse *et al.*, 2013). However, there are several challenges posed by fragmented transcriptome assemblies that make this method impractical. First, different transcriptomes may include different fragments of the same transcript that do not overlap. Second, it is often difficult to distinguish splice variants (transcripts mapping to the same gene model) from paralagous transcripts (transcripts mapping to different gene models with similar sequences) in a fragmented transcriptome. Both of these issues make it difficult to construct the multiple sequence alignments necessary to use, for example, in orthoDB.

As an alternative strategy, we used a hybrid approach between two methods to establish pairwise orthology between each species and *D. melanogaster*, which is by far the most complete, error-free and best-annotated insect genome. We assume that evidence for pairwise orthology to *D. melanogaster* across all 10 non-*melanogaster* species constitutes evidence for one-to-one orthology across all included species. Given the fragmented assemblies, we first removed redundancy in the transcriptome assemblies (different splice variants or different, non-overlapping fragments mapping to the same gene model) by consolidating transcripts with Blastx hits with an *e*-value < 0.05 to the same *D. melanogaster* gene model, and no secondary hits to other gene models with an *e*-value $< 1 \times 10^{-105}$. This latter threshold was established by blasting *D. melanogaster* transcripts against themselves to determine an *e*-value above which annotation to a similar paralogue was unlikely. In the first method, we assigned orthology based on reciprocal best hits (RBH) using Blastx searches against the *D. melanogaster* protein database. This is probably the most common method for orthology assignment (Tatusov *et al.*, 1997), although it lacks some sensitivity because it is based on local alignments that may be much shorter than global alignments of whole transcripts (Wall & Deluca, 2007). Therefore, we also applied a second method, reciprocal smallest distance (RSD), where orthology is assigned based on reciprocally smallest evolutionary distances as estimated from global, in-frame alignments (Wall & Deluca, 2007). For transcripts that were collapsed into single gene models, we used only the longest inferred isoform, and we used ESTscan (Iseli *et al.*, 1999) to translate transcript to protein sequences. Finally, we merged the results of the two assignment methods, retaining all sequences where RBH or RSD assigned orthology.

Overall, our strategy for orthology assignment is less conservative than the orthoDB approach, although our goal was to maximize accuracy of orthologous relationships at the same time as maximizing the number of gene models included in the analysis. There are likely some miss-assignments of orthology (associating paralogues with each other) across species, as well as instances of collapsing transcripts originating from paralagous genes within each species. However, these miss-assignments are more likely with highly similar sequences, and highly similar paralogues tend to retain very similar functions, although that relationship does break down at greater evolutionary distance (Gabaldón & Koonin, 2013). Furthermore, our primary goal was to use whole transcriptome signals to test for physiological (i.e. functional) similarity among diapause responses. Errors in orthology assignments do not bias our comparisons that leverage hundreds to thousands of transcripts to make inferences about average responses.

Species clustering

Clustering relationships of expression profiles among species were assessed using hierarchical clustering implemented in the R package 'pvclust', using the 'complete' agglomerative algorithm and calculating distance as 1 – absolute value of the Pearson correlation. To assess confidence in the clustering relationships, we performed 1000 bootstrap re-sampled replicates (the 'BP' method in pvclust).

Gene clustering

We attempted to assess clustering relationships of expression profiles among genes using a number of different distance calculations and clustering methods, including normal mixture modelling (Fraley & Raftery, 1999) and weighted gene coexpression networks (Langfelder & Horvath, 2008). However, clustering relationships were shallow and unstable. The resulting co-expression clusters had no clear biological interpretations and were often driven by extreme values in one or two species (for WGCNA clustering, see Supporting information, Fig. S1; for gene expression patterns per module summarized by eigengenes, see Supporting information, Table S2). Given these difficulties with noisy data, we took a simpler, categorical approach that (i) identifies subsets of genes exhibiting similar directionality across species and (ii) provides a robust statistical test for similarity of expression profiles across species. First, we determined the subset of genes exhibiting all positive and all negative log-fold changes ('all up-' and 'all down-' regulated) across all species in the data set, allowing for a maximum of one mismatch in sign. Log-fold changes reflected the ratio of diapause over nondiapause expression, such that 'all up-' and 'all down-' regulated genes were up- and down-regulated in diapause compared with nondiapause, respectively. Next, we assessed statistical significance by calculating the percentile of the point estimate (i.e. the number of genes in the 'all up-' and 'all down-' regulated categories) in a random empirical distribution generated using 1000 permutations randomly shuffling rows (genes) independently within each column (species). This procedure estimates the number of 'all up-' and 'all down-' regulated genes expected by chance alone, given the empirical multivariate distribution of expression values. Point estimates less than or greater than the 0.5% and 99.5% percentiles, respectively, were assigned statistical significance (two-tailed test, $\alpha = 0.01$).

Enrichment analysis

Gene lists identified as 'all up-' or 'all down-' regulated using the procedure described above were tested for enrichment of functional categories (including Gene Ontology, KEGG pathways, INTERPRO domains, etc.) using DAVID (Huang *et al.*, 2008), applying a Benjamini and Hochberg multiple correction to enrichment *P*-values.

Results

Orthologous gene expression data set

Allowing for up to three species lacking a representative orthologue, we identified 5614 orthologues based on the RBH strategy and 3379 orthologues based on the RSD strategy, yielding 6657 unique total orthologues. We further filtered this data set to allow a maximum of three missing expression values (log-fold change, diapause/nondiapause) per gene across species, yielding a final 11 species by 4791 genes data set.

Species clustering

Clustering of the different insect species based on expression profiles for all included orthologues revealed several patterns. First, rather than clustering as outgroups, the two nondipteran species were nested within two different clusters of dipterans, although branch lengths at this level were short and bootstrap support was low (Fig. 2). Second, only the three Drosophilid species (*D. melanogaster*, *Drosophila montana* and *Chymomyza costata*) clustered together according to taxonomic affinity. Two of the most closely-related species, *D. melanogaster* and *D. montana*, clustered most closely together of any two species in the analysis on a relatively long branch with high bootstrap support (100% of bootstrap replicates included this pairing). Third, of the three represented diapause stages, only pupal-stage diapausing species grouped in a single cluster, although branches were shallow and bootstrap support was low. This group included the relatively closely-related Muscoid (*Sarcophaga crassipalpis* and *Delia antiqua*) and Tephritid (*Rhagoletis pomonella* and *Bactrocera minax*) flies and the hymenopteran *Megachile rotundata*. Even within this group, species did not cluster according to taxonomy, with the Muscoid and Tephritid flies not clustering most closely with each other as they do in a phylogenetic tree (Fig. 1).

Gene clustering

As noted in the Materials and methods, approaches to clustering based on continuous log-fold change value yielded very shallow clustering relationships of limited biological meaning (see

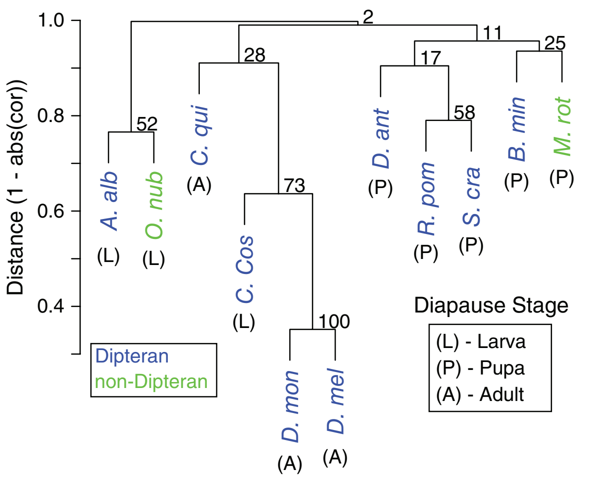

Fig. 2. Hierarchical clustering of included insect species based on log-fold changes (diapause versus nondiapause) in transcript expression across all orthologous genes, with bootstrap (percentage) values indicated at each node. Distance was calculated as one minus the absolute value of the Pearson correlation. Dipterans are depicted in blue; non-dipterans are depicted in green. The stereotypical diapause life stage of each species is indicated by the larval, pupal and adult icons. [Colour figure can be viewed at wileyonlinelibrary.com].

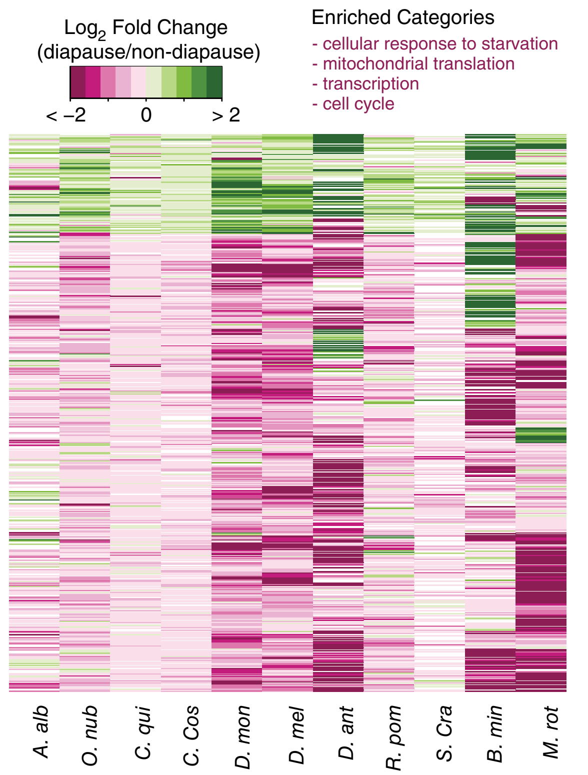

Fig. 3. Heatmap of all 542 genes majority up- or down-regulated (allowing one mismatch) in diapause relative to nondiapause in all 11 included insect species. Categories represent functional groups statistically enriched in the down-regulated (pink) gene lists; there was no significant enrichment in the up-regulated gene list. [Colour figure can be viewed at wileyonlinelibrary.com].

Supporting information, Fig. S1 and Table S2). Thus, we applied a categorical approach, identifying genes that were either all up-regulated or all down-regulated across all species, allowing for a maximum of one mismatch. This criterion identified 98 'all up-' and 444 'all down-' regulated genes (Fig. 3; see also Supporting information, Table S3). Both of these values were statistically different (larger) from random expectations; none of the 1000 random permutations resulted in values as extreme (median random expectations: 42 'all up-' and 281 'all down-'). The greater number of 'all down-' versus 'all up-' regulated transcripts reflects the overall greater abundance of negative log-fold changes, or down-regulation in diapause versus nondiapause.

Given the above observation that the diapause transcriptomic response for all species with a pupal diapause clustered together, we also used the same permutation test applied above to investigate whether there were more 'all up-' and 'all down-' regulated genes than predicted by chance alone, for only the pupal-diapausing species. The 613 total genes identified (the combined set of 'all up-' and 'all down-' regulated genes) were not significantly more than the random expectation ($P = 0.32$).

Gene enrichment analysis

Enrichment analysis identified no statistical enrichment in the 'all up-' regulated gene list, although it identified a number of functional categories statistically enriched in the 'all down-' regulated gene list. The full results are provided in the Supporting information (Table S4). The categories identified in Fig. 3 summarize the broadest and most statistically significant categories directly related to development and metabolism, including the cellular response to starvation, mitochondrial translation (largely containing mitochondrial ribosomal proteins), transcription and cell cycle.

Discussion

The combination of expression data collected using a diverse set of methods, analyzed using diverse statistical models and performed in a diverse set of laboratories produces a fairly noisy data set in which the distribution of log-fold changes has longer tails extending to more extreme values in a few of the species, even after scale normalization (see Supporting information, Fig. S2). This presents less of a problem for a cross-species clustering analysis because these distributions do not bias correlations in expression pattern across species. However, the data structure causes severe bias in the estimation of correlations among genes, with genes having extreme values in two or more species tending to cluster together (see eigengene patterns in the Supporting information, Table S2). Thus, in the present study, we apply traditional clustering approaches to cluster species, whereas we apply our categorical approach in order to identify sets of genes up- or down-regulated across the majority of included species. This latter approach has the advantage of producing patterns with simple biological interpretations, identifying genes and associated functional categories/pathways that appear to play a role in insect diapause, with the same directionality across species.

The inherent, technical noise in the data set as a result of diverse experimental and statistical approaches also limits inferences about the species specificity of the diapause transcriptome. Species-specific transcriptomic signatures could reflect not only the multiple evolutionary origins of diapause, but also study-to-study differences in approach and analysis that we cannot completely eliminate with a standardized metric. Likewise, noise in the data may obscure phylogenetic signal. Thus, we temper any interpretations based on dissimilarities.

Clustering based on taxonomy and diapause stage

Clustering of species expression profiles across all orthologues generally aligns better with the stage of diapause than with phylogenetic affinity. Larval and adult diapausing species do not all cluster together, although all pupal diapausing species cluster in a single group, albeit with relatively weak bootstrap support. Even within this group of species with a common diapausing life stage, species do not cluster according to their phylogenetic relationships. For example, the two Muscoid flies (*S. crassipalpis* and *D. antiqua*) cluster more closely with one of the two Tephritid flies (*R. pomonella* and *B. minax*) than with each other (Fig. 2). Moreover, *M rotundata*, a Hymonopteran pupal diapauser, nests within the Dipteran pupal diapause group.

Clustering based on diapause stage rather than phylogenetic similarity is consistent with the hypothesized multiple evolutionary origins of diapause across insecta, where, for example, different taxa have evolved independently to diapause as pupae. As noted above, the noisiness inherent in the data may overwhelm any phylogenetic signal. Even if such a signal exists, these results indicate that the diapause transcriptome is better predicted by the stage of diapause, at least in this data set. Most previous reviews settle on the same conclusion (Danks, 1987; Tauber *et al.*, 1986) based on the observations that (i) most major orders of insects have members that diapause in all different life-cycle stages; (ii) most species exhibit high levels of intraspecific variation in diapause regulation, and some species may even diapause in multiple life-cycle stages (Sims, 1982); and (iii) diapause responses are lost rapidly during laboratory evolution (Kingsolver & Nagle, 2007) and during invasion of tropical habitats (Lounibos *et al.*, 2003).

There are arguments for single origins of diapause at shallow levels of phylogenetic divergence, assuming that similarity in diapause stage is consistent with a single origin. Carey (1994) reports an analyses of diapausing species within Lycaenid butterflies and shows a weak but detectable signal of conservation of diapause stage within taxonomic clades. Similarly, almost all species in the genus *Drosophila* that diapause do so in the adult stage. In the present study, *D. melanogaster* and *D. montana* have highly similar expression profiles for adult reproductive diapause, which is consistent with a single evolutionary origin. However, some of this similarity is probably driven simply by the comparison being between a fly with and without developed ovaries (Zhao *et al.*, 2015; Kankare *et al.*, 2016). Goto *et al.* (1999) report that diapause responses have diversified dramatically over the approximately 15 million years, separating members of the *obscura* group of Drosophilids (Gao *et al.*, 2007); the group includes species that diapause as adults, diapause as pupae, and do not diapause at all, and these phenotypes are interspersed across phylogenetic relationships. This is a substantially shorter divergence time than the estimate of 40 million years ago for the most recent common ancestor of *D. melanogaster* and the *virilis* group containing *D. montana* (Vieira *et al.*, 1997). No studies that we are aware of (including the present study) are able to test definitively whether diapause responses across any two species are of common evolutionary origin.

Transcriptomic responses and the 'shared toolkit'

An alternative explanation for similarity in diapause expression profiles is that, during independent evolution of a diapause response, species tend to converge on similar physiological, regulatory machinery, tracing evolutionary paths of least resistance (Schluter, 1996). This is supported by the increasingly common observations of a shared 'genomic toolkit' or a set of transcripts differentially expressed during diapause in several species (Ragland *et al.*, 2010; Poelchau *et al.*, 2013; Amsalem *et al.*, 2015). In the present study, our identification of gene sets up- and down-regulated in the majority of the 11 insect species and three represented diapause stages provides statistical evidence for shared transcriptomic profiles across insect diapause. Ragland *et al.* (2010) previously identified a set of 10 metabolically and developmentally related genes that are more than two-fold differentially expressed during diapause in two fly species (*D. melanogaster* and *S. crassipalpis*) and one worm species (*C. elegans*) as candidate members of the universal diapause 'toolkit'. Surprisingly, none of these 10 are represented in the 'all up-' and 'all down-' regulated gene sets identified in the present study. Nevertheless, the current analysis identifies a robust signal of shared transcriptional responses containing member genes that fit into predictable functional categories. Indeed, the 'all down-' regulated gene set matches well with the most conspicuous phenotype associated with diapause: developmental suppression (Denlinger, 2002). This set is enriched in 'cellular response to starvation', containing mainly genes involved in developmental processes, transcription and translation. The category 'mitochondrial translation' is also enriched, reflecting mainly down-regulation of mitochondrial ribosomal proteins, as are the categories 'transcription' and 'cell cycle'. Down-regulation of these categories is consistent with the suppression of gene/protein expression, cell division and tissue proliferation. There is no overt signature of metabolic depression or remodelling (the other conspicuous phenotype associated with diapause) in the enrichment analysis; major enzymes in the glycolysis/gluconeogenesis pathway differentially regulated in some diapause responses (Ragland *et al.*, 2010; Hahn & Denlinger, 2011) are not up- or down-regulated across the majority of species in the current analysis.

The signal of shared transcriptomic responses is clear. However, does this signal reflect convergence of developmental pathways that initiate/terminate diapause or does it reflect metabolic remodelling and/or cessation of growth associated with diapause maintenance? The hierarchical model of Losos (1992) predicts that evolutionary convergence is more likely for a given phenotype than for the specific developmental processes that produce that phenotype. In the sense that diapause initiation and termination developmentally transition into and out of diapause maintenance, we might expect that evolutionary convergence would be most likely in the diapause maintenance stage. In other words, there may be many developmental pathways that lead to similar outcomes of, for example, reduced cell cycling, which is the most prominent feature of the shared transcriptomic response that we describe in the present study. Indeed, decades of cancer research illustrate the diversity of mechanisms that can lead to similar patterns of cell cycling in mammalian species (Heng *et al.*, 2010). Moreover, the phenotypic outcome of suppressed growth may follow from the transduction of diverse combinations of environmental signals, and thus different regulatory pathways during diapause.

There is evidence for marked variability in several core developmental pathways and hormonal cascades commonly influencing the initiation and termination of diapause in many insect species. For example, an increased diapause hormone titre induces embryonic diapause in *Bombyx mori* (Saturniidae), but terminates pupal diapause in several Noctuid moths (Zhang *et al.*, 2004). Similarly, although insulin signalling is associated with diapause regulation in multiple taxa (Sim & Denlinger, 2013), the directionality of transcriptional changes in insulin pathway genes is variable among species, at least at the transcriptional level (Meyers *et al.*, 2016). The diversity of environmental conditions and developmental time points that make up the current data set preclude distinguishing between shared initiation, maintenance and termination responses. However, transcriptomic approaches could address this question in principle if they were applied in a systematic experimental design. This would involve surveying tissue-specific expression at multiple time points during diapause development within a taxonomic group in which all members diapause in the same stage and respond to similar combinations of environmental cues.

Table 2. Candidate genes exhibiting majority (maximum of one mismatch) up- or down-regulation in diapause versus nondiapause in all 11 insect species included in the analysis.

Direction of regulation	Function/pathway	Gene name
Up	Circadian rhythm	*Pdp1*
Up	Wnt	*spen*
Up	Insulin/tor	*melt*
Up	Insulin/tor	*Nlaz*
Up	Insulin/tor	*sga*
Up	Insulin/tor	*rhea*
Down	Insulin/tor	*Lnk*
Down	Insulin/tor	*Nprl3*
Down	Insulin/tor	*ns1*
Down	Wnt	*lgs*
Down	Wnt	*lqfR*
Down	Wnt	*CkIIalpha*
Down	Wnt	*nmo*
Down	Wnt	*Roc1a*
Down	Wnt	*SoxN*
Down	Ecdysone signalling	*EcR*
Down	Ecdysone signalling	*Hr78*
Down	Ecdysone synthesis	*phm*

The pathway or function in which each gene participates and the direction of regulation (up- or down-regulated in diapause relative to nondiapause) are included.

Candidate genes

Although the 'all up-' and 'all down-' regulated gene sets likely contain false positives (Type I errors), it provides a valuable resource for comparisons with future transcriptomic studies and a list of candidate genes involved in the diapause response. As discussed above, we cannot currently distinguish transcriptomic changes associated with diapause initiation, maintenance and termination, although, in the present study, we highlight genes with roles in several key pathways and/or categories known to influence developmental progression. Genes integral to cell cycle progression are highly enriched in the 'all down-' regulated list and, at some level, these genes are likely to play a role in diapause developmental arrest (e.g. through cell cycle arrest) (Tammariello & Denlinger, 1998). For example, normal progression through the cell cycle may be halted at several checkpoints, where stimuli such as DNA damage may prevent the transition from one cell cycle phase to another (Murray, 1994). Many genes with vertebrate homologues identified as tumor suppressors play a key role in 'enforcing' these check points (Edgar & Ofarrell, 1989; Kastan & Bartek, 2004). However, the specific controls of cell cycle progression often depend on the developmental context (Ninov *et al.*, 2009) and thus are difficult to predict for diapause in general. Moreover, many of these genes (e.g. cyclins and cyclin-dependent kinases) are likely downstream targets of regulatory pathways. In the present study, we focus on the genes that are likely to play regulatory roles in pathways previously implicated in diapause regulation, including insulin/Tor signalling, ecdysone signalling and synthesis, and circadian rhythms, as summarized in Table 2. We emphasize that this is a very limited list necessarily excluding many pathway members that are not present in our final, filtered data set.

Insulin signalling appears to be integral to diapause responses in several insect species and in the worm *C. elegans*. Both the fly and worm pathways are very similar to those characterized in vertebrates, with insulin-like peptides binding to the insulin receptor (*Inr* in flies, *daf-2* in worms), which initiates a signaling cascade that regulates aspects of metabolism (Anton, 2014) and interacts with several other signaling pathways. Principally, insulin signalling interacts with the Tor pathway to regulate growth and development (Wullschleger *et al.*, 2006). Repression of *Inr* via dsRNA interference produces a diapause-like phenotype in nondiapausing *Culex pipiens* mosquitoes (Sim & Denlinger, 2008) and naturally segregating sequence variation in the gene encoding phosphatidylinositol 3-kinase (PI3-kinase, central to the insulin signalling cascade) correlates with the frequency of diapause phenotypes in natural populations of *D. melanogaster* (Williams *et al.*, 2006). In *C. elegans* worms, *daf-2* function is necessary for entrance into the dauer stage, the analogue of insect diapause (Gottlieb & Ruvkun, 1994).

Both the 'all up-' and 'all down-' regulated gene sets in the present study contain members of the insulin/Tor signalling pathways. The 'all up-' regulated list contains the genes *melt* and *Nlaz*, and both modulate insulin signalling through interactions with the Tor and JNK pathways, respectively. The *melt* gene is necessary for proper lipid storage (Teleman *et al.*, 2005), whereas *Nlaz* mediates JNK antagonism of insulin signalling, which represses growth (Hull-Thompson *et al.*, 2009). The 'all up-' regulated list also contains *sga* and *rhea*, which are predicted to interact with insulin signalling based on a genetic interaction with a protein central to Tor signalling (Tsc1) and a computationally-predicted insulin receptor binding domain, respectively. The 'all down-' regulated list contains *Lnk*, *ns1* and *Nprl3*. The *Lnk* and *ns1* genes are both positive regulators of insulin signalling (Kaplan *et al.*, 2008; Werz *et al.*, 2009), whereas *Nprl3* is a negative regulator of Tor signalling (Wei & Lilly, 2014).

Wnt signalling is a highly conserved developmental regulatory pathway (Wodarz & Nusse, 1998) that plays a highly pleiotropic role in tissue proliferation in insect embryogenesis and metamorphosis (Siegfried & Perrimon, 1994). The gene *wg* is a secreted ligand that initiates a number of signalling cascades, principally the canonical (β-catenin-dependent) pathway, which interacts with a number of different pathways including insulin signalling (Yoon *et al.*, 2010). Members of the canonical and calcium-dependent Wnt signalling pathway are differentially expressed at the transcript level during diapause termination in the European corn borer (*Ostrinia nubilalis*) and the apple maggot fly (*R. pomonella*), both represented in the current data set (Ragland *et al.*, 2011; Wadsworth & Dopman, 2015). Furthermore, protein levels and phosphorylation states of genes in the Wnt pathway are associated with diapause in additional moth species (Lin *et al.*, 2009; Chen & Xu, 2014).

The 'all down-' regulated list included six Wnt pathway genes down-regulated during diapause. This general signal of down-regulation during a developmentally suppressed state matches with predictions based on the positive influence of Wnt signalling on developmental progression (Wodarz & Nusse, 1998). However, one central member of the Wnt pathway, *spen*, is mainly up-regulated during diapause. This directionality is not predicted based on the typical function of *spen* as a positive regulator of Wnt signalling (Chang *et al.*, 2008). However, *spen* acts pleiotropically to repress Notch signalling, which is another central pathway regulating tissue proliferation and developmental patterning (Chang *et al.*, 2008). Moreover, Meyers *et al.* (2016) report that many Wnt signalling genes are differentially expressed during pupal diapause between populations of *R. pomonella* flies differing in their rates of diapause development. In particular, *spen* is the most differentially regulated gene (highest log-fold change) in the Wnt pathway and is up-regulated in the population that exhibits a more intense and longer pupal diapause (Meyers *et al.*, 2016).

Pulses of the hormone ecdysone are associated with molting and developmental transitions in all insects, and concomitantly play a role in transitions into and out of diapause (Truman & Riddiford, 2002; Denlinger *et al.*, 2005). For example, pupal diapause in the flesh fly *S. crassipalpis* is characterized by low ecdysone titres and is terminated by increased ecdysone titres (Denlinger, 2002). Upstream of ecdysone, the expression of several of the the so-called 'halloween' ecdysone biosynthesis genes is also shown to be down-regulated during diapause in the Drosophilid *C. costata* and the Tephritid *B. minax*, and both are included in the current data set (Dong *et al.*, 2014; Poupardin *et al.*, 2015). The 'all down-' regulated gene set includes one ecdysone biosynthesis gene, *phm*, one of the two transcripts coding for the heterodimeric ecdysone receptor complex, *Ecr*, and the nuclear receptor *Hr78*, which is essential for ecdysone signalling during metamorphosis (Fisk & Thummel, 1998). Hormone titres and receptor expression are sometimes elevated and sometimes depressed during diapause, depending on the diapause stage (Denlinger, 2002). Thus, the common pattern of regulation observed in the present study may not hold true during diapause in all insect species.

Finally, genes in the two primary feedback loops determining circadian rhythmicity are repeatedly linked to insect photoperiodic responses in general and diapause in particular (Goto, 2013; Meuti & Denlinger, 2013). The two loops involve cyclic expression of (i) *tim* and *per* and (ii) *vri* and *pdp1*; both loops interact with the core genes *Clk* and *cyc* (Panda *et al.*, 2002). The specific link between the time-counting mechanism in circadian rhythms versus the mechanism for measuring changes in day length remains unknown, although several potential models are proposed (Bradshaw & Holzapfel, 2007; Goto, 2013; Meuti & Denlinger, 2013). A number of studies report that expression levels of the canonical circadian rhythm genes are correlated with photoperiodism and diapause development. For example, mutant lines of *C. costata* carrying a 1855-bp deletion in the 5′-UTR of *tim* demonstrate drastically reduced *tim* gene expression and do not enter diapause under short-day conditions, the typical wild-type phenotype (Stehlík *et al.*, 2008). In addition, allelic variants of *tim* vary predictably with geography and associate with propensity to enter diapause in *D. melanogaster* (Sandrelli *et al.*, 2007; Tauber *et al.*, 2007). The 'all up-' regulated list in the present study contains *Pdp1*, another cycling, canonical circadian rhythm gene whose expression levels also vary with different day length conditions and diapause status in the linden bug *Pyrrhocoris apterus* (Dolezel *et al.*, 2008). Furthermore, *Pdp1* appears to play a central role as a negative modulator insulin signalling that affects dauer diapause in *C. elegans* (Narasimhan *et al.*, 2011).

Summary

The extensive diversity of diapause strategies in diapausing insect species interspersed with many nondiapausing species has led to the reasonable but largely untested hypothesis that diapausing life-cycle stages have evolved independently many times during insect evolutionary diversification. Despite this diversity of diapause stages and phenotypes and the presumed multiple evolutionary origins, there are notable physiological similarities in diapause responses across insect species. Our whole-transcriptome analysis of diapause across 11 insect species has not detected any phylogenetic signal, although such a signal might be detectable in a more standardized experiment or at shallower phylogenetic distances. Despite the lack of phylogenetic signal, there is a clear signal of transcriptional similarity in diapause responses across species for a core set of transcripts. From this core set, we are able to identify candidate genes participating in regulatory pathways such as insulin and Wnt signalling, which would appear to support common regulatory mechanisms. However, the available data do not allow us to test whether similarities in expression patterns are driven primarily by similarities in diapause initiation, maintenance, or termination across species. Studies that apply more standardized, time series analysis of diapause development in focused taxonomic comparisons could overcome this limitation.

Shared data

In addition to the Supporting information, all orthology mappings and the full, unfiltered gene expression matrix have been archived in the Dryad Digital Repository: doi:10.5061/dryad.rq3sp.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: DOI: 10.1111/phen.12193

Table S1. Metadata for transcriptomic data sets.

Table S2. Eigengenes identified using the R package WGCNA.

Table S3. Log-fold changes (diapause/nondiapause) for all genes regulated in the same direction across all data sets, allowing one mismatch in direction.

Table S4. Tables of functional categories enriched in the 'all down-' regulated gene list as identified using DAVID.

Script S1. R script applying an edgeR model to fit expression data from *Drosophila montana*.

Figure S1. Dendrogram and cluster assignment (different colors represent different clusters) of orthologous genes based on co-expression patterns across species, as determined using the R package WGCNA.

Figure S2. Distributions of log-fold changes within each species in the combined data set.

Acknowledgements

We are grateful to Shin Goto and Daniel Hahn for organizing the 2016 International Congress of Entomology symposium 'Photoperiodic Induction of Diapause and Seasonal Morphs', as well as to all symposium participants for useful and stimulating discussion. We wish to thank Erik Dopman, You-Jin Hao, David Kang, Maaria Kankare, Vladimír Košťál, Monica Poelchau, Rodolphe Poupardin, Cheolho Sim, George Yocum and Crista Wadsworth for sharing sequence and expression data, as well as Shin Goto, Daniel Hahn and two anonymous reviewers for their valuable comments on a previous version of the manuscript. This work was supported by NSF grants IOS-1700773 and DEB-1638951 to G. J. R. and a 2016 Kansas State University summer REU fellowship to E. K. The authors declare that they have no conflicts of interest.

References

Amsalem, E., Galbraith, D.A., Cnaani, J. *et al.* (2015) Conservation and modification of genetic and physiological toolkits underpinning diapause in bumble bee queens. *Molecular Ecology*, **24**, 5596–5615.

Anton, S. (2014) The coming of age of insulin-signaling in insects. *Frontiers in Physiology*, **5**, 216.

Ayala, F.J. (1997) Vagaries of the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 7776–7783.

Baker, D.A. & Russell, S. (2009) Gene expression during *Drosophila melanogaster* egg development before and after reproductive diapause. *Bmc Genomics*, **10**, 242.

Bradshaw, W.E. & Holzapfel, C.M. (2007) Evolution of animal photoperiodism. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 1–25.

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

Carey, D.B. (1994) Diapause and the host plant affiliations of lycaenid butterflies. *Oikos*, **69**, 259–266.

Chang, J.L., Lin, H.V., Blauwkamp, T.A. & Cadigan, K.M. (2008) Spenito and Split ends act redundantly to promote Wingless signaling. *Developmental Biology*, **314**, 100–111.

Chen, W. & 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.

Danilevsky, A., Goryshin, N. & Tyshchenko, V. (1970) Biological rhythms in terrestrial arthropods. *Annual Review of Entomology*, **15**, 201–244.

Danks, H.V. (1987) *Insect Dormancy: An Ecological Perspective. Biological Survey of Canada (Terrestrial Arthropods)*. Entomological Society of Canada, Canada.

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

Denlinger, D.L., Yocum, G.D. & Rinehart, J.P. (2005) Hormonal control of diapause. *Comprehensive Molecular Insect Science: Endocrinology* (ed. by L.I. Gilbert, K. Iatrou and S.S. Gill). Elsevier, U.K.

Ding, S., Li, X., Wang, N. *et al.* (2015) The phylogeny and evolutionary timescale of muscoidea (Diptera: Brachycera: Calyptratae) inferred from mitochondrial genomes. *PLoS ONE*, **10**, e0134170.

Dolezel, D., Zdechovanova, L., Sauman, I. & Hodkova, M. (2008) Endocrine-dependent expression of circadian clock genes in insects. *Cellular and Molecular Life Sciences*, **65**, 964–969.

Dong, Y., Desneux, N., Lei, C. & Niu, C. (2014) Transcriptome characterization analysis of *Bactrocera minax* and new insights into its pupal diapause development with gene expression analysis. *International Journal of Biological Sciences*, **10**, 1051–1063.

Edgar, B.A. & Ofarrell, P.H. (1989) Genetic control of cell division patterns in the *Drosophila* embryo. *Cell*, **57**, 177–187.

Fisk, G.J. & Thummel, C.S. (1998) The DHR78 nuclear receptor is required for ecdysteroid signaling during the onset of *Drosophila* metamorphosis. *Cell*, **93**, 543–555.

Fraley, C. & Raftery, A.E. (1999) MCLUST: software for model-based cluster analysis. *Journal of Classification*, **16**, 297–306.

Gabaldón, T. & Koonin, E.V. (2013) Functional and evolutionary implications of gene orthology. *Nature Reviews Genetics*, **14**, 360–366.

Gao, J.-J., Watabe, H.-A., Aotsuka, T. *et al.* (2007) Molecular phylogeny of the Drosophila obscura species group, with emphasis on the Old World species. *BMC Evolutionary Biology*, **7**, 1.

Gomi, T. & Takeda, M. (1996) Changes in life-history traits in the fall webworm within half a century of introduction to Japan. *Functional Ecology*, **10**, 384–389.

Goto, S.G. (2013) Roles of circadian clock genes in insect photoperiodism. *Entomological Science*, **16**, 1–16.

Goto, S.G., Yoshida, T., Beppu, K. & Kimura, M.T. (1999) Evolution of overwintering strategies in Eurasian species of the *Drosophila obscura* species group. *Biological Journal of the Linnean Society*, **68**, 429–441.

Gottlieb, S. & Ruvkun, G. (1994) Daf-2, daf-16 and daf-23 – genetically interacting genes controlling dauer formation in *Caenorhabditis elegans*. *Genetics*, **137**, 107–120.

Hao, Y.J., Zhang, Y.J., Si, F.L., Fu, D.Y., He, Z.B. & Chen, B. (2016) Insight into the possible mechanism of the summer diapause of *Delia antiqua* (Diptera: Anthomyiidae) through digital gene expression analysis. *Insect Science*, **23**, 438–451.

Hahn, D.A. & Denlinger, D.L. (2011) Energetics of insect diapause. *Annual Review of Entomology*, **56**, 103–121.

Heng, H., Stevens, J.B., Bremer, S.W. *et al.* (2010) Evolutionary mechanisms and diversity in cancer. *Advances in Cancer Research*, **112**, 217–253.

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

Hull-Thompson, J., Muffat, J., Sanchez, D. *et al.* (2009) Control of metabolic homeostasis by stress signaling is mediated by the lipocalin NLaz. *PLoS Genetics*, **5**, e1000460.

Iseli, C., Jongeneel, C.V. & Bucher, P. (1999) ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proceedings International Conference on Intelligent Systems for Molecular Biology*, 138–148.

Kang, D.S., Cotten, 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. *et al.* (2016) Transcriptional differences between diapausing and non-diapausing *D. montana* females reared under the same photoperiod and temperature. *PLoS ONE*, **11**, e0161852.

Kaplan, D.D., Zimmermann, G., Suyama, K. *et al.* (2008) A nucleostemin family GTPase, NS3, acts in serotonergic neurons to regulate insulin signaling and control body size. *Genes & Development*, **22**, 1877–1893.

Kastan, M.B. & Bartek, J. (2004) Cell-cycle checkpoints and cancer. *Nature*, **432**, 316–323.

Kingsolver, J.G. & Nagle, A. (2007) Evolutionary divergence in thermal sensitivity and diapause of field and laboratory populations of *Manduca sexta*. *Physiological and Biochemical Zoology*, **80**, 473–479.

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

Langfelder, P. & Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 1.

Lin, J.-L., Lin, P.-L. & 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.

Losos, J.B. (1992) The evolution of convergent structure in Caribbean Anolis communities. *Systematic Biology*, **41**, 403–420.

Lounibos, L., Escher, R. & Lourenço-de-Oliveira, R. (2003) Asymmetric evolution of photoperiodic diapause in temperate and tropical invasive populations of *Aedes albopictus* (Diptera: Culicidae). *Annals of the Entomological Society of America*, **96**, 512–518.

Meuti, M.E. & Denlinger, D.L. (2013) Evolutionary links between circadian clocks and photoperiodic diapause in insects. *Integrative and Comparative Biology*, **53**, 131–143.

Meyers, P.J., Powell, T.H., Walden, K.K. *et al.* (2016) Divergence of the diapause transcriptome in apple maggot flies: winter regulation and post-winter transcriptional repression. *Journal of Experimental Biology*, **219**, 2613–2622.

Misof, B., Liu, S., Meusemann, K. *et al.* (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, **346**, 763–767.

Murray, A. (1994) Cell cycle checkpoints. *Current Opinion in Cell Biology*, **6**, 872–876.

Nakagaki, M., Takei, R., Nagashima, E. & Yaginuma, T. (1991) Cell cycles in embryos of the silkworm, *Bombyx mori*: G2-arrest at diapause stage. *Roux's Archives of Developmental Biology*, **200**, 223–229.

Narasimhan, S.D., Yen, K., Bansal, A. *et al.* (2011) PDP-1 links the TGF-β and IIS pathways to regulate longevity, development, and metabolism. *PLoS Genetics*, **7**, e1001377.

Ninov, N., Manjon, C. & Martin-Blanco, E. (2009) Dynamic control of cell cycle and growth coupling by ecdysone, EGFR, and PI3K signaling in *Drosophila* histoblasts. *PLoS Biology*, **7**, 892–903.

Panda, S., Hogenesch, J.B. & Kay, S.A. (2002) Circadian rhythms from flies to human. *Nature*, **417**, 329–335.

Parker, D.J., Ritchie, M.G. & Kankare, M. (2016) Preparing for winter: the transcriptomic response associated with different day lengths in *Drosophila montana*. *G3-Genes Genomes Genetics*, **6**, 1373–1381.

Peterson, B.K., Hare, E.E., Iyer, V.N. *et al.* (2009) Big genomes facilitate the comparative identification of regulatory elements. *PLoS ONE*, **4**, e4688.

Poelchau, M.F., Reynolds, J.A., Elsik, C.G. *et al.* (2013) Deep sequencing reveals complex mechanisms of diapause preparation in the invasive mosquito, *Aedes albopictus*. *Proceedings of the Royal Society of London Series B, Biological Sciences*, **280**, 20130143.

Poupardin, R., Schottner, K., Korbelova, J. *et al.* (2015) Early transcriptional events linked to induction of diapause revealed by RNAseq in larvae of drosophilid fly, *Chymomyza costata*. *BMC Genomics*, **16**, 720.

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

Ragland, G.J., Egan, S.P., Feder, J.L. *et al.* (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.

Reidenbach, K.R., Cook, S., Bertone, M.A. *et al.* (2009) Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae) based on nuclear genes and morphology. *BMC Evolutionary Biology*, **9**, 1.

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

Sandrelli, F., Tauber, E., Pegoraro, M. *et al.* (2007) A molecular basis for natural selection at the timeless locus in *Drosophila melanogaster*. *Science*, **316**, 1898–1900.

Schluter, D. (1996) Adaptive radiation along genetic lines of least resistance. *Evolution*, **50**, 1766–1774.

Siegfried, E. & Perrimon, N. (1994) *Drosophila* wingless: a paradigm for the function and mechanism of Wnt signaling. *Bioessays*, **16**, 395–404.

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

Sim, C. & Denlinger, D.L. (2013) Insulin signaling and the regulation of insect diapause. *Frontiers in Physiology*, **4**, 189.

Sims, S.R. (1982) Larval diapause in the eastern tree-hole mosquito, *Aedes triseriatus*: latitudinal variation in induction and intensity. *Annals of the Entomological Society of America*, **75**, 195–200.

Stehlík, J., Závodská, R., Shimada, K. *et al.* (2008) Photoperiodic induction of diapause requires regulated transcription of timeless in the larval brain of *Chymomyza costata*. *Journal of Biological Rhythms*, **23**, 129–139.

Tammariello, S.P. & 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.

Tatusov, R.L., Koonin, E.V. & Lipman, D.J. (1997) A genomic perspective on protein families. *Science*, **278**, 631–637.

Tauber, M.J., Tauber, C.A. & Masaki, S. (1986) *Seasonal Adaptations of Insects*. Oxford University Press, New York, New York.

Tauber, E., Zordan, M., Sandrelli, F. *et al.* (2007) Natural selection favors a newly derived timeless allele in *Drosophila melanogaster*. *Science*, **316**, 1895–1898.

Teleman, A.A., Chen, Y.W. & Cohen, S.M. (2005) *Drosophila* melted modulates FOXO and TOR activity. *Developmental Cell*, **9**, 271–281.

Truman, J.W. & Riddiford, L.M. (2002) Endocrine insights into the evolution of metamorphosis in insects. *Annual Review of Entomology*, **47**, 467–500.

Vieira, C.P., Vieira, J. & Hartl, D.L. (1997) The evolution of small gene clusters: evidence for an independent origin of the maltase gene cluster in *Drosophila virilis* and *Drosophila melanogaster*. *Molecular Biology and Evolution*, **14**, 985–993.

Wadsworth, C.B. & Dopman, E.B. (2015) Transcriptome profiling reveals mechanisms for the evolution of insect seasonality. *Journal of Experimental Biology*, **218**, 3611–3622.

Wall, D.P. & Deluca, T. (2007) Ortholog detection using the reciprocal smallest distance algorithm. *Methods in Molecular Biology*, **396**, 95–110.

Waterhouse, R.M., Tegenfeldt, F., Li, J. *et al.* (2013) OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Research*, **41**, D358–D365.

Wei, Y. & Lilly, M.A. (2014) The TORC1 inhibitors Nprl2 and Nprl3 mediate an adaptive response to amino-acid starvation in *Drosophila*. *Cell Death and Differentiation*, **21**, 1460–1468.

Werz, C., Kohler, K., Hafen, E. & Stocker, H. (2009) The *Drosophila* SH2B family adaptor lnk acts in parallel to chico in the insulin signaling pathway. *PLoS Genetics*, **5**, e1000596.

Williams, K.D., Busto, M., Suster, M.L. *et al.* (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. & Nusse, R. (1998) Mechanisms of Wnt signaling in development. *Annual Review of Cell and Developmental Biology*, **14**, 59–88.

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

Yocum, G.D., Rinehart, J.P., Horvath, D.P. *et al.* (2015) Key molecular processes of the diapause to post-diapause quiescence transition in the alfalfa leafcutting bee *Megachile rotundata* identified by comparative transcriptome analysis. *Physiological Entomology*, **40**, 103–112.

Yoon, J.C., Ng, A., Kim, B.H. *et al.* (2010) Wnt signaling regulates mitochondrial physiology and insulin sensitivity. *Genes & Development*, **24**, 1507–1518.

Zhang, T.-Y., Sun, J.-S., Zhang, Q.-R. *et al.* (2004) The diapause hormone-pheromone biosynthesis activating neuropeptide gene of *Helicoverpa armigera* encodes multiple peptides that break, rather than induce, diapause. *Journal of Insect Physiology*, **50**, 547–554.

Zhao, X., Bergland, A.O., Behrman, E.L. *et al.* (2015) Global transcriptional profiling of diapause and climatic adaptation in *Drosophila melanogaster*. *Molecular Biology and Evolution*, **33**, 707–720.

Accepted 20 February 2017
First published online 17 April 2017