Submission as an article (Discoveries section)

FUNCTIONAL DIVERGENCE OF THE MIRNA TRANSCRIPTOME AT THE ONSET OF *DROSOPHILA* METAMORPHOSIS

Shu-Dan Yeh [1,5], Marcin von Grotthuss [1,4], Kania A. Gandasetiawan [1], Suvini Jayasekera [1], Xiao-Qin Xia [2], Carolus Chan [1], Vivek Jayaswal [3], José M. Ranz [1,5]

[1] Department of Ecology and Evolutionary Biology, University of California Irvine, CA 92697

[2] Institute of Hydrobiology, Chinese Academy of Sciences, 7 Donghu South Road, Wuhan 430072, China

[3] School of Mathematics and Statistics, The University of Sydney, Sydney, NSW, Australia 2006

[4] Present address: The Broad Institute, 7 Cambridge Center, Cambridge, MA 02142

[5] Corresponding author: jranz@uci.edu

Keywords: *Drosophila*, evolution of expression profiles, miRNAs, metamorphosis, miRNA-mRNA associations

Running head: Evolution of miRNA expression levels

ABSTRACT

MicroRNAs (miRNAs) are endogenous RNA molecules that regulate gene expression post-transcriptionally. To date, the emergence of miRNAs and their patterns of sequence evolution have been analyzed in great detail. However, the extent to which miRNA expression levels have evolved over time, the role different evolutionary forces play in shaping these changes, and whether this variation in miRNA expression can reveal the interplay between miRNAs and mRNAs remain poorly understood. This is especially true for miRNA expressed during key developmental transitions. Here, we assayed miRNA expression levels immediately before (≥18BPF) and after (PF) the increase in the hormone ecdysone responsible for triggering metamorphosis. We did so in four strains of *Drosophila melanogaster* and two closely related species. In contrast to their sequence conservation, ~25% of miRNAs analyzed showed significant within-species variation in male expression levels at ≥18BPF and/or PF. Additionally, ~33% showed modifications in their pattern of expression bias between developmental timepoints. A separate analysis of the ≥18BPF and PF stages revealed that changes in miRNA abundance accumulate linearly over evolutionary time at PF but not at ≥18BPF. Importantly, ≥18BPF-enriched miRNAs showed the greatest variation in expression levels both within and between species, so are the less likely to evolve under stabilizing selection. Functional attributes, such as expression ubiquity, appeared more tightly associated with lower levels of miRNA expression polymorphism at PF than at ≥18BPF. Furthermore, ≥18BPF- and PF-enriched miRNAs showed opposite patterns of covariation in expression with mRNAs, which denoted the type of regulatory relationship between miRNAs and mRNAs. Collectively, our results show contrasting

patterns of functional divergence associated with miRNA expression levels during *Drosophila* ontogeny.

3

INTRODUCTION

Precise regulation of gene expression is instrumental for proper execution of the majority of biological processes including cell differentiation and homeostasis. A key parameter of this regulation is mRNA abundance, which is influenced post-transcriptionally by microRNAs (miRNAs) . MiRNAs are small (~22 nt) non-coding RNA trans-acting factors that induce mRNA decay or translation inhibition by base pairing with complementary regions on the mRNA molecule . This complementarity involves Watson-Crick pairing with particular motifs at the 3' untranslated region (UTR) of the mRNA, although it can also occur at the 5’ UTR or coding region . Importantly, miRNAs found in distantly related taxa tend to exhibit a remarkable degree of sequence conservation, especially in the so-called “seed” motif, which is close to the 5' end of the miRNA and is critical for the interaction with the targeted mRNAs .

MiRNAs are thought to ameliorate expression noise in expression networks and, consequently, increase the robustness of developmental systems contributing to phenotypic stability . It is therefore of special interest to determine the limits to which changes in miRNA expression attributes such as the expression level can be accommodated over evolutionary time. Changes in miRNA abundance can contribute to inter-individual variation in expression of miRNA-regulated targets ultimately impacting on protein levels and distribution . Some changes in miRNA abundance have been shown to account for variability in platelet reactivity and drug sensitivity in humans , elicit common disorders such as cancer , or underlie the intraspecific variation of morphological characters . Among closely related species, divergence in miRNA abundance has also been linked to important functional and phenotypic consequences. In the natural occurring hybrid of two sister *Arabidopsis*

species, miRNA expression diversity from the parental species results in novel phenotypes that contribute to adaptation . In primates, it has been proposed that miRNA-mediated differences in mRNA abundance underlie partly the evolution of human cognitive functions . To date, the malleability in miRNA expression levels during key developmental transitions requiring precise regulation of gene expression remains largely unexplored both at the intra- and interspecific levels.

Metamorphosis is an intricate biological process in which large-scale tissue remodeling and organogenesis are orchestrated . In *Drosophila,* fluctuation of 20-hydroxyEcdysone (20E) level induces the transition from larva to immobile pupa . This transition is accompanied by multiple changes in mRNA abundance , with some occurring in a sex-dependent fashion . Evolutionary changes in mRNA abundance have been reported both within and between closely related species of the *Drosophila melanogaster* species subgroup at the onset of metamorphosis . MiRNAs play critical roles during insect metamorphosis and in fact their expression profiles have also been characterized at the onset of *D. melanogaster* metamorphosis . However, neither the extent to which miRNAs can accommodate intra- and inter-specific changes in abundance nor the interplay of these evolved changes with fluctuation in target mRNA abundance have been elucidated. This is especially relevant when considering the contrast between the two stages that define this transition: late third instar larva stage, mostly characterized by the transition into a post-feeding stage, wandering, and finding a place to glue; and puparium formation, mostly characterized by an effective deployment of part of the developmental blueprint that leads to the formation of an adult individual. Here, we address these issues by examining miRNA expression profiles and their evolutionary patterns within the *D. melanogaster* species group. We generate a portrait of how miRNA

expression levels have been shaped by different evolutionary mechanisms at larva and pupa stages and how they have been impacted by different factors during this key organismal transition.

RESULTS AND DISCUSSION

MiRNA expression profiles at the onset of metamorphosis

Expressed miRNAs at the onset of metamorphosis. We surveyed miRNA expression at late third instar larvae (18 hours before puparium formation; ≥18BPF) and white prepuparium (at puparium formation; PF) in both sexes of *D. melanogaster* Oregon-R and in males of its close relatives *D. simulans* and *D. yakuba* using Illumina RNA-seq (Material and Methods; supplementary table S1 and dataset S1, Supplementary Material online). These two species shared ancestor with *D. melanogaster* ~5.4 and ~12.8 million years ago . We found reads corresponding to sequences of ~76% (130 out of 171) of the miRNAs registered in miRBase release 15 for *D. melanogaster* , which is comparable to other surveys (supplementary figure S1A, Supplementary Material online). Of the miRNAs with evidence of expression, ~73% (94 out of 130) were supported by sequence reads both in our Illumina dataset and in previously generated sequence reads by 454 and J.M. Ranz and M. Ashburner, unpublished results; supplementary table S2 and alignments, Supplementary Material online). Interestingly, ~38% (41 out of 107) and ~29% (35 out of 120) of the miRNA genes with sequence reads in the ≥18BPF and PF Illumina datasets, respectively, do so in one sex only, suggesting some potential sex-biased expression (supplementary figure S1B, Supplementary Material online).

In the case of *D. simulans* and *D. yakuba*, we documented the expression of 81 and 95 orthologous miRNAs, respectively (supplementary dataset S1, Supplementary Material online). Among these orthologs, 9 in *D. simulans* and 33 in *D. yakuba* were either not listed in miRBase release 15 or listed with an associated hairpin sequence for which we found no support due to nucleotide differences relative to the reference genomes sequence of these species (*e.g. miR-277-3p* and *miR-305-5p* in *D. simulans;* supplementary figure S2, Supplementary Material online).

As a preamble to our analyses on polymorphism and divergence (see below), we assayed levels of miRNA expression at ≥18BPF and PF in males using a microarray platform that included: i) reporters for miRNAs registered in miRBase release 15 across insect species; ii) 100 additional reporters based on previous deep-sequencing results from several *Drosophila* species; and iii) several controls (Material and Methods; supplementary figure S3 and table S3, Supplementary Material online). We examined four strains of *D. melanogaster,* including one African strain to better reflect the recent demographic history of the species , and one strain of each of its relatives *D. simulans* and *D. yakuba* (Material and Methods; supplementary table S1, Supplementary Material online). A total of 280 reporters representing 132 and 148 sequences of the 5' and 3' arms, respectively, of different miRNA genes relevant to the six strains of the *D. melanogaster* species subgroup were subject to downstream analyses. The potential impact of nucleotide mismatches on expression estimates across *D. melanogaster* strains was found to be limited to two reporters (supplementary text, table S4, and figure S4, Supplementary Material online). Within the subset of reporters considered, 120 were identical in sequence for the three species while the rest provided reliable information for two or one of the species only. Per species, we found similar counts

of expressed miRNA reporters: 77 in *D. melanogaster;* 73 in *D. simulans;* and 70 in *D. yakuba.* Among the 67 miRNA reporters that are conserved in sequence across the three species and could be assayed in our arrays, 55 (83%) were confirmed to be expressed at either ≥18BPF, PF, or both in all four *D. melanogaster* strains (supplementary dataset S1, Supplementary Material online). The reliability of our expression measures using an array platform was supported by the high positive correlations between biological replicates (supplementary figure S3, Supplementary Material online), by replicating some experiments with qRT-PCR (supplementary text, figures S5-S6, and table S5, Supplementary Material online), and by the good agreement between the expression estimates obtained between microarrays and deep-sequencing experiments (Spearman's ρ= 0.6171, $P<0.0001$; supplementary figure S7, Supplementary Material online).

MiRNA expression variability during male development. Microarray profiling across males from six *Drosophila* strains revealed the miRNA reporters differentially expressed between ≥18BPF and PF (*i.e.* ≥18BPF- and PF-enriched; figure 1A). Concordant with early single-strain studies , known ecdysone-induced miRNAs *let-7-5p*, *miR-125-5p*, and *miR-100-5p* , were PF-enriched across *D. melanogaster* strains, as they were in *D. simulans* and *D. yakuba.* Conversely, *miR-34-5p*, which is down-regulated by the transcription factor Broad in high ecdysone titer conditions, was found to be ≥18BPF-enriched across strains, a pattern also displayed by *miR-8-5p* . Nevertheless, the precise fraction of developmentally regulated miRNA reporters varied from strain to strain ranging from 32% in *D. yakuba* to 62% in the *D. melanogaster* strain Zimbabwe-109, being the median 42% (table 1; supplementary figure S8, Supplementary Material online).

Among *D. melanogaster* males, we confirmed several types of variation associated with miRNA expression levels. ~25% (15 out of 60) of the expressed miRNAs suitable for comparison across the four *D. melanogaster* strains showed significant differences in expression levels at ≥18BPF (6), at PF (4), or at both (5) (P_{adj}<0.01; Material and Methods). Further inspection revealed that only ~66% (42 out of 64) of the miRNA reporters with detectable level of expression in at least three strains showed consistency in the type of developmental expression pattern, *e.g.* PF enrichment across all the strains (supplementary table S6, Supplementary Material online). The remaining 33% of miRNAs (22 out of 64) harbored differences in their developmental expression pattern among strains. This is the case of *miR-956-3p*, which is ≥18BPF-enriched in all strains of *D. melanogaster* but in Zimbabwe-109 (figure 1B). Only the reporter corresponding to *miR-289-5p* showed opposite developmental regulation, *i.e.* ≥18BPF or PF enrichment depending on the strain. No difference in the ratio of developmentally to non-developmentally regulated miRNAs (21:21) was observed among reporters categorized as consistent.

Among the strains assayed, Zimbabwe-109 showed the largest proportion of PF-enriched miRNAs among those expressed (supplementary figure S8, Supplementary Material online), a difference that is statistically significant (randomization test of goodness-of-fit, $P_{adj}=2.9\times10^{-2}$ when Zimbabwe-109 is included and $P_{adj}=3.1\times10^{-1}$ when excluded; P_{adj}<0.05 when any other strain is omitted). To determine whether the lineage leading to Zimbabwe-109 or that leading to the other three strains of *D. melanogaster* accumulated the most alterations in developmental expression pattern, we included the strains of *D. simulans* and *D. yakuba* to phylogenetically polarize the differences recorded (supplementary figure S9, Supplementary Material online). Assuming maximum parsimony, the branch associated with Zimbabwe-109

appears to have accumulated more differences in developmental expression pattern than that leading to the other three *D. melanogaster* strains.

Consistency in the developmental expression pattern across strains does not preclude more subtle changes in miRNA abundance. For example, *miR-1012-5p* was categorized as non-developmentally enriched across all strains, however it entailed statistically significant differences at PF among some of them (ANOVA, P_{adj}<0.01; figure 1B). To evaluate the extent of these more subtle changes in expression, we focused on those miRNAs showing consistent patterns of expression not only within *D. melanogaster* but also across *D. simulans* and *D. yakuba*. Thirty such miRNA reporters were documented (8 ≥18BPF-enriched; 5 PF-enriched; 18 non-developmentally enriched) of which 17% (2 ≥18BPF-enriched, 2 PF-enriched, and 1 non-developmentally enriched) showed significant differences at ≥18BPF, PF, or both (P_{adj}<0.01; supplementary dataset S2, Supplementary Material online). Taken together, all these forms of variation pointed towards a malleable miRNA transcriptome at the onset of metamorphosis.

Gender differences in miRNA expression. We assayed miRNA expression levels in females from two strains of *D. melanogaster* and one of *D. simulans* finding a similar fraction of developmentally regulated miRNAs to that in males (table 1; supplementary figure S10, Supplementary Material online). 78% (54 out of 69) of miRNAs with detectable level of expression in females of the two *D. melanogaster* strains showed a consistent expression pattern between developmental stages with 24 of them (45%) exhibiting developmental regulation (11 ≥18BPF-enriched and 13 PF-enriched). Unlike in males, the proportion of miRNAs displaying different developmental expression patterns did not differ across strains (supplementary

figure S8, Supplementary Material online; randomization test of goodness-of-fit, $P_{adj}=1.5\times10^{-1}$; males -including the same three strains as in females-, $P_{adj}=6.0\times10^{-3}$).

Sex-biased gene expression for the protein-coding fraction of the genome has been examined at PF , but no equivalent analysis has been performed for miRNAs. Although limited, we did find evidence of miRNA sex-biased expression occurring as early as ≥18BPF (supplementary figure S11A, Supplementary Material online). Most miRNAs though, whether developmentally regulated or not, showed no evidence of sex-bias in expression (supplementary figure S12A-B, Supplementary Material online). At $P_{adj}<0.01$, 10 miRNA reporters showed significant sex-biased expression in at least one of the six strains by sex combinations assayed, a number that increased up to 22 at $P_{adj}<0.05$ (supplementary figure S11B and dataset S2, Supplementary Material online). Among miRNA reporters showing sex-bias in expression, we found cases such as *miR-964-5p*, which displayed differences in expression between the genders at a single developmental stage, and others like *miR-312-3p*, which did at both ≥18BPF and PF (supplementary figure S12C-D, Supplementary Material online). No case involving reversal in the pattern of developmental enrichment between the sexes, *e.g.* from ≥18BPF enrichment in males to PF enrichment in females, was detected.

Among the miRNA reporters showing sex-biased gene expression, the cluster *miR-310* - *miR-313* stood out. Each of the four constituent miRNA genes showed statistically significant male-biased expression in at least two of the six strains by developmental stage combinations assayed (supplementary figure S12D, Supplementary Material online). Although the sequence similarity of some of the miRNA genes in the cluster (*miR-310-3p, miR-311-3p,* and *miR-312-3p*) could give rise to some apparent co-expression, inspection of the deep-sequencing data for the

247strain Oregon-R (supplementary alignments, Supplementary Material online) and
248the inferred co-expression of the mir-310 cluster based on deep-sequencing data
249ruled out this possibility. Therefore, the observed sex-bias expression pattern
250strongly suggests a sex-dependent co-regulation of the mir-310 cluster at PF. This
251interpretation agrees well with the influence of this cluster on male fertility by
252modulating the Wingless signaling pathway, which is required for cell differentiation
253of the somatic and germline tissues in testis . Intriguingly, the expression of *miR-*
254*313-5p* could represent a case of sex-dependent arm-switching.

255

256**Evolution of miRNA expression profiles**

257***Differentiation of expression levels at ≥18BPF and PF.*** The multiple
258developmental changes occurring during early pupation compared to late instar
259larvae parallel changes in mRNA abundance of many genes during this
260developmental transition . Expression profiles of regulatory genes at PF should be
261especially refractory to change due to potential detrimental effects. We tested
262whether this hypothesis was reflected in miRNA expression levels by examining the
263way significant differences have accumulated over evolutionary time and by
264comparing the magnitude of expression differences between stages.

265We calculated expression distances among the males of the six strains surveyed at
266≥18BPF and PF separately as well as the divergence time for each strain pair
267(Material and Methods; supplementary text and table S7, Supplementary Material
268online). At PF, unlike at ≥18BPF, expression divergence in miRNA abundance did
269correlate with divergence time (figure 2A and supplementary figure S13,

Supplementary Material online). This result suggests that the way changes in miRNA abundance accumulate over evolutionary time differs between the stages.

We subsequently estimated the magnitude of miRNA expression differences across the six strains as the coefficient of variation (CV), *i.e.* the ratio of the standard deviation to the mean. We compared the log10-transformed CV at ≥18BPF and PF finding no significant difference (average CV: ≥18BPF, 0.761; PF, 0.729; one-way ANOVA, P=0.630). Nevertheless, absence of significant differences in the global levels of variation does not rule out more subtle patterns of differentiation among particular groups of miRNAs based on their expression attributes. According to this, we tested for differences in expression levels among miRNAs showing consistent developmental expression patterns in *D. melanogaster*. Both at ≥18BPF (one-way ANOVA, P=0.012) and PF (one-way ANOVA, P=0.005), we found the same rank of differentiation in miRNA expression levels: ≥18BPF-enriched > PF-enriched > non-developmentally enriched miRNAs (figure 2B). *Post hoc* tests underscored the difference between ≥18BPF- and non-developmentally enriched miRNAs (Tukey-Kramer HSD; ≥18BPF, P=0.019; PF, P=0.003). Collectively, these results suggest that the two developmental stages sampled might be subject to different evolutionary dynamics, which affect the pace at which expression changes accumulate over evolutionary time and the global levels of differentiation of particular groups of miRNAs.

Evolutionary mode and developmental expression pattern. We determined the mode of evolution (*i.e.*, stabilizing selection, genetic drift, or directional selection) of miRNA expression levels at ≥18BPF and PF. We used a two-step ANOVA-based approach in which differences in expression levels were tested first within and then between species (supplementary text, Supplementary Material

online). ~75% (39 out of 48) and ~80% (47 out of 55) miRNAs showed a mode of evolution consistent with the action of stabilizing selection at ≥18BPF and PF respectively (figure 3A), a result reminiscent of previous observations for mRNA levels . Our results suggest that directional selection has acted on the expression level of a few miRNAs: *miR-34-5p* and *miR-956-3p* at ≥18BPF, and *miR-34-5p*, *miR-312-3p,* and *miR-995-3p* at PF. The remaining miRNAs exhibited patterns of variation within and between species compatible with genetic drift alone or with other evolutionary scenarios such as relaxation of constraints in a lineage-dependent manner or combinations of genetic drift and some form lineage-dependent selection (figure 3B).

If miRNAs with different developmental expression patterns differ in their degree of variation in expression levels (fig. 2B), they should also show a different propensity to evolve under stabilizing selection. We confirmed this non-random association especially at PF (supplementary figure S14, Supplementary Material online). At this developmental stage, non-developmentally enriched and PF-enriched miRNAs largely evolve under stabilizing selection (~83% in both cases) while only 50% of the ≥18BPF-enriched miRNAs follow this mode of evolution. The difference in the proportion of miRNAs evolving under stabilizing selection is statistically significant between ≥18BPF- and non-developmentally enriched miRNAs (two-tailed Fisher's exact test, FET; ≥18BPF, P=0.040; PF, P=0.011). Why the expression level of ≥18BPF-enriched miRNAs is more variable overall than that of non-developmentally enriched miRNAs is unclear at this time. Nevertheless, this pattern reinforces the notion that miRNAs with different developmental expression patterns are exposed to different evolutionary pressures.

Evolutionary mode and miRNA evolutionary age. We examined whether the evolutionary mode in miRNA expression is related to evolutionary age, *i.e.* the moment at which we can parsimoniously date the emergence of a miRNA. Presumably, more ancient miRNAs should be more stably integrated into regulatory networks than younger miRNAs and thus should more likely evolve under stabilizing selection (Chen and Rajewsky 2007). We dated the emergence of miRNAs assuming a maximum parsimony framework (supplementary text and dataset S3, Supplementary Material online) distinguishing between two main age classes. The first class included miRNAs inferred to have emerged during the evolution of the subgenus *Sophophora* after the split with the subgenus *Drosophila* -and therefore less likely to evolve under stabilizing selection-, and the second class included miRNAs inferred to have been present in the common ancestor of both subgenera -and therefore more likely to be already stably integrated into the regulatory network-. Subsequently, we examined whether these two age classes differ in the extent to which they evolve under stabilizing selection versus other evolutionary modes. MiRNAs inferred to have emerged during the evolution of the subgenus *Sophophora* (6 expressed at ≥18BPF and 8 at PF) do evolve less often under stabilizing selection than more ancient miRNAs (46 expressed at ≥18BPF and 51 at PF), which is confirmed at PF but not at ≥18BPF (FET; ≥18BPF, P=0.157; PF, P=0.046). Although this result should be taken cautiously due to the limited count of the *Sophophora* specific miRNAs present in the analysis, it suggests that the expression levels of more recently evolved miRNAs have not had enough time to be shaped by natural selection to the extent the levels of expression of more ancient miRNAs have been.

15

Determinants of miRNA expression polymorphism in* D. melanogaster *males. The regulatory role of recently emerged miRNAs is more likely to be still evolving compared to that of miRNAs found in many phyla. Due to the potentially detrimental effects on organismal fitness of recently evolved miRNAs if expressed to a high level or across multiple tissues/organs, these miRNAs should exhibit lower expression levels and narrower spatiotemporal expression profiles than ancient miRNAs . Consistent with this notion, recently evolved miRNAs have been shown to be expressed at a low level in primates and *Drosophila* species . Further, the introduction of the relatively young *mir-310* family from *D. pseudoobscura* into the *D. melanogaster* genome resulted in misexpression of numerous genes and in lower organismal fitness . Based on these premises, miRNAs that are either present across most metazoans, exhibit a detrimental phenotype if misexpressed, or are more ubiquitously expressed across tissues and/or developmental stages should possess lower expression polymorphism.

We tested for a negative relationship between the intraspecific $\log_{10}$CV in expression of 62 miRNAs expressed at least in one of two studied developmental stages in *D. melanogaster* and evolutionary age, having associated a gain-of-function phenotype , and expression breadth across developmental stages and tissues . We did so for each developmental stage separately. For the evolutionary age, we distinguished broadly between *Drosophila* evolved miRNAs versus miRNAs inferred to have been present in the ancestor to the *Drosophila* genus and *Anopheles gambiae* (age classes *young* and *ancient* in figure 4A), finding evidence of a significant negative association with the level of expression polymorphism (one-way ANOVA; ≥18BPF, P=0.086; PF, P=0.008). When the number of age classes considered was further divided into four based on additional phylogenetic

partitions (supplementary text, Supplementary Material online), the trend was similar (one-way ANOVA; ≥18BPF, P=0.069; PF, P=0.034). In this case, pairwise *post hoc* tests revealed, at least for PF, that the diametrically opposed relationship between the expression polymorphism of miRNAs originated during the evolution of the subgenus *Sophophora* versus that of the most ancient miRNAs is the main factor contributing to the pattern found (supplementary figure S15A and table S8, Supplementary Material online). Considering age as a continuous variable did not alter this observation (supplementary figure S15B, Supplementary Material online).

Phenotypic effects upon inducing miRNA misexpression or being ubiquitously expressed were associated similarly with miRNA expression polymorphism in *D. melanogaster*. MiRNAs displaying gain-of-function phenotypes were found to harbor significantly lower CVs in expression than miRNAs with no phenotype especially at PF (one-way ANOVA; ≥18BPF, P=0.089; PF, P=0.017) (figure 4B). Furthermore, for the expression breadth, we found a significant negative correlation between expression ubiquity and $\log_{10}$CV, a pattern confined to the PF stage (r^2=0.228, P=0.004; ≥18BPF, r^2=0.021, P=0.426) (figure 4C). These results confirm that miRNAs that are more necessary for obtaining a wild type phenotype and miRNAs with more ubiquitous expression profiles, which are likely to be exposed to conflicting functional requirements across tissues , are more constrained in their capability to accommodate segregating expression variance during population differentiation. Importantly, these functional constraints are more apparent at PF than at ≥18BPF.

The negative correlations observed between miRNA expression polymorphism and evolutionary age, gain-of-function phenotypes, and expression ubiquity are unlikely to be independent from one another. This would be the case if more recently

originated miRNAs occupy less relevant positions in the regulatory network and have narrower spatiotemporal expression profiles, which can result in better accommodating higher levels of expression polymorphism. We did find statistical evidence of the association among these variables pointing to this parsimonious view (supplementary text, Supplementary Material online), which is consistent with a higher association of ancient miRNAs with disease phenotypes and with broader expression profiles across tissues, compared to recently originated miRNAs, reported in humans and *Diptera* respectively .

Using intraspecific variation in expression levels to uncover the landscape of the miRNA-mRNA regulatory network

To uncover the interface between miRNAs and mRNAs at the expression level at the onset of metamorphosis, we leveraged the intraspecific variation in expression levels for both molecules. We identified developmentally regulated miRNAs upon pooling the miRNA expression data from the males of the four strains of *D. melanogaster* (P_{adj}<0.05; supplementary text, Supplementary Material online). Then, we assayed levels of mRNA abundance from the same biological samples and used a mixed-effects linear model to estimate the expression association between developmentally regulated miRNAs and the transcripts for which the microarray platform had distinctive probesets, *i.e.* the so-called mRNA exemplars (Material and Methods supplementary; figure S3 and text, Supplementary Material online). An mRNA exemplar might correspond to one or more transcripts. Permutation tests helped determine whether the observed miRNA-mRNA association values were higher than expected by chance alone (P_{adj}<0.05; supplementary text, Supplementary Material online).

We identified 617 mRNA exemplars, representing 545 genes, as significantly associated with 43 developmentally regulated miRNAs (supplementary datasets S4-S5). The number of significant associations with mRNA exemplars varied remarkably among miRNAs (average ± SD / median; ~217 ± ~133 / 212; supplementary figure S16), with *miR-34-3p* displaying the highest number − 612. These statistically significant associations between expression levels of miRNAs and mRNAs may reflect concurrent co-regulation by a common upstream factor in the transcriptional hierarchy or a *bona fide* causal regulatory relationship, which can result from either direct or indirect targeting, the latter as it might occur between a miRNA regulating a transcription factor and the battery of genes under the transcription factor's control. We investigated the interplay between expression associations and their potential causal regulatory nature by considering the sign of the association, *i.e.* positive or negative, the link with miRNA developmental expression patterns, and the associated biological coherent patterns.

We distinguished between positive and negative associations in expression for each miRNA-mRNA exemplar pair. Importantly, when a miRNA covaried in expression with an mRNA exemplar, it was more likely to show a negative than a positive association (supplementary figure S16, Supplementary Material online). As reported in other organisms such as primates , positive expression associations were also abundant, which highlights the ambiguous nature of the interplay between expression levels of miRNAs and mRNAs . Intriguingly, developmentally regulated miRNAs showed a bimodal distribution for the proportion of negative significant associations over the total exhibited by each miRNA (figure 5A). We investigated whether this bimodal distribution was related to miRNA developmental expression patterns in a non-random manner. Specifically, we calculated the

correlation between the proportion of negative significant associations and the propensity of a miRNA in showing a particular type of developmental expression pattern across the four strains assayed (figure 5B). We found that the higher is the number of strains showing miRNA expression enrichment at PF, the higher is the proportion of negative miRNA-mRNA expression associations (Spearman's Rho=0.687, $P<0.0001$). Accordingly, ≥18BPF-enriched miRNAs across *D. melanogaster* strains were more likely to exhibit positive expression associations with mRNAs while PF-enriched miRNAs were more likely to exhibit negative associations.

These patterns could denote distinctive relationships with the covariating mRNAs (figure 5C). Down-regulation of a miRNA alone does not result directly in an increase of its targets' abundance unless that miRNA is involved in a feedback loop to inhibit the transcription of its targets. Therefore, positive miRNA-mRNA expression associations may often reflect the concurrent down-regulation of both molecules at PF denoting no causal regulatory relationship (figure 5C, left panel). Conversely, the up-regulation of a miRNA at PF may be important to facilitate the degradation of truly regulated targets that are not necessary at this stage, resulting in a decreasing abundance (figure 5C, right panel). This second pattern is reminiscent of the degradation of maternally deposited mRNAs by a set of zygotically expressed miRNAs in the *Drosophila* embryo . In consequence, negative associations of PF-enriched miRNAs and mRNAs should more likely represent *bona fide* causal regulations by miRNAs. To test this, we examined whether PF-enriched miRNAs showing negative expression associations with predicted targets among the 617 mRNA exemplars were significantly overrepresented relative to ≥18BPF-enriched miRNAs displaying equivalent properties. Indeed, we found that there are

more predicted targets among mRNAs negatively associated with PF-enriched miRNAs than among those with ≥18BPF-enriched miRNAs, a pattern not shown in positive expression associations (Randomization test of goodness-of-fit; P=0.013 and P=0.246 respectively; supplementary table S9, Supplementary Material online). This difference reinforces the possibility that positive expression associations are less likely than negative expression associations in denoting *bona fide* causal miRNA regulation at the onset of metamorphosis.

We further examined the biological properties of the miRNA-mRNA exemplar associations by searching for biological coherent patterns in genes grouped by their patterns of expression association with the 43 developmentally regulated miRNAs. Briefly, ten clusters of mRNA exemplars were identified by hierarchical cluster analysis based on the sign of their expression associations (supplementary figure S17, Supplementary Material online). Subsequently, functional enrichment for functional rubrics in each cluster was tested with DAVID under several degrees of stringency (Materials and Methods). We found enrichment for biological processes and other functional rubrics unambiguously related to the onset of metamorphosis in nine of the ten clusters (supplementary table S10, Supplementary Material online). Nucleotide biosynthesis and energy production pathways, structural and regulatory genes related to muscle formation, and genes involved in molting formation were enriched among those down-regulated at PF in clusters *1*, *3*, *4*, and *8-10*. On the other hand, histolysis upon tissue apoptosis and innate immune response pathway related genes were found to be overrepresented among those up-regulated at PF in clusters *5* and *7*.

A closer inspection of several clusters unveiled the complexity of the regulatory interactions that occur during the transition from larva to pupa. For instance, 66

mRNA exemplars in Cluster *8* were negatively associated with up to 27 miRNAs and positively associated with up to 14 miRNAs. This cluster is enriched for genes involved in the functional rubric "molting cycle" and the cellular component rubric "muscle myosin complex" (supplementary table S11, Supplementary Material online). Among the seven genes in Cluster *8* annotated as part of the functional rubrics relevant to muscle development ("muscle myosin complex" or "muscle protein"), five were predicted to harbor binding sites in their 3'UTRs according to TargetScan. Three of these genes showed negative associations in expression with miRNAs while the other two showed positive associations. The remaining two genes in Cluster *8* (*Tm2* and *Mlc2*) did not have any predicted miRNA binding site. Similar patterns were found for the constituent genes of Cluster *10*, which are annotated as part of the functional rubric "contractile fiber" (supplementary table S11, Supplementary Material online).

An intriguing aspect is that none of the genes that are a part of functional rubrics related to muscle development in Clusters *8* and *10* possess binding sites for *let-7* complex miRNAs in their 3'UTRs or ORFs. Nevertheless, the expression levels of these genes are negatively associated with those of the *let-7* complex miRNAs. The *let-7* complex is required for the maturation of neuromuscular junction and deformation of abdominal neuromusculature, which at least in part is achieved by downregulating the BTB-zinc finger transcription factor Ab . Several non-mutually exclusive explanations may account for this observation. First, miRNAs other than *let-7* may contribute to the regulation of these genes during metamorphosis (supplementary table S11, Supplementary Material online). Second, the *let-7* complex may regulate these genes indirectly through the control of their upstream

transcription factor(s) during metamorphosis. Third, the expression levels of these genes are not regulated post-transcriptionally by miRNAs denoting false positives.

The innate immune response of *Drosophila* at the onset of metamorphosis is well represented in the significant expression associations found. This response is regulated by ecdysone and juvenile hormone and mainly consists of two components: localized melanization and antimicrobial peptides production . We found that Cluster *7* is enriched for genes involved in these two components. The genes *Dat*, *e*, and *ple* were found to participate in the functional rubric "dopamine metabolic process", which contributes to melanization among other biological functions . Three other genes encode antimicrobial peptides: *Drs*, *Drsl2*, and *Drsl5*. The expression levels of these six genes are significantly higher at PF as expected and this work). Interestingly, positive expression associations were mostly found between genes and miRNAs that are predicted to bind their ORFs while the only two negative associations involve genes (*Drsl2* and *ple*) presumably bound at their 3’UTRs by miRNAs (supplementary table S12, Supplementary Material online). Further, the putative miRNAs regulating upstream genes of the innate immune pathway have been studied *in silico* . Some of these miRNAs were present in our miRNA-mRNA exemplars association list such as the *miR-2* family, *miR-9a* (both arms), *miR-125-5p*, *miR-279-3p*, and *miR-281-2-5p*. Thus, the expression levels of these innate immune response genes may be miRNA-regulated directly or through their upstream regulators in those same immune pathways.

Our characterization of the landscape of miRNA-mRNA exemplars associations is limited in two ways. First, it is dependent on the differential miRNA expression between the developmental stages compared. Second, some miRNAs are expressed in tissue- or cell-type-specific manners and therefore expression

associations may happen in multiple organs but in different directions resulting in a blurry association signal if any. It is remarkable therefore that we are still able to capture significant expression associations from our whole-body assays, which in some cases are suggestive of *bona fide* causal regulation.

We have generated a portrait of the intra- and inter-specific differences in expression levels of sequence conserved miRNAs at the onset of *Drosophila* metamorphosis. In spite of the documented phylogenetic differentiation in expression levels, we find that the evolution of miRNA abundance is driven mainly by stabilizing selection. This agrees well with the stabilizing role that miRNAs play by repressing leaky expression or fine-tuning transcript levels. Notably, the type of developmental expression pattern of a miRNA appears to be an excellent predictor of the degree to which a miRNA can accommodate variation in expression level during the evolutionary process. The expression level**s** of some miRNAs are still evolving, which in a few cases seems compatible with the optimization of their functional role by directional selection. Whether this functional optimization goes beyond canalizing expression levels during metamorphosis contributing as well to the phenotypic diversification in the genus *Drosophila* remains to be established. Overall, we find distinct patterns of differentiation among miRNA expression levels between late third instar larva and white prepupa, which is also evidenced in how these evolved changes are linked to relevant proxies for the integration of miRNAs into the regulatory network. In addition, miRNAs with different developmental expression patterns exhibit marked differences in how they covariate with mRNAs in expression, which might reflect the type of regulatory relationship between both kind of molecules. Taken together, the profound differences in the biological

requirements of these two developmental stages seem to dictate the properties in miRNA expression levels at the onset of *Drosophila* metamorphosis.

MATERIALS AND METHODS

Fly stocks and husbandry. Six strains representing three species of the *D. melanogaster* species group were used (supplementary table S1, Supplementary Material online). Flies were grown in standard corn meal medium, constant lighting conditions, and at 25ºC. Third instar larvae were identified as previously reported . Briefly, larvae were raised under non-crowded conditions in medium with 0.05% bromophenol blue. Wandering larvae exhibiting dark blue gut, which corresponds to ~18 hours before puparium formation, were collected, rinsed with water, and separated by sex. For the 0-1 hr white prepuparia, wandering larvae with light blue to white gut were separated by sex, placed in a Petri dish on damped light-wipe tissue until appropriate moment for collection . The presence of visible male gonads was used for sex identification; trial collections were performed to assure process accuracy. Samples were snap frozen in liquid nitrogen and stored at -80C until RNA isolation.

RNA extractions. Two rounds of RNA collection were performed: the first, for small RNA-seq; and the second, for the remaining expression profiling approaches used (see below). For each collection, ~120 mg (60-80 individuals depending on the strain) of frozen flies for each biological replicate were grinded using motorized pestles and the total RNAs extracted and purified with *miR*Vana miRNA isolation kit (Ambion Inc.), which allows the optional separation of mRNAs from small RNAs. In the first collection, the enriched small RNAs were isolated according to the

25

manufacturer indications for subsequent small RNA library preparation. In the second collection, both total RNAs and small RNAs from four biological replicates for each strain/sex were isolated and used in all three expression profiling approaches. Concentration, quality, and integrity of the RNA samples were assessed using the NanoDrop 8000 Spectrophotometer and the RNA 6000 Nano and Small RNA kits (Agilent Technologies) in an Agilent 2100 Bioanalyzer. RNA extractions for different strains were performed independently to avoid cross-contamination.

Small RNA-seq. Adaptor-ligated cDNAs were prepared according to Illumina small RNA preparation protocol (Preparing Samples for Analysis of Small RNA Using the Oligo Only Kit, http://www.illumina.com). Briefly, 16-28nt small RNAs were size selected from a denaturing polyacrylamide gels, 5' adaptors with four nucleotide indexes and 3' adaptors were ligated using T4 RNA ligase (Ambion 2140) and the resulting molecules were purified by size selection from denaturing polyacrylamide gels after each ligation reaction. Four samples with different indexes were pooled for 3' adaptor ligation and further preparation steps. The ligated small RNAs were reversed transcribed using Superscript II reverse transcriptase (Invitrogen 18064), subsequently PCR amplified with a Phusion high fidelity PCR DNA polymerase (Finnzymes F-506) through 10 cycles, and purified from denaturing polyacrylamide gels. Library preparation for different strains was performed independently to avoid cross-contamination. The quality of the cDNAs was evaluated with the High Sensitivity DNA kit (Agilent Technologies) using an Agilent 2100 Bioanalyzer before high-throughput sequencing using a Genome Analyzer II (Illumina) at the Genome Center of the University of California, Davis. Twelve libraries, named as L5-L16 in supplementary table S2 (Supplementary Material online), were sequenced in three

lanes. Reproducibility of sequencing results was evaluated at biological and technical levels.

Reads were sorted by distinctive indexes before their 5' and 3' adaptors sequences were trimmed. Reads matching yeast rRNAs and *D. melanogaster* 2S rRNA were discarded. The remaining reads were aligned against the stem-loop sequences of known miRNAs according to miRBase release 15 using Bowtie 0.12.5 . Because of the potential errors in the reference genome of *D. simulans* and *D. yakuba,* we also used the *D. melanogaster* genome sequence as a reference in the characterization of the libraries of these two species. Alignments and read numbers were recorded. Four libraries sequenced by 454, named as L1-L4 in supplementary table S2 (Supplementary Material online), were also included and J.M. Ranz and M. Ashburner, unpublished results). In-house Perl scripts were used for sequence processing.

miRNA microarray profiling. The miRNArthropoda_15_UC_100610 array from LC Sciences based on the annotations of miRBase release 15 was used. In addition, 100 custom probes were added to the array (supplementary table S3, Supplementary Material online). Custom probes include previously dubbed passenger sequences of known *D. melanogaster* miRNAs, putative miRNAs found in deep-sequencing experiments by us and others , and control reporters for the 2S rRNA gene harboring nucleotide differences in number and position to evaluate the impact of mismatches on hybridization kinetics. Probes for annotated and predicted miRNAs were present in triplicates. Internal quality controls included 43 spikes and six additional positive controls replicated either 4 or 16 times on the array.

Four biological replicates per developmental stage per strain were used in competitive hybridizations. Labeling, hybridization, and image acquisition were

performed by LC Sciences. For a given array, the Cy3 and Cy5 fluorescent intensity values of each array were first adjusted by subtracting local background and then normalized according to a locally-weighted regression approach . The adjusted values were further log2-transformed and normalized across arrays using the quantile method implemented in JMP Genomics 5.0 . The expression values for the 280 reporters relevant to the six strains of interest were obtained by averaging over the three technical replicates spotted on separate blocks of the array. A mean intensity value lower than 32 was not considered to be reflective of expression above the background. A given miRNA was not considered to be expressed if it lacked evidence of expression across all sex by strain by developmental stage combinations assayed.

A linear model was used to test for differences in expression levels between developmental stages in any given strain by sex combination and for differences in expression levels between the sexes in any given strain by developmental stage combination. For this purpose, a linear mixed-effects gene model that takes into account both array and dye-specific effects was considered –

$$y_{ijklm} = \mu + A_i + T_j + S_k + D_l + Z_m + S_kD_l + S_kZ_m + D_lZ_m + S_kD_lZ_m + \varepsilon_{ijklm}$$

where y_{ijklm} denotes the miRNA expression for the i^{th} array, j^{th} dye, k^{th} strain, l^{th} developmental stage and m^{th} sex. Also, μ is the baseline expression, A_i is the effect of the i^{th} array, T_j is the effect of the j^{th} dye, S_k is the effect of the k^{th} strain, D_l is the effect of the l^{th} developmental stage, and Z_m is the effect of the m^{th} sex. A was implemented as a random effect while T, S, D, Z, and the interaction effects were implemented as fixed effects in JMP Genomics 5.0. Next, the appropriate contrasts were made to obtain the differences of interest. The differences of interest were considered to be statistically significant at a false discovery rate (FDR) of 0.01.

28

For a given strain, to test the null hypothesis that the difference in magnitude of developmental change was the same in both sexes, a linear mixed-effects model was developed as follows –

$$y_{ijlm} = \mu + A_i + T_j + D_l + Z_m + D_lZ_m + \varepsilon_{ijlm}$$

where y_{ijlm} denotes the miRNA expression for the i^{th} array, j^{th} dye, l^{th} developmental stage and m^{th} sex. Also, μ is the baseline expression, A_i is the effect of the i^{th} array, T_j is the effect of the j^{th} dye, D_l is the effect of the l^{th} developmental stage, and Z_m is the effect of the m^{th} sex. A was implemented as a random effect while T, D, Z, and the interaction effect were implemented as fixed effects. Since there were two developmental stages and two sexes the interaction term corresponded to the difference of interest. The difference in magnitude of developmental change in expression across sexes was considered to be statistically significant at a FDR of 0.05.

qRT-PCR. Triplicate total RNA samples for each strain by developmental stage by sex analyzed were polyadenylated and reverse-transcribed using oligo-dT as a primer following manufacturer conditions (Exiqon Universal cDNA Synthesis Kit, 203300). The qRT-PCR step was performed in a CFX-96 real-time instrument (BioRad) using Locked Nucleic Acid (LNA) primers (Exiqon; supplementary table S5, Supplementary Material online) and SYBR Green chemistry (Exiqon SYBR Green Master Mix, 203450). Expression levels of the miRNAs analyzed were estimated relative to the reference genes *miR-1-3p* and *miR-995-3p,* which were chosen based on two criteria: i) expression uniformity in microarray experiments across all strain by developmental stage by sex combinations; and ii) because they cover two differentiated levels of expression (*miR-1-3p* is substantially more expressed than *miR-995-3p* according to the small-RNA sequencing experiments performed in this

study). Estimates were calculated using the $-2^{\Delta\Delta Cq}$ method implemented in the Bio-Rad CFX manager software and statistically significant differences among samples were interrogated in JMP Genomics 5.0 (SAS Institute Inc.).

mRNA microarray profiling. We assayed levels of mRNA abundance for *D. melanogaster* males at ≥18BPF and PF (supplementary table S1, Supplementary Material online). Three out of four biological samples used for assaying miRNA levels of expression were randomly chosen for this purpose. Per sample, 10 µg of total RNA were reverse transcribed into cDNA using the SuperScript® Double-Stranded cDNA Labeling Kit (Invitrogen). The quality of the cDNAs was evaluated with the DNA 12000 kit (Agilent Technologies) using an Agilent 2100 Bioanalyzer. Probe labeling, hybridization, array scanning, and data extraction were performed by Roche NimbleGen Service Group in Iceland. Single color hybridizations were performed onto species-specific 12x135k NimbleGen oligonucleotide arrays corresponding to the FlyBase release 5.7 for *D. melanogaster*.

The raw mRNA expression data were pre-processed using NimbleGen's DEVA software suite, which includes background correction, quantile normalization , and summarization of probeset expression using the robust multi-array average (RMA) method . The random probes present on the NimbleGen arrays were removed prior to mRNA data analysis.

A linear mixed-effects model was used to test for differences in expression levels between developmental stages per strain as follows –

$$y_{kl} = \mu + S_k + D_l + S_k D_l + \varepsilon_{kl}$$

where y_{kl} denotes the mRNA expression for the k^{th} strain and l^{th} developmental stage. Also, μ is the baseline expression, S_k is the effect of the k^{th} strain, and D_l is

the effect of the l^{th} developmental stage. *S*, *D*, and interaction terms were implemented as fixed effects in JMP Genomics 5.0 . The difference in expression levels was considered to be statistically significant at a FDR of 0.05.

Genomic datasets. Putative target genes with conserved miRNA binding sites were retrieved from TargetScan 6.2. . Association with gain-of-function phenotypes following miRNA overexpression was determined based on positive results in at least one of three screens (eye, wing, and ubiquitous activation) performed in *D. melanogaster* . MiRNA expression ubiquity values, τ, were taken from . These values were calculated as reported based on expression values derived from 28 libraries of small RNAs representing different developmental stages and tissues of *D. melanogaster*. 1-τ values were used here, which range from 0 to 1; high 1-τ values correspond to more ubiquitously expressed miRNAs.

Functional enrichment analysis. Enrichment for Gene Ontology terms (biological process, molecular function, and cellular localization), and KEGG pathways in sets of protein-coding genes was evaluated with DAVID 6.7 . The false discovery rate adjustment was used to account for multiple testing at $P_{adj}<0.05$; the stringency was set to "Highest". As a background list, all the *D. melanogaster* genes on the NimbleGen array were used. To increase the stringency of the analysis, the list of 543 genes that turned out to show significant associations in expression with miRNAs was also used as background list in a second analysis.

Expression distance among strains. Statistically significant differences in miRNA abundance were recorded across strains in a pairwise fashion (one-way ANOVA, $P<0.01$). One hundred and twenty miRNA reporters conserved in sequence across *D. simulans, D. yakuba,* and *D. melanogaster* (including the strain Zimbabwe-109) were deemed suitable for the analysis.

Accession numbers. Small RNA sequencing output, LC Sciences expression data, and NimbleGen expression data have been deposited in the NCBI GEO database under accession numbers GSE57438, GSE55562, and GSE55398, respectively. Sequence data for protein-coding and miRNA loci have been deposited in NCBI GenBank under accession numbers KJ767237-KJ67254 and KJ774561-KJ774633.

SUPPLEMENTARY MATERIAL

Supplementary text, supplementary alignments, supplementary figures S1-S17, supplementary tables S1-S12, and supplementary datasets S1-S5 are available online.

ACKNOWLEDGMENTS

We thank Vera Chan for technical help; Kevin Thornton for providing the Zimbabwe-109 strain; Yang Shen for providing the expression breadth values; and Bryan Clifton, Anthony Long, Hsiu-Ching Ma, Antonio Marco, Andrei Tatarenkov, John True, and Zhaoxia Yu for discussions on different aspects of the manuscript. This work was supported by the National Science Foundation grants (DEB-0949365 and MCB-1157876 to J.R.) and the Foundation for Polish Science (to M.G.).

Table 1. List of miRNAs showing differential expression between ≥18BPF and PF stages

	Males						Females		
MiRNA	**CS**	**ORR**	**Sam**	**ZW**	**sim**	**yak**	**ORR**	**ZW**	**sim**
let-7-5p	PF	PF	PF	PF	PF	PF	PF	PF	PF
miR-2a-3p	PF	PF	PF	PF	PF	PF	PF	PF	PF
miR-125-5p	PF	PF	PF	PF	PF	PF	PF	PF	PF
miR-279-3p	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18B
miR-8-5p	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	PF ≥18B
miR-14-5p	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	PF ≥18B
miR-34-5p	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	PF ≥18B
miR-313-5p	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	PF ≥18B
miR-958-3p	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	PF ≥18B
miR-958-5p	F	F	F	F	F	F	F	F	PF
miR-2c-3p	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF	PF ≥18BP	ns ≥18BP	PF ≥18BP	PF ≥18BP	PF ≥18B
miR-34-3p	F ≥18BP	F ≥18BP	F ≥18BP	ne	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	PF ≥18B
miR-281-2-5p	F ≥18BP	F ≥18BP	F ≥18BP	ns ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F	PF ≥18B
miR-284-5p	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	ns ≥18BP	PF
miR-316-3p	F	F	F	F	F	F	F	F	ns
miR-100-5p	ne	PF	PF	PF	PF	PF	PF	PF	PF
miR-2b-3p	PF	PF	PF	PF	ns	PF	PF	PF	ns
miR-2b-2-5p	PF	ns	PF	PF	PF	ns	PF	PF	PF
miR-92a-3p	PF	PF	PF	PF	PF	ns	PF	PF	ns
miR-276b-3p	PF ≥18BP	PF ≥18BP	PF ≥18BP	PF	PF ≥18BP	PF ≥18BP	PF ≥18BP	ns	ns ≥18B
miR-956-3p	F ≥18BP	F	F ≥18BP	ns ≥18BP	F ≥18BP	F ≥18BP	F ≥18BP	ns ≥18BP	PF
miR-1000-3p	F	ne	F	F	F	F	F	F	ne
miR-92b-3p	PF	PF	PF ≥18BP	PF ≥18BP	ns ≥18BP	ns	PF	PF	ns ≥18B
miR-289-5p	PF	ns	F	F	F	ns	PF	ns	PF
miR-306-5p	PF	ns	PF	PF	PF	ns	PF	PF	ns
miR-2a-1-5p	PF	ne	PF	PF	PF	ns	ns	PF	ns
miR-7-5p	PF	ns	PF	PF	ns	ns	ns	PF	PF

33

Table 1. List of miRNAs showing differential expression between ≥18BPF and PF stages

	Males						Females		
MiRNA	**CS**	**ORR**	**Sam**	**ZW**	**sim**	**yak**	**ORR**	**ZW**	**sim**
miR-9a-5p	PF	ns	PF	PF	ns	ns	PF	PF	ns
miR-305-5p	ns	PF	PF	PF	ns	PF	PF	ns	ns
miR-1012-3p	PF	ns	PF	PF	N/A	ne	PF	ns	N/A
miR-276a-3p	ns	PF	PF	ns	ns	PF	PF	ns	ns
miR-1000-5p	ns	PF	PF	PF	PF	ns	ns	ns	ns
miR-1010-3p	ne	ne	ne	PF	N/A	PF	PF	PF	N/A
miR-79-3p	PF	ne	ns	ne	PF	ns	ns	PF	ns
miR-9a-3p	ns	ns	PF	PF ≥18BP	ns	ns	ns ≥18BP	PF ≥18BP	ns
miR-964-5p	ns	ns	ns	F	ns	ns	F	F	ne
miR-988-3p *miR-314-5p-*	PF ≥18BP	ne	ne	ne	ne	N/A	ns ≥18BP	PF	ns
dme *miR-2493-5p-*	F ≥18BP	ne	ne	ns ≥18BP	N/A	N/A	F	ns	N/A
dme	F	ne	ns	F	N/A	N/A ≥18BP	ns ≥18BP	ns	N/A
miR-277-3p	ns	ns	ns	ns	ns	F	F	ns	ns
miR-998-3p *miR-314-5p-*	ns	ns	PF	PF	ns ≥18BP	ns ≥18BP	ns	ns	ns
dsi	N/A	N/A	N/A	N/A	F	F	N/A	N/A	ne
miR-2a-2-5p	PF	ne	ne	ne	ne	ns	ns	ne	ne
bantam-5p	ns	ns	ne	PF	ns ≥18BP	ns	ns	ne	ns
miR-10-3p	ns	ns	ns	ns	F	ns	ns	ns	ns ≥18B
miR-10-5p	ns	ns	ns	ns	ns	ns	ns	ns	PF
miR-11-3p	ns	ns	ns	PF	ns	ns	ns	ns	ns
miR-31a-5p	ns	ns	ns	PF	ns	ns	ns	ns	ns
miR-31b-5p	ns	ns	ns	PF	ns	ns	ns	ns	ns
miR-252-5p	ns	ns	ns	PF	ns	ns	ns	ns	ns
miR-275-3p	ns	ns	ns	ns	ns	ns	ns	PF	ns
miR-996-5p	ns	ns	ns	PF	ns	ns	ns	ns	ns
miR-998-5p	ns	ns	ns	PF	ns	ns	ne	ne	ns
miR-9b-5p	ne	ne	ne	ne	N/A	ne	ne	PF	N/A
miR-993-3p	ne	PF	ne	ne	ns	N/A	ne	ne	ns

CS, Canton-S; ORR, Oregon-R; Sam, Samarkand; ZW, Zimbabwe-109; sim, *D. simulans*; yak, *D. yakuba*.

Differences in expression between developmental stages at P_{adj}<0.01 (Materials and Methods). ne, non-expressed; ns, non-developmentally enriched; PF, enriched at PF; ≥18BPF, enriched at ≥18BPF; N/A, the analysis is not applicable to this miRNA due to its absence or sequence differences.

763

35

FIGURE LEGENDS

Fig. 1. Developmental changes in miRNA expression at the onset of metamorphosis across six *Drosophila* males. (A) Differences in expression levels between ≥18BPF and PF. *x*-axis, difference in normalized $\log_2$-transformed expression levels between ≥18BPF and PF; *y*-axis, significance of the difference as $-\log_{10}(P \text{ value})$. Statistically significant differences were determined using a one-way ANOVA. Red dotted line, $P_{adj}=0.01$. (B) Examples of variation in miRNA abundance in males. The average expression level and the standard error of the mean are shown. Left, interstrain difference affecting the developmental expression pattern. *miR-956-3p* is ≥18BPF-enriched in all strains but in Zimbabwe-109. In this strain, there is no significant difference in expression level between stages due to an increase in the level of expression at PF ($P_{adj}=2\text{x}10^{-8}$). Right, interstrain difference not associated with differences in developmental expression patterns. *miR-1012-5p* is non-developmentally enriched in all strains but statistical significant differences in expression level were confirmed at PF ($P_{adj}=7.1\text{x}10^{-3}$). Canton-S, Oregon-R, Samarkand, and Zimbabwe-109 are strains of *D. melanogaster.*

Fig. 2. Distinct patterns of functional divergence in miRNA expression levels at the onset of *Drosophila* metamorphosis. (A) Expression distance in miRNA abundance and divergence time are not related in the same way at ≥18BPF and PF. Branch lengths from Neighbor-Joining trees were used as a surrogate for divergence time among strains. Unlike at PF, expression changes do not accumulate in a linear fashion over time at ≥18BPF. This result holds when estimated divergence times among strains are used instead of branch lengths and when an alternative statistical approach is used (supplementary text and figure S13, Supplementary Material online). (B) Comparison of the variation in expression levels among

36

miRNAs sets with different developmental expression patterns. Only miRNAs with consistent developmental expression patterns across *D. melanogaster* strains and identical in sequence across species were considered.

Fig. 3. Evolutionary modes of miRNA expression at the onset of male metamorphosis in the *D. melanogaster* species subgroup. A) Counts of miRNAs with patterns of intra- and inter-specific variation compatible with particular evolutionary modes. B) MiRNAs illustrating four main evolutionary modes. Box plots are used to show the dispersion around the median level of expression for each strain by miRNA combination. CS, Canton-S; ORR, Oregon-R; Sam, Samarkand; Zw, Zimbabwe-109; and sim, *D. simulans*. Within and between species differences in expression levels were tested respectively by one-way ANOVA. Whether evidence of significant intra- and inter-specific variation was found is indicated on the *x*- and *y*-axis respectively. See supplementary text (Supplementary Material online) for the rationale followed to categorize each miRNA under a particular evolutionary mode. MiRNAs showing non-significant and significant differences in expression between *D. melanogaster* and *D. simulans* are shown on top and bottom, respectively. Depending on the combination of significant intra- and inter-specific differences in expression levels, miRNAs are categorized as evolving under stabilizing selection (top left), directional selection (bottom left), genetic drift (bottom right), or associated with complex scenarios (top right). *y*-axis, $\log_2$-tranformed expression values in an arbitrary scale. Examples shown correspond to miRNAs expressed at the PF stage.

Fig. 4. Relationship between the level of miRNA expression polymorphism in four males of *D. melanogaster* and evolutionary age, gain-of-function phenotype, and expression breadth. (A) Young miRNAs exhibit higher levels of polymorphism in gene expression than ancient miRNAs in *D. melanogaster* males. Left, ≥18BPF;

right, PF. Average±SEM values of expression polymorphism are indicated on top of the chart. *Ancient* class, miRNAs inferred to have been already present in the ancestor to all currently existing *Drosophila* species and other insects and/or metazoans. *Young* class, miRNAs inferred to have been present in the ancestor to all *Drosophila* species but not in that of other insects or metazoans, or to have evolved subsequently during the radiation of the genus *Drosophila.* (B) Gain-of-function phenotype upon miRNA overexpression. Showing a gain-of-function phenotype in at least one of the three screens performed did result in a significantly lower level of expression polymorphism at both developmental stages being specially acute at PF. (C) Expression breadth. As miRNA expression ubiquity increases, the level of expression polymorphism decreases. This pattern is observed at PF but not at ≥18BPF. The use of a non-parametric correlation statistic made no difference (Spearman's ρ; PF, P=0.006; ≥18BPF, P=0.154). Expression ubiquity values, 1 - τ, are close to 1 if expression is ubiquitous and close to 0 if expression is rather specific. τ values were taken from .

Fig. 5. Relationship between the percentage of miRNA-mRNA negative expression associations and miRNA developmental expression pattern. (A) Distribution of the number of miRNAs based on the percentage of negative significant associations with mRNA exemplars relative to the total number of statistically significant associations. Thirty-eight miRNAs exhibited a proportion of negative over positive associations deviating significantly from the 1:1 ratio while five miRNAs did not. (B) MiRNAs with higher percentage of negative expression associations to mRNAs are more likely to be PF-enriched than ≥18BPF-enriched (Spearman's Rho=0.6895, P<0.0001). *x*-axis, developmental expression index. This index is calculated by counting the number of *D. melanogaster* strains that exhibit ≥18BPF- or PF-

38

enrichment and adding -1 or +1, respectively. For example, an index value of -3 indicates that the miRNA was categorized as ≥18BPF-enriched in three strains and non-developmentally enriched in one strain. *y*-axis, percentage of negative expression associations shown by miRNAs. (C) Ideogram representing separately the expression associations between miRNAs and mRNAs across the two developmental timepoints assayed. Left panel, the positive association between miRNA (blue) and mRNA (red) expression levels identified in miRNAs with lower expression level at PF may often reflect the concurrently transcriptional down-regulation of both kind of molecules. Right panel, the up-regulation of a miRNA at PF may be important for dampening mRNAs from genes that are not to be expressed at this stage. This second pattern is more likely to reflect a causal regulatory relationship between the miRNA and the mRNA, which can be exerted directly through miRNA binding or indirectly throughout a third party.

39

REFERENCES

Ambros V 2004. The functions of animal microRNAs. Nature 431: 350-355.

Ameres SL, Zamore PD 2013. Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol 14: 475-488. doi: 10.1038/nrm3611

Andres AJ, Thummel CS 1994. Chapter 29 Methods for quantitative analysis of transcription in larvae and prepupae. Methods in Cell Biology 44: 565-573.

Arbeitman MN, Furlong EE, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP 2002. Gene expression during the life cycle of Drosophila melanogaster. Science 297: 2270-2275.

Arif S, Murat S, Almudi I, Nunes MD, Bortolamiol-Becet D, McGregor NS, Currie JM, Hughes H, Ronshaugen M, Sucena E, et al. 2013. Evolution of mir-92a underlies natural morphological variation in Drosophila melanogaster. Curr Biol 23: 523-528. doi: 10.1016/j.cub.2013.02.018

Axtell MJ, Westholm JO, Lai EC 2011. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome Biol 12: 221. doi: 10.1186/gb-2011-12-4-221

Baehrecke EH 1996. Ecdysone signaling cascade and regulation of Drosophila metamorphosis. Arch Insect Biochem Physiol 33: 231-244.

Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP 2008. The impact of microRNAs on protein output. Nature 455: 64-71.

874Bainbridge SP, Bownes M 1981. Staging the metamorphosis of Drosophila
875melanogaster. J Embryol Exp Morphol 66: 57-80.

876Bartel DP 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell
877116: 281-297.

878Bartel DP 2009. MicroRNAs: target recognition and regulatory functions. Cell 136:
879215-233.

880Bashirullah A, Pasquinelli AE, Kiger AA, Perrimon N, Ruvkun G, Thummel CS 2003.
881Coordinate regulation of small temporal RNAs at the onset of Drosophila
882metamorphosis. Dev Biol 259: 1-8.

883Beckstead RB, Lam G, Thummel CS 2005. The genomic response to 20-
884hydroxyecdysone at the onset of Drosophila metamorphosis. Genome Biol 6: R99.

885Berezikov E, Robine N, Samsonova A, Westholm JO, Naqvi A, Hung JH, Okamura K,
886Dai Q, Bortolamiol-Becet D, Martin R, et al. 2010. Deep annotation of Drosophila
887melanogaster microRNAs yields insights into their processing, modification, and
888emergence. Genome Res 21: 203-215.

889Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, Cuppen E, Plasterk
890RH 2006. Diversity of microRNAs in human and chimpanzee brain. Nat Genet 38:
891375-1377.

892Bolstad BM, Irizarry RA, Astrand M, Speed TP 2003. A comparison of normalization
893methods for high density oligonucleotide array data based on variance and bias.
894Bioinformatics 19: 185-193.

Borel C, Deutsch S, Letourneau A, Migliavacca E, Montgomery SB, Dimas AS, Vejnar CE, Attar H, Gagnebin M, Gehrig C, et al. 2011. Identification of cis- and trans-regulatory variation modulating microRNA expression levels in human fibroblasts. Genome Res 21: 68-73. doi: 10.1101/gr.109371.110

Bushati N, Stark A, Brennecke J, Cohen SM 2008. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. Curr Biol 18: 501-506. doi: 10.1016/j.cub.2008.02.081

Carrington JC, Ambros V 2003. Role of microRNAs in plant and animal development. Science 301: 336-338.

Caygill EE, Johnston LA 2008. Temporal regulation of metamorphic processes in Drosophila by the let-7 and miR-125 heterochronic microRNAs. Curr Biol 18: 943-950.

Chawla G, Sokol NS 2012. Hormonal activation of let-7-C microRNAs via EcR is required for adult Drosophila melanogaster morphology and function. Development 139: 1788-1797. doi: 10.1242/dev.077743

Chen K, Rajewsky N 2007. The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 8: 93-103.

Chen Z, Chen LY, Dai HY, Wang P, Gao S, Wang K 2012. miR-301a promotes pancreatic cancer cell proliferation by directly inhibiting Bim expression. J Cell Biochem 113: 3229-3235. doi: 10.1002/jcb.24200

Duursma AM, Kedde M, Schrier M, le Sage C, Agami R 2008. miR-148 targets human DNMT3b protein coding region. RNA 14: 872-877.

42

Ebert MS, Sharp PA 2012. Roles for microRNAs in conferring robustness to biological processes. Cell 149: 515-524. doi: 10.1016/j.cell.2012.04.005

Filipowicz W, Bhattacharyya SN, Sonenberg N 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 9: 102-114.

Flatt T, Heyland A, Rus F, Porpiglia E, Sherlock C, Yamamoto R, Garbuzov A, Palli SR, Tatar M, Silverman N 2008. Hormonal regulation of the humoral innate immune response in Drosophila melanogaster. J Exp Biol 211: 2712-2724. doi: 10.1242/jeb.014878

Forman JJ, Legesse-Miller A, Coller HA 2008. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc Natl Acad Sci U S A 105: 14879-14884.

Fullaondo A, Lee SY 2012. Identification of putative miRNA involved in Drosophila melanogaster immune response. Dev Comp Immunol 36: 267-273. doi: 10.1016/j.dci.2011.03.034

Garbuzov A, Tatar M 2010. Hormonal regulation of Drosophila microRNA let-7 and miR-125 that target innate immunity. Fly (Austin) 4: 306-311.

Gomez-Orte E, Belles X 2009. MicroRNA-dependent metamorphosis in hemimetabolan insects. Proc Natl Acad Sci U S A 106: 21678-21682. doi: 10.1073/pnas.0907391106

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of Drosophila melanogaster. Nature 471: 473-479. doi: 10.1038/nature09715

Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Chapman EJ, Carrington JC, Chen X, Wang XJ, Chen ZJ 2009. Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proc Natl Acad Sci U S A 106: 17835-17840.

Hornstein E, Shomron N 2006. Canalization of development by microRNAs. Nat Genet 38 Suppl: S20-24. doi: 10.1038/ng1803

Hsouna A, Lawal HO, Izevbaye I, Hsu T, O'Donnell JM 2007. Drosophila dopamine synthesis pathway genes regulate tracheal morphogenesis. Dev Biol 308: 30-43. doi: 10.1016/j.ydbio.2007.04.047

Hu HY, Guo S, Xi J, Yan Z, Fu N, Zhang X, Menzel C, Liang H, Yang H, Zhao M, et al. 2011. MicroRNA expression and regulation in human, chimpanzee, and macaque brains. PLoS Genet 7: e1002327. doi: 10.1371/journal.pgen.1002327

Huang da W, Sherman BT, Lempicki RA 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57.

Huang RS, Gamazon ER, Ziliak D, Wen Y, Im HK, Zhang W, Wing C, Duan S, Bleibel WK, Cox NJ, et al. 2011. Population differences in microRNA expression and biological implications. RNA Biol 8: 692-701. doi: 10.4161/rna.8.4.16029

957Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP
9582003. Exploration, normalization, and summaries of high density oligonucleotide
959array probe level data. Biostatistics 4: 249-264.

960Jin H, Kim VN, Hyun S 2012. Conserved microRNA miR-8 controls body size in
961response to steroid signaling in Drosophila. Genes Dev 26: 1427-1432. doi: 10.1101/
962gad.192872.112

963Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G 2001. The
964contributions of sex, genotype and age to transcriptional variance in Drosophila
965melanogaster. Nat Genet 29: 389-395. doi: 10.1038/ng766

966Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, Weiss G,
967Lachmann M, Paabo S 2005. Parallel patterns of evolution in the genomes and
968transcriptomes of humans and chimpanzees. Science 309: 1850-1854. doi: 10.1126/
969science.1108296

970Kheradpour P, Stark A, Roy S, Kellis M 2007. Reliable prediction of regulator targets
971using 12 Drosophila genomes. Genome Res 17: 1919-1931.

972Kozomara A, Griffiths-Jones S 2011. miRBase: integrating microRNA annotation and
973deep-sequencing data. Nucleic Acids Res 39: D152-157. doi: 10.1093/nar/gkq1027

974Langmead B, Trapnell C, Pop M, Salzberg SL 2009. Ultrafast and memory-efficient
975alignment of short DNA sequences to the human genome. Genome Biol 10: R25.
976doi: 10.1186/gb-2009-10-3-r25

977Lappalainen T, Sammeth M, Friedlander MR, t Hoen PA, Monlong J, Rivas MA,
978Gonzalez-Porta M, Kurbatova N, Griebel T, Ferreira PG, et al. 2013. Transcriptome

979and genome sequencing uncovers functional variation in humans. Nature 501: 506-
980511. doi: 10.1038/nature12531

981Lebo MS, Sanders LE, Sun F, Arbeitman MN 2009. Somatic, germline and sex
982hierarchy regulated gene expression during Drosophila metamorphosis. BMC
983Genomics 10: 80.

984Li H, Stephan W 2006. Inferring the demographic history and rate of adaptive
985substitution in Drosophila. PLoS Genet 2: e166. doi: 10.1371/journal.pgen.0020166

986Liang H, Li WH 2009. Lowly expressed human microRNA genes evolve rapidly. Mol
987Biol Evol 26: 1195-1198. doi: 10.1093/molbev/msp053

988Livak KJ, Schmittgen TD 2001. Analysis of relative gene expression data using real-
989time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.
990doi: 10.1006/meth.2001.1262

991Lu J, Clark AG 2012. Impact of microRNA regulation on variation in human gene
992expression. Genome Res 22: 1243-1254. doi: 10.1101/gr.132514.111

993Lu J, Shen Y, Wu Q, Kumar S, He B, Shi S, Carthew RW, Wang SM, Wu CI 2008a. The
994birth and death of microRNA genes in Drosophila. Nat Genet 40: 351-355.

995Lu M, Zhang Q, Deng M, Miao J, Guo Y, Gao W, Cui Q 2008b. An analysis of human
996microRNA and disease associations. PLoS One 3: e3420. doi:
99710.1371/journal.pone.0003420

998Lyu Y, Shen Y, Li H, Chen Y, Guo L, Zhao Y, Hungate E, Shi S, Wu CI, Tang T 2014.
999New MicroRNAs in Drosophila-Birth, Death and Cycles of Adaptive Evolution. PLoS
1000Genet 10: e1004096. doi: 10.1371/journal.pgen.1004096

1001Maroni G, Stamey SC 1983. Use of blue food to select synchronous late 3rd instar
1002larvae. Drosophila Information Service 59: 142-143.

1003Nagalla S, Shaw C, Kong X, Kondkar AA, Edelstein LC, Ma L, Chen J, McKnight GS,
1004Lopez JA, Yang L, et al. 2011. Platelet microRNA-mRNA coexpression profiles
1005correlate with platelet reactivity. Blood 117: 5189-5197. doi: 10.1182/blood-2010-
100609-299719

1007Orom UA, Nielsen FC, Lund AH 2008. MicroRNA-10a binds the 5'UTR of ribosomal
1008protein mRNAs and enhances their translation. Mol Cell 30: 460-471.

1009Pancratov R, Peng F, Smibert P, Yang S, Jr., Olson ER, Guha-Gilford C, Kapoor AJ,
1010Liang FX, Lai EC, Flaherty MS, et al. 2013. The miR-310/13 cluster antagonizes beta-
1011catenin function in the regulation of germ and somatic cell differentiation in the
1012Drosophila testis. Development 140: 2904-2916. doi: 10.1242/dev.092817

1013Parts L, Hedman AK, Keildson S, Knights AJ, Abreu-Goodger C, van de Bunt M,
1014Guerra-Assuncao JA, Bartonicek N, van Dongen S, Magi R, et al. 2012. Extent,
1015causes, and consequences of small RNA expression variation in human adipose
1016tissue. PLoS Genet 8: e1002704. doi: 10.1371/journal.pgen.1002704

1017Rao Y, Lee Y, Jarjoura D, Ruppert AS, Liu CG, Hsu JC, Hagan JP 2008. A comparison
1018of normalization techniques for microRNA microarray data. Stat Appl Genet Mol Biol
10197: Article22.

1020Rifkin SA, Houle D, Kim J, White KP 2005. A mutation accumulation assay reveals a
1021broad capacity for rapid evolution of gene expression. Nature 438: 220-223. doi:
102210.1038/nature04114

1023Rifkin SA, Kim J, White KP 2003. Evolution of gene expression in the Drosophila
1024melanogaster subgroup. Nat Genet 33: 138-144.

1025Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC 2007. Evolution,
1026biogenesis, expression, and target predictions of a substantially expanded set of
1027Drosophila microRNAs. Genome Res 17: 1850-1864.

1028Ryazansky SS, Gvozdev VA, Berezikov E 2011. Evidence for post-transcriptional
1029regulation of clustered microRNAs in Drosophila. BMC Genomics 12: 371. doi:
103010.1186/1471-2164-12-371

1031Schertel C, Rutishauser T, Forstemann K, Basler K 2012. Functional characterization
1032of Drosophila microRNAs by a novel in vivo library. Genetics 192: 1543-1552. doi:
103310.1534/genetics.112.145383

1034Schnall-Levin M, Zhao Y, Perrimon N, Berger B 2010. Conserved microRNA targeting
1035in Drosophila is as widespread in coding regions as in 3'UTRs. Proc Natl Acad Sci U S
1036A 107: 15751-15756. doi: 10.1073/pnas.1006172107

1037Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N 2008.
1038Widespread changes in protein synthesis induced by microRNAs. Nature 455: 58-63.
1039doi: 10.1038/nature07228

1040Sempere LF, Dubrovsky EB, Dubrovskaya VA, Berger EM, Ambros V 2002. The
1041expression of the let-7 small regulatory RNA is controlled by ecdysone during
1042metamorphosis in Drosophila melanogaster. Dev Biol 244: 170-179.

1043Sempere LF, Sokol NS, Dubrovsky EB, Berger EM, Ambros V 2003. Temporal
1044regulation of microRNA expression in Drosophila melanogaster mediated by
1045hormonal signals and broad-Complex gene activity. Dev Biol 259: 9-18.

1046Shen Y, Lv Y, Huang L, Liu W, Wen M, Tang T, Zhang R, Hungate E, Shi S, Wu CI
10472011. Testing hypotheses on the rate of molecular evolution in relation to gene
1048expression using microRNAs. Proc Natl Acad Sci U S A 108: 15942-15947. doi:
104910.1073/pnas.1110098108

1050Sokol NS, Xu P, Jan YN, Ambros V 2008. Drosophila let-7 microRNA is required for
1051remodeling of the neuromusculature during metamorphosis. Genes Dev 22: 1591-
105221596.

1053Takahashi A 2013. Pigmentation and behavior: potential association through
1054pleiotropic genes in Drosophila. Genes Genet Syst 88: 165-174.

1055Tamura K, Subramanian S, Kumar S 2004. Temporal patterns of fruit fly (Drosophila)
1056evolution revealed by mutation clocks. Mol Biol Evol 21: 36-44. doi:
105710.1093/molbev/msg236

1058Tang T, Kumar S, Shen Y, Lu J, Wu ML, Shi S, Li WH, Wu CI 2010. Adverse
1059interactions between micro-RNAs and target genes from different species. Proc Natl
1060Acad Sci U S A 107: 12935-12940. doi: 10.1073/pnas.1007591107

1061Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I 2008. MicroRNAs to Nanog, Oct4
1062and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 455:
106331124-1128.

1064Thummel CS 1996. Flies on steroids--Drosophila metamorphosis and the
1065mechanisms of steroid hormone action. Trends Genet 12: 306-310.

1066Thummel CS 2001. Molecular mechanisms of developmental timing in C. elegans
1067and Drosophila. Dev Cell 1: 453-465.

1068Varghese J, Cohen SM 2007. microRNA miR-14 acts to modulate a positive
1069autoregulatory loop controlling steroid hormone signaling in Drosophila. Genes Dev
107021: 2277-2282.

1071Waddington CH 1959. Canalization of development and genetic assimilation of
1072acquired characters. Nature 183: 1654-1655.

1073Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ
10742009. The deep evolution of metazoan microRNAs. Evol Dev 11: 50-68.

1075White KP, Rifkin SA, Hurban P, Hogness DS 1999. Microarray analysis of Drosophila
1076development during metamorphosis. Science 286: 2179-2184.

1077Wittkopp PJ, Beldade P 2009. Development and evolution of insect pigmentation:
1078genetic mechanisms and the potential consequences of pleiotropy. Semin Cell Dev
1079Biol 20: 65-71. doi: 10.1016/j.semcdb.2008.10.002

1080Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C,
1081Paules RS 2001. Assessing gene significance from cDNA microarray expression data
1082via mixed models. J Comput Biol 8: 625-637. doi: 10.1089/106652701753307520

1083Wu CI, Shen Y, Tang T 2009. Evolution under canalization and the dual roles of
1084microRNAs: a hypothesis. Genome Res 19: 734-743. doi: 10.1101/gr.084640.108

Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E, et al. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics 21: 650-659. doi: 10.1093/bioinformatics/bti042

51

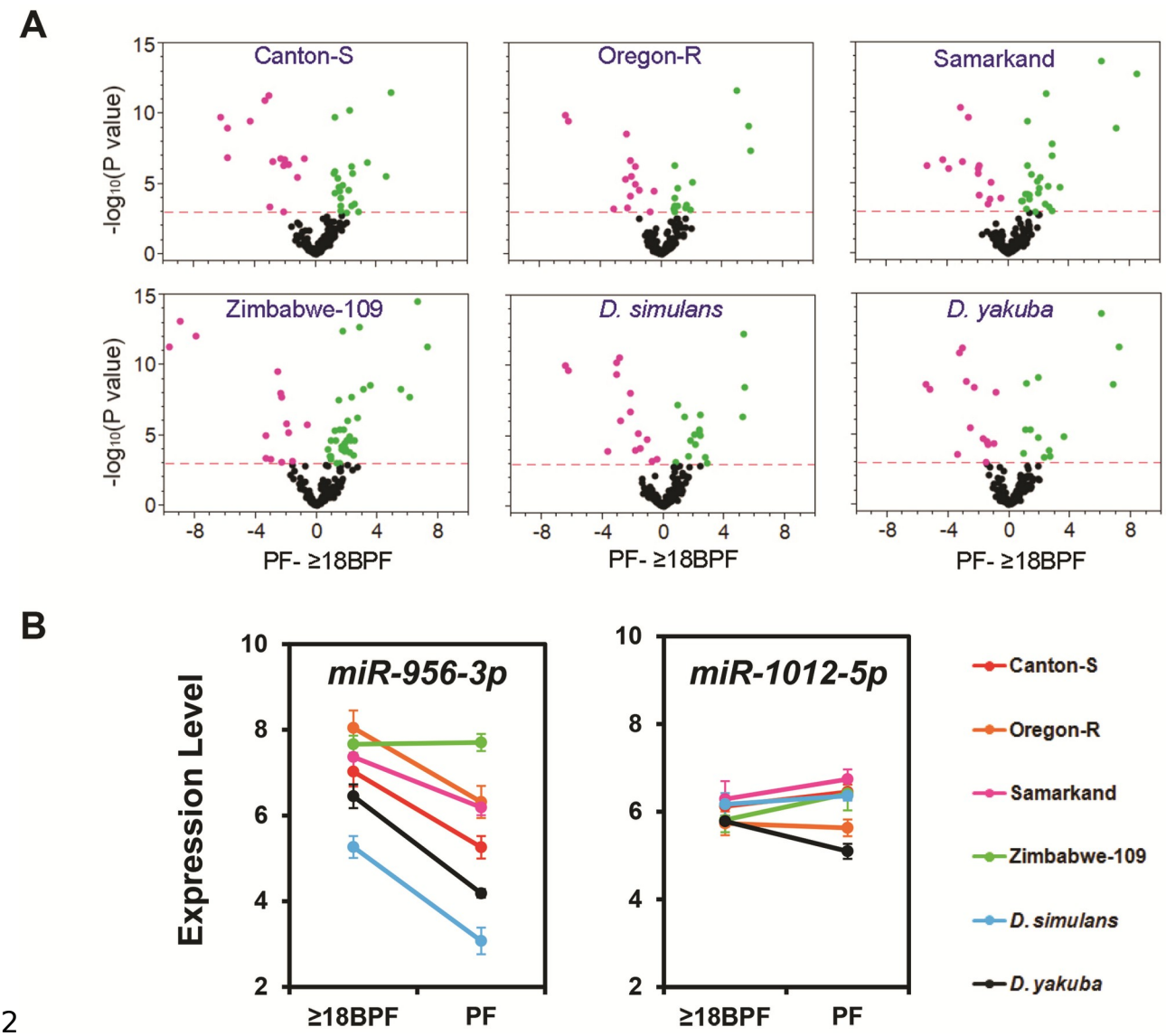

A
Canton-S
Oregon-R
Samarkand
Zimbabwe-109
D. simulans
D. yakuba
-log10(P value)
PF- ≥18BPF
B
miR-956-3p
miR-1012-5p
Expression Level
≥18BPF
PF
Canton-S
Oregon-R
Samarkand
Zimbabwe-109
D. simulans
D. yakuba

Fig. 1

52

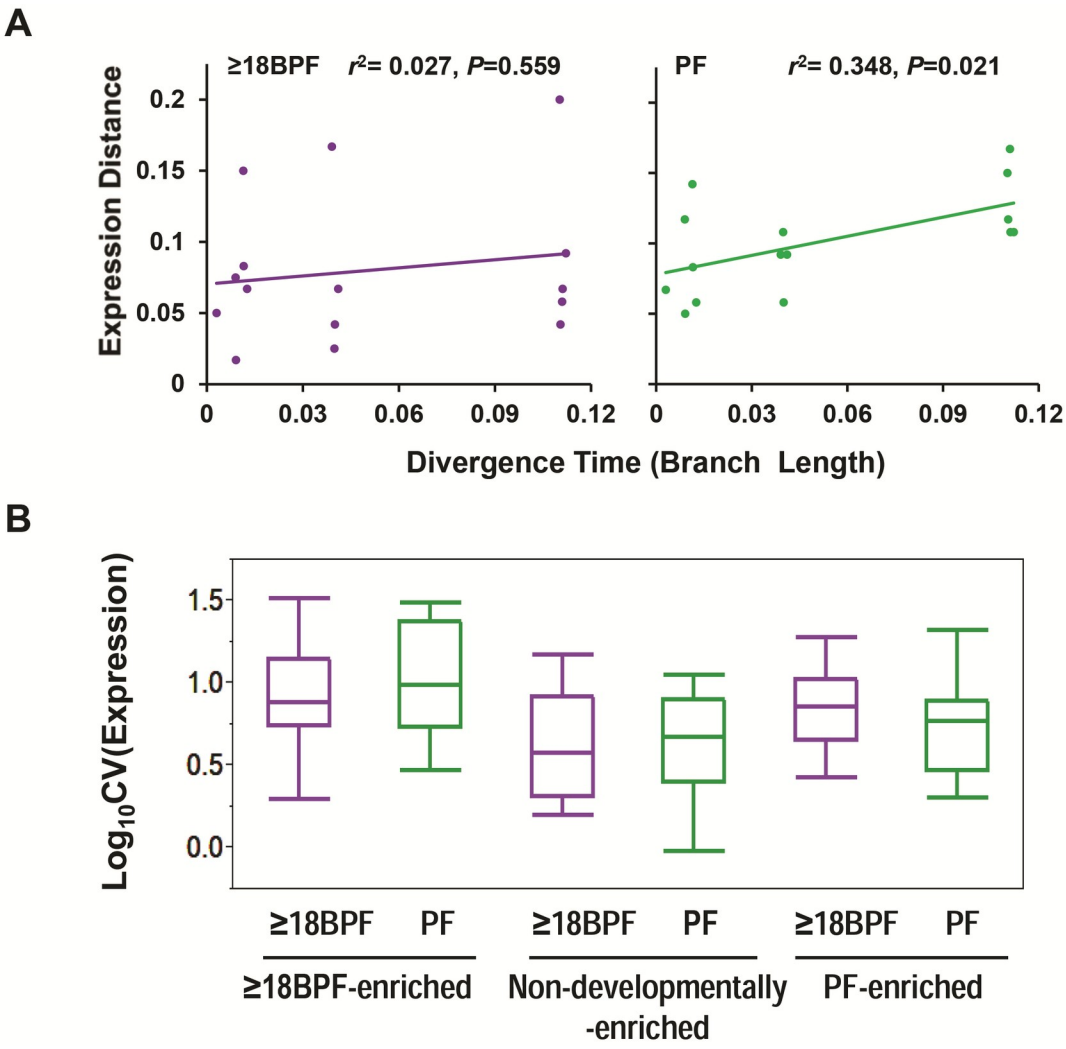

A
≥18BPF r^2= 0.027, P=0.559
PF r^2= 0.348, P=0.021
Expression Distance
0.2
0.15
0.1
0.05
0
0 0.03 0.06 0.09 0.12
0 0.03 0.06 0.09 0.12
Divergence Time (Branch Length)
B
Log_{10}CV(Expression)
1.5
1.0
0.5
0.0
≥18BPF PF ≥18BPF PF ≥18BPF PF
≥18BPF-enriched
Non-developmentally -enriched
PF-enriched

Fig. 2

53

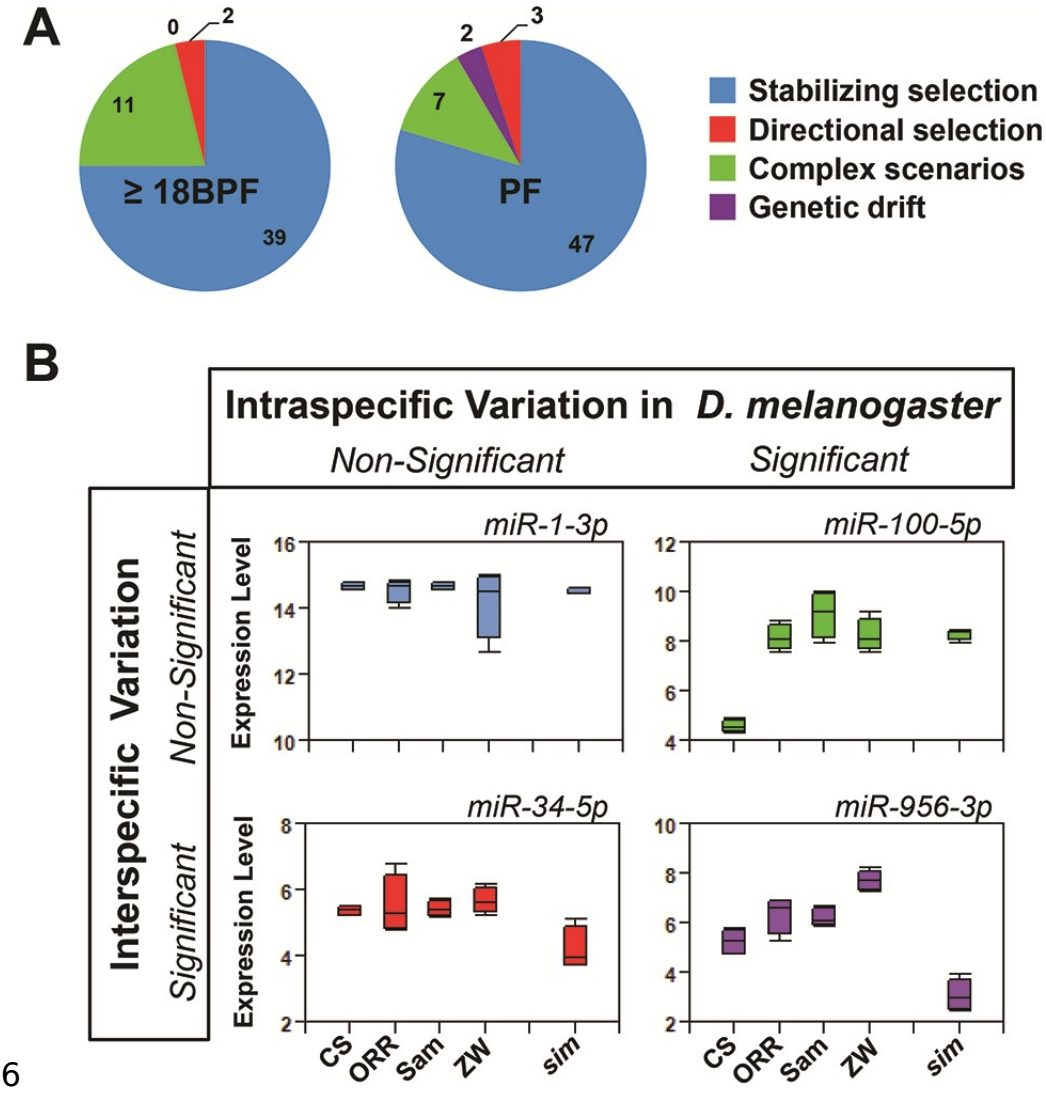
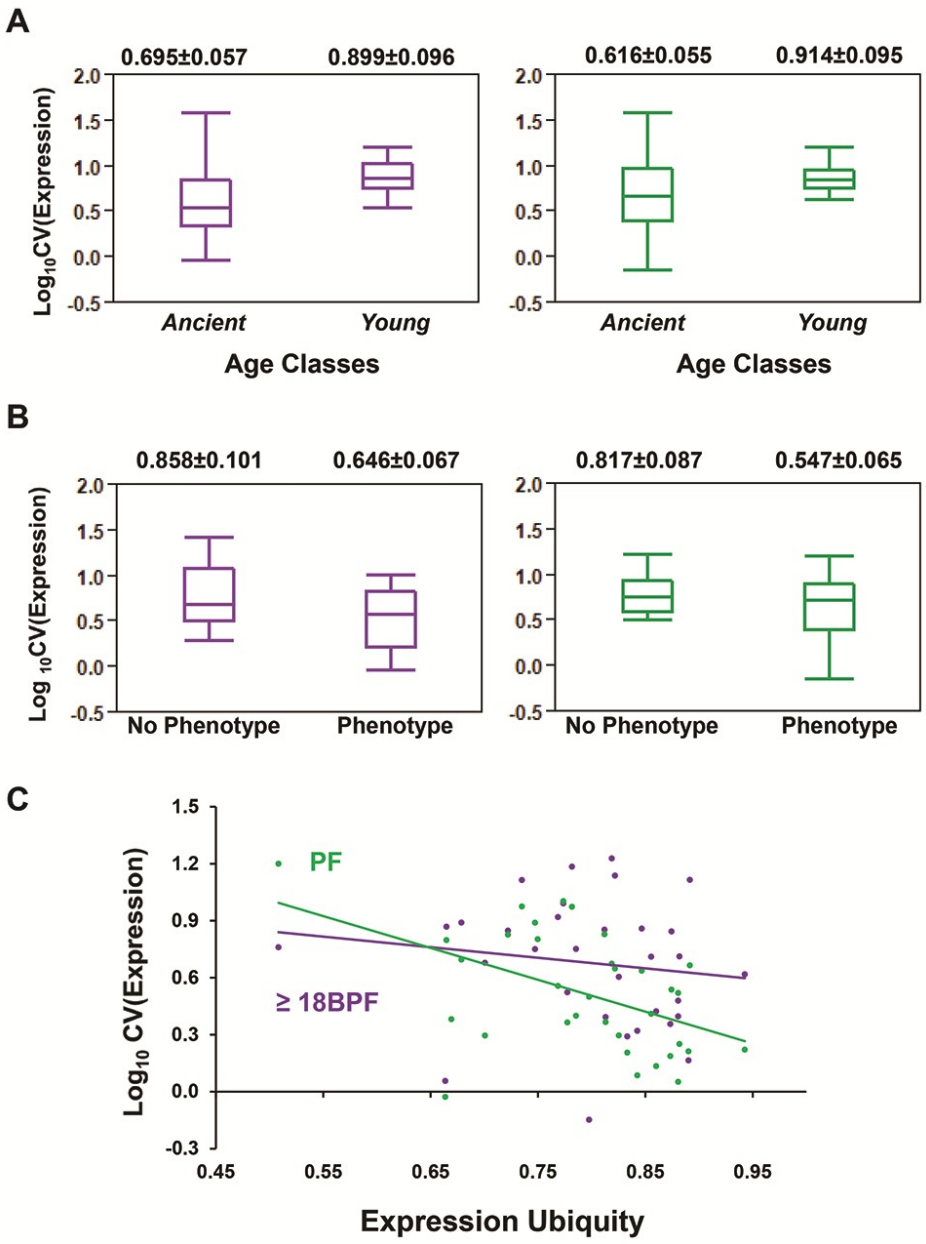

A
0
2
11
≥ 18BPF
39
2
3
7
PF
47
Stabilizing selection
Directional selection
Complex scenarios
Genetic drift
B
Intraspecific Variation in D. melanogaster
Non-Significant
Significant
Interspecific Variation
Non-Significant
Significant
miR-1-3p
miR-100-5p
miR-34-5p
miR-956-3p
Expression Level
CS
ORR
Sam
ZW
sim

Fig. 3

54



1113

1114**Fig. 4**

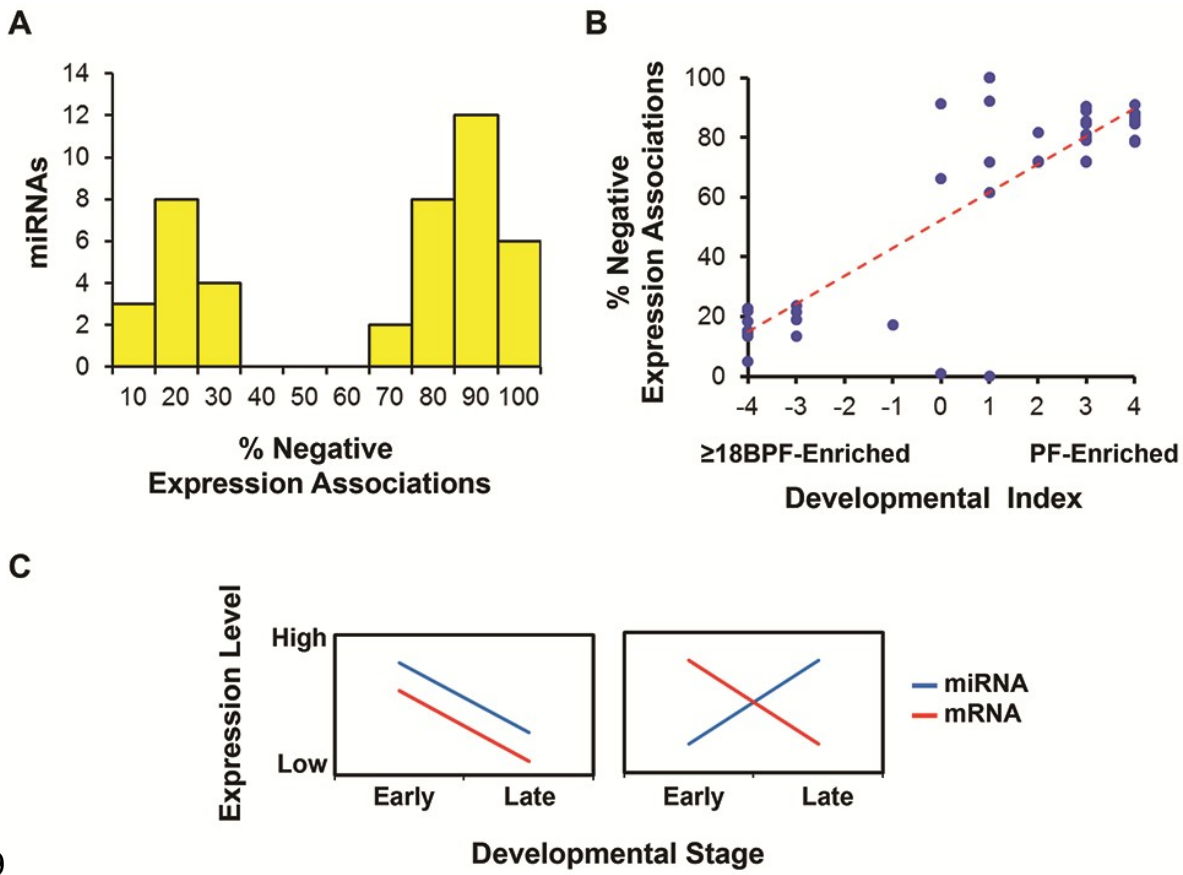
1115

1116

1117

1118

55



1119

1120**Fig. 5**

1121

1122

56