

Regulation of Large Number of Weak Targets—New Insights from Twin-miRNAs

Yixin Zhao^{1,3}, Pei Lin¹, Zhongqi Liufu¹, Hao Yang¹, Yang Lyu^{1,4}, Xu Shen¹, Chung-I Wu^{1,2,*}, and Tian Tang^{1,*}

¹State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China

²Department of Ecology and Evolution, University of Chicago

³Present address: Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA.

⁴Present address: Department of Molecular and Integrative Physiology and Geriatrics Center, University of Michigan, Biomedical Sciences and Research Building, Ann Arbor, MI

*Corresponding authors: E-mails: ciwu@uchicago.edu; lsstt@mail.sysu.edu.cn.

Accepted: April 23, 2018

Data deposition: All RNA-seq data generated for this study was deposited in GEO under the accession number GSE97364.

Abstract

Each animal microRNA (miRNA) targets many genes for repression. Down-regulation of most of these targets is weak and has no detectable individual phenotypic effect. Whether this extensive weak repression is biologically relevant is a central issue in the debate on miRNA functionality. In the “small (target) pool” view, weak repression is nonfunctional and should be gradually removed during evolution. However, since the selective advantage of removing individual targets is small, testing this hypothesis is a challenge. We propose a novel approach by using miRNAs we call twin-miRs, which produce two mature products from the hairpin of the same miRNA precursor. Loss of the minor miR partner would affect all its targets and thus could be visible to selection. Since the minor miRs repress all their targets weakly, the “small pool” hypothesis would predict the elimination of twin-miRs over time. Surveying and sequencing 45 small RNA libraries in *Drosophila*, we found that nearly 40% of miRNAs produce twin-miRs. The minor forms are expressed in nontrivial abundance and repress their targets weakly. Interestingly, twin-miRs are often evolutionarily old, highly conserved, and comparable to solo-miRs in expression. Since there is no measurable trend toward reduction in target pool size, we conclude that at least some of the weak repression interactions are functional. A companion study using the May–Wigner theory of network stability suggests that distributed weak repression cumulatively contributes to stability of gene regulatory networks.

Key words: microRNA, twin miRNA, target number, evolution.

Introduction

MicroRNAs (miRNAs) are a class of (~22 nt) endogenous RNAs that posttranscriptionally repress target gene expression (Bartel 2004; Bushati and Cohen 2007). In animals, miRNAs recognize their targets by base pairing between the so-called seed region (2–8 nt of the miRNA) and the sites located in 3′-UTRs of target transcripts (Bartel 2009). Hundreds of targets are usually predicted for each miRNA (Bartel 2009), the vast majority of them are only weakly repressed (Baek et al. 2008; Selbach et al. 2008; Hausser and Zavolan 2014). Despite their large number, only a few targets seemingly contribute to phenotypes associated with miRNA deletions (Karres et al. 2007; Varghese et al. 2010; Ecsedi et al. 2015). These

observations gave rise to the view that most miRNA-target interactions are biologically irrelevant (Pinzon et al. 2017). However, since negative results on phenotypic measurements do not necessarily imply lack of fitness consequences in a population over evolutionary time, it is still unclear why miRNAs have so many targets but repress them so weakly (Baek et al. 2008; Selbach et al. 2008). Finding answers to this question is crucial to understanding the biology of miRNAs in general.

If most miRNA-mediated repression offers no selective advantage, natural selection, or genetic drift should drive down target pool size. In particular, weakly repressed genes, considered nonfunctional by many, should be

subjected to removal from the network, for example by target-site mutation. We refer to this view as the “small (functional) pool” hypothesis favored by Pinzón *et al.* and others (Ecsedi *et al.* 2015; Pinzon *et al.* 2017). On the other hand, an alternative “large (functional) pool” hypothesis (Wu *et al.* 2009; Osella *et al.* 2011; Ebert and Sharp 2012; Pelaez and Carthew 2012; Posadas and Carthew 2014; Chen *et al.* 2017) proposes that weak targets are collectively beneficial.

While several studies have shown that target genes may evolve, one by one, to avoid repression (Chen and Rajewsky 2007; Roux *et al.* 2012; Barbash *et al.* 2014), others could not find evidence for target pool reduction (Meunier *et al.* 2013) and often observed the opposite trend (Shomron *et al.* 2009; Nozawa *et al.* 2016). Even though there may be a selective advantage in decreasing pool size, removing or adding a single target likely has negligible fitness consequences. The rejection of fitness neutrality has been shown to be extremely difficult as it become weaker (Wang *et al.* 2017; Wen *et al.* 2018). Therefore, a pool size that is stable over time and involves continued target turnover is compatible with both the “small pool” and “large pool” hypotheses.

In this context, “twin-miRNAs” (twin-miRs for short) may provide a useful test. Biogenesis of miRNAs is generally thought to involve the production of a single mature miRNA from a hairpin RNA precursor (Khvorova *et al.* 2003; Schwarz *et al.* 2003), here referred to as “solo-miR.” Not infrequently, however, both strands are used to produce mature miRNAs (fig. 1, see also [supplementary fig. S1A, Supplementary Material](#) online) (Ro *et al.* 2007; Okamura *et al.* 2008; Yang *et al.* 2011). The two products are designated as the major and minor miRs, depending on their relative abundance. Minor miRs roughly correspond to miR* in the older literature (see also Materials and Methods). Although minor miRs have sometimes been characterized as biogenic errors (Khvorova *et al.* 2003; Schwarz *et al.* 2003; Shen *et al.* 2011), they are able to repress target gene expression (Ro *et al.* 2007; Okamura *et al.* 2008; Kuchenbauer *et al.* 2011; Yang *et al.* 2011). They would thus substantially increase the number of miRNA targets and spread the total repression even further.

Since all targets of minor miRs are weakly repressed, their *en-masse* elimination by reduction from a twin- to solo-miR could lead to large overall fitness consequences, even if derepression of individual targets has little effect (Chen and Rajewsky 2007; Roux *et al.* 2012; Barbash *et al.* 2014). The “small pool” hypothesis would then predict that twin-miRs are evolutionarily transient, observable mainly among newly emerged miRNAs (Lu *et al.* 2008; Lyu *et al.* 2014). We therefore identify and study the evolution of twin-miRs in the *Drosophila* clade. We find that these miRNAs are evolutionarily stable and functional, arguing against the idea that miRNA pool sizes are under selection pressure to decrease.

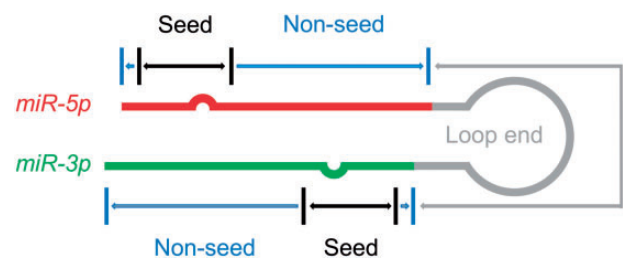


FIG. 1.—Precursor miRNA producing two mature products. The hairpin of a precursor miRNA has the potential to produce two mature products from both arms, which are denoted as miR-5p (in red) and miR-3p (in green). Both miR-5p and miR-3p can be separated into the seed (2–8nt at the 5'-end of the miRs) and the nonseed region (the remainder 14nt), which are indicated with black and blue lines, respectively.

Materials and Methods

Data Collection and Small RNA Sequencing

To identify twin-miR candidates, we collected deep sequencing data from 40 small RNA libraries representing all major *D. melanogaster* tissue types (head, body, ovary, and testis) from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>; last accessed April 30, 2018) ([supplementary table S1, Supplementary Material](#) online). Five additional small RNA libraries were sequenced for this study, including ovaries and testes of 3- to 5-day-old Canton-S and Z56 adults, and Canton-S carcasses. Total RNA was extracted using the TRIzol reagent (Invitrogen, Carlsbad, CA) and analyzed using an Agilent 2100 Bioanalyzer. Small RNA libraries were constructed using the Small RNA Sample Preparation kit (Illumina, San Diego, CA) and sequenced with a HiSeq 2000 (Illumina, San Diego, CA) at the Beijing Genomics Institute (Shenzhen, China). All sequence data generated for this study were deposited in GEO under the accession number GSE97364.

miRNA Expression Analyses

The fly genome (r6.04) was retrieved from FlyBase (<http://flybase.org/>) (Attrill *et al.* 2016). High-confidence miRNA precursor and mature sequences were retrieved from miRBase Release 21 (<http://www.mirbase.org/>; last accessed April 30, 2018) (Kozomara and Griffiths-Jones 2014). 149 miRNA precursors that have miRBase mature miRNA annotations produced from both arms (miR-5p and miR-3p) were used for expression analyses. The expression of each known mature miRNA was measured using the Mapper and Quantifier modules of miRDeep2 version 2.0.0.7 (Friedlander *et al.* 2012) with default parameters, normalized by the total reads matching all the miRNAs per library, and scaled as Reads Per Million (RPM). The expression ratio of miR-5p to the sum of miR-5p and miR-3p (in RPM units) was plotted as a heat map using Excel 2013. Principal component analysis (PCA) of the ratio for each miRNA was conducted in Python 2.7 using the sklearn package version 0.17.1.

Identification of Solo and Twin miRNAs

After removing weakly expressed miRNA precursors (average RPM < 10 across all libraries in every tissue) and duplicated miRNAs with identical mature sequences, the remaining 123 miRNAs were used to identify twin- and solo-miRs (workflow shown in [supplementary fig. S1B, Supplementary Material](#) online). We defined a miRNA as a twin-miR if the mature miRNA with lower within-tissue average RPM, referred to as “minor miR,” was expressed at >30% of the total read counts of the precursor in two libraries of a given tissue type. This criterion ensures the coexpression of two mature products.

The designation of the minor product used to be miR*, which has been largely abandoned. This may be because “star” in miR* appears to connote “noise” (as in “star activity” of restriction enzymes). However, the current notation of 5p or 3p is uninformative about their relative abundance. We therefore suggest major versus minor miRs, which informs about relative abundance without implying functionality. Except for a few miRNAs that produce the two forms in near parity, the major/minor designation should generally be unambiguous.

Comparative analyses of solo-miRs and twin-miRs in *Drosophila melanogaster*

Drosophila melanogaster miRNAs orthologous sequences in *D. pseudoobscura* and *D. persimilis*, as well as their ages were adopted from Mohammed *et al.* (Mohammed *et al.* 2013). To compare expression levels of solo-miRs and twin-miRs ([table 1](#)), the RPMs of two mature products produced from each miRNA precursor were added up per library and then averaged across all libraries in each tissue type, followed by log₂ transformation. To estimate conservation, precursor sequences in both *D. melanogaster* and either *D. pseudoobscura* or *D. persimilis* were aligned using MUSCLE with default parameters (Edgar 2004). Each miRNA precursor was divided into five parts: the seed (2–8 nt) and nonseed regions (the remaining 14 nt) from both major and minor miRs ([fig. 1](#) and [supplementary fig. S4, Supplementary Material](#) online), and the loop end region. For each of the five regions, the number of substitutions between *D. melanogaster* and either *D. pseudoobscura* or *D. persimilis* was counted, and then used to estimate the degree of sequence conservation as (total sites–substitutions)/total sites × 100%. The conservation difference between the seed and nonseed regions was calculated for twin-miRs ($n = 23, 24$) and solo-miRs ($n = 59, 58$) in both *D. pseudoobscura* and *D. persimilis*. The significance of any differences was determined by using a bootstrap analysis, in which each region of twin-miRs or solo-miRs was resampled with replacement 1,000 times, followed by conservation calculation.

Table 1

Summary of Twin-miRs across *Drosophila melanogaster* Tissues

miRNA	Head	Body	Ovary	Testis
<i>dme-mir-10</i>	√	√	√	√
<i>dme-mir-1012</i>	√	√	√	√
<i>dme-mir-313</i>	√	√	√	√
<i>dme-mir-92a</i>	√	√	√	√
<i>dme-mir-959</i>	√	√	√	√
<i>dme-mir-988</i>	√	√	√	√
<i>dme-mir-954</i>	√	√	√	×
<i>dme-mir-974</i>	√	√	×	√
<i>dme-mir-304</i>	√	×	√	√
<i>dme-mir-305</i>	√	×	√	√
<i>dme-mir-4968</i>	√	×	√	√
<i>dme-mir-960</i>	√	×	√	√
<i>dme-mir-965</i>	√	×	√	√
<i>dme-mir-984</i>	√	×	√	√
<i>dme-mir-306</i>	√	√	×	×
<i>dme-mir-5</i>	√	√	×	×
<i>dme-mir-927</i>	√	√	×	×
<i>dme-mir-962</i>	√	√	×	×
<i>dme-mir-998</i>	√	√	×	×
<i>dme-mir-308</i>	×	√	√	×
<i>dme-mir-100</i>	×	×	√	√
<i>dme-mir-1003</i>	×	×	√	√
<i>dme-mir-1006</i>	×	×	√	√
<i>dme-mir-1008</i>	×	×	√	√
<i>dme-mir-125</i>	×	×	√	√
<i>dme-mir-190</i>	×	×	√	√
<i>dme-mir-2279</i>	×	×	√	√
<i>dme-mir-277</i>	×	×	√	√
<i>dme-mir-283</i>	×	×	√	√
<i>dme-mir-8</i>	×	×	√	√
<i>dme-mir-958</i>	×	×	√	√

√, miRNAs identified as twin-miRs in a given tissue type; ×, miRNAs not defined as twin-miRs in the tissue.

Repression Effects of Both miRs from *dme-mir-313*

The *dme-mir-310/311/312/313* cluster (denoted as *Dm310s*) knockout and control flies were constructed previously (Tang *et al.* 2010). Among the cluster members, *dme-mir-313* is a twin-miR identified in multiple tissues (see [fig. 2](#) and [table 1](#)). We evaluated repression effects of both miRs produced by *dme-mir-313* using RNA-seq analysis of the *Dm310s* knockout and control flies (accession numbers listed in [supplementary table S2, Supplementary Material](#) online). Total RNA was extracted from testes using the Trizol reagent (Invitrogen, Carlsbad, CA). RNA-seq was performed using the Illumina GA-II platform at the Beijing Genomics Institute (Shenzhen, China). Raw data from RNA-seq were mapped to the *D. melanogaster* genome (r6.04), and normalized using Tophat (Trapnell *et al.* 2009) and SeqMonk version 1.36.0 (<http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/>; last accessed April 30, 2018). The expression level of each gene was calculated in Reads Per



Fig. 2.—miRNA arm usage in *Drosophila melanogaster*. The expression levels of miR-5p and miR-3p from each precursor miRNA was measured as Reads Per Million (RPM) in 45 small RNA libraries from heads, bodies, ovaries, and testes of *D. melanogaster*. The heat map shows relative expression ratios of miR-5p to the sum of miR-5p and miR-3p (in RPM units). miRNAs that were expressed almost exclusively from either the 5p or 3p arm of the precursor are in red (5p dominant) or green (3p dominant), and those that were undetectable are in gray. miRNA names, miRNA ages, and tissue types surveyed are labeled on the left, right, and bottom of the heat map, respectively. Twin-miR names are shaded in yellow.

Kilobase per Million (RPKM) as previously described (Mortazavi et al. 2008). 3'-UTR sequences of *D. melanogaster* were retrieved from FlyBase (r6.04) (Attrill et al. 2016) and miR-313 target genes were predicted using the TargetScan algorithm (Ruby et al. 2007). Note that the seed region of *dme-miR-313-3p* is identical with other members in *Dm310s*, and *dme-miR-313-5p* has a distinct seed region (supplementary fig. S5B, Supplementary Material online). To examine the repression effects of *dme-miR-313-5p* and *dme-miR-313-3p*, all predicted targets with 8mer, 7mer-m8, and 7mer-1A sites were used. Genes with $\log_2(\text{RPKM}) > 1$ were used to calculate expression fold changes between the *Dm310s* knockout and control flies, and cumulative distributions were plotted in R version 3.3.2 (R Core Team 2016).

Comparative Analyses of Solo- and Twin-miR Target Number

The solo-miRs, as well as minor and major products of twin-miRs, were used for target prediction by TargetScan (Ruby et al. 2007) and microT-CDS (Paraskevopoulou et al. 2013) with default settings. For TargetScan prediction, multiple alignments of 3'-UTRs in *Drosophila* species were retrieved from TargetScanFly (Release 6.2, http://www.targetscan.org/cgi-bin/targetscan/data_download.cgi?db=fly_12; last accessed April 30, 2018). Only 8mer, 7mer-m8, and 7mer-1A sites that are conserved between *D. melanogaster* and *D. pseudoobscura* were used in further comparisons of miRNA target numbers. For microT-CDS (Reczko et al. 2012; Paraskevopoulou et al. 2013) prediction, all targets were retrieved from the official website with a threshold of 0.7 (default) for target identification. The target numbers of major and minor miRs were counted separately. Overlapping targets were counted only once when calculating the total target number of both products of twin-miRs. Mann-Whitney *U* test, as implemented in R version 3.3.2 (R Core Team 2016), was used for comparisons of target numbers between twin-miRs and solo-miRs.

Results

To identify miRNAs that produce twin products, we collected deep sequencing data from 40 small RNA libraries and supplemented this set with five additional libraries that were sequenced for this study. These 45 *D. melanogaster* libraries together represent a systematic coverage of the major tissue types (head, body, and sexual organs; see supplementary table S1, Supplementary Material online). We defined twin-miRs as miRNAs that express their minor miR at >30% of the total read count of both mature products in at least two samples from a given tissue (see Materials and Methods). The criterion ensures the coexpression of major and minor miRs under at least some conditions. We classified

miRNAs by their age. Old miRNAs are found in both the *Sophophora* and *Drosophila* subgenus, medium aged ones are present only in *Sophophora*, and young miRNAs have emerged only after the split of the *melanogaster* subgroup. We wish to know whether twin-miR production is evolutionarily young and transient.

Survey of twin-miRs in *Drosophila melanogaster*

Of the 300 million reads we collected from the 45 small RNA libraries, ~100 million were mapped to miRNA precursors (Materials and Methods). As shown in figure 2, nearly 40% (49/123) of the expressed miRNAs are identified as twin-miRs (names highlighted in yellow), including 25 expressed in heads, 18 in bodies, 25 in ovaries, and 32 in testes. The ratio of miRNA-5p reads (in unit of RPM) to the total miR reads for each miRNA is also shown in figure 2. When this ratio is close to 1 or 0, miRs are expressed almost exclusively from either the 5p or 3p arm of the precursor. A ratio near 0.5 indicates coexpression of the two arms. A principal component analysis of the ratio for each miRNA from the 45 libraries separates the 123 miRNAs into two clusters that represent solo-miRs, indicating that arm usage bias is an intrinsic feature that varies among miRNAs (supplementary fig. S2, Supplementary Material online). Other miRs fall outside the clusters and represent twin-miRs. A more comprehensive compilation is shown in supplementary figure S3, Supplementary Material online.

The proportion of twin-miRs among young miRNAs (10/22, or 45%) is slightly higher than that among old miRNAs (36/97 = 37%), but this difference is not statistically significant (two-sided Fisher's exact test, $P = 0.48$). Birth-death dynamics of miRNAs (Lu et al. 2008; Lyu et al. 2014) resulted in too few medium-aged miRs to analyze (four), but it is notable that three of them are twin-miRs. Given that 37% of the old miRNAs remain twin-miRs after >60 Myr of evolution, the production of twin-miRs should not be considered evolutionarily transient. Twin-miR production appears to be a normal facet of miRNA biogenesis.

We further ask whether minor miRs might be highly tissue-specific. Over 60% (31/49) of twin-miRs are identified as twin-miRs in at least two tissue types (table 1). About 6 of the 31 miRNAs (*dme-mir-10*, *dme-mir-1012*, *dme-mir-313*, *dme-mir-92a*, *dme-mir-959*, *dme-mir-988*) produce twin-miRs in every tissue examined. Thus, twin-miR production is not restricted to specific tissues. Since these 31 miRNAs are more prevalent in twin miRs production, we then further compared their expression, evolutionary conservation and target number with solo-miRs.

Expression of Twin-miRs

The first question is whether twin-miRs are expressed at a comparable level to solo-miRs. If minor miRs are

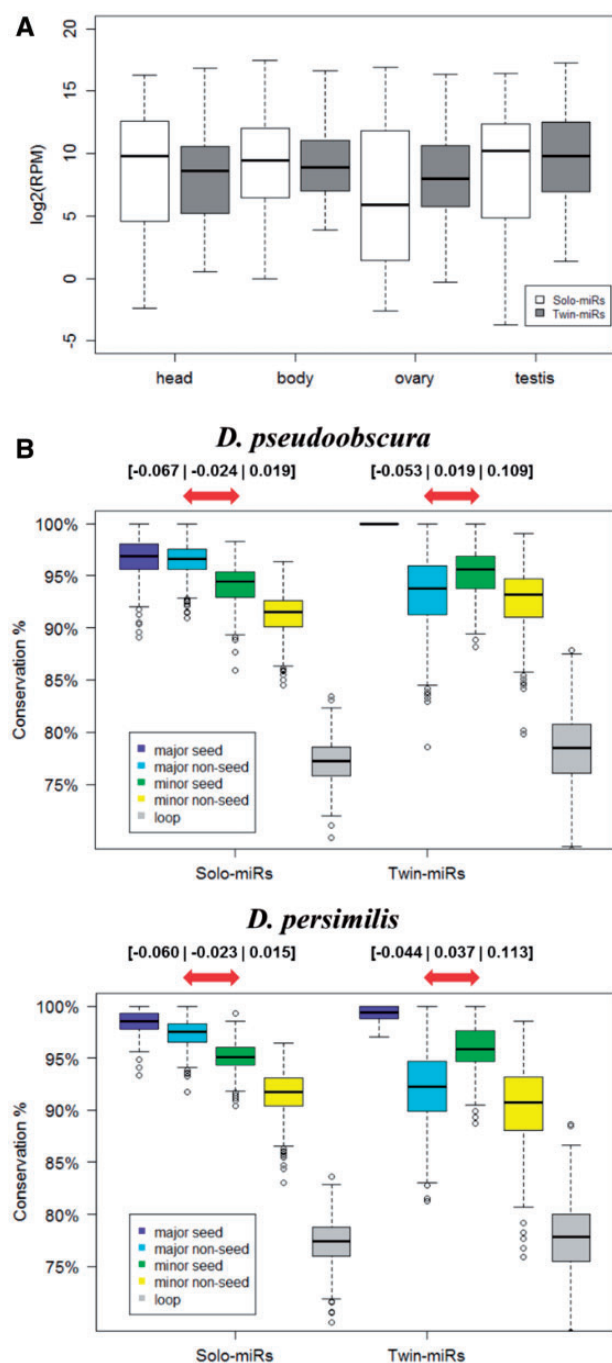


FIG. 3.—Comparisons of expression and conservation between solo- and twin-miRs. (A) The expression levels of solo- and twin-miRs in heads, bodies, ovaries, and testes of *Drosophila melanogaster*. The boxplot shows log₂(RPM) of 74 solo-miRs and 31 twin-miRs that were identified in at least two tissue types (table 1). In each tissue type, the expression level of a solo- or twin-miR was measured as the sum of the within-tissue average RPMs of two mature miRs produced from the same miRNA precursor. The expression difference between solo- and twin-miRs was insignificant in all comparisons (Mann–Whitney *U* test, all $P > 0.05$). (B) Sequence conservation of solo- and twin-miRs between *D. melanogaster* and either *D. pseudoobscura* or *D. persimilis*. The boxplot shows the degree of conservation

processing errors and deleterious, we would expect that only weakly expressed miRNAs would produce minor miRs without a strong selective pressure (Shen et al. 2011). However, there is no statistically significant difference in expression level between solo-miRNAs and twin-miRs (fig. 3A, Mann–Whitney *U* test, all $P > 0.05$ for solo-miRs and twin-miRs in table 1).

Evolutionary Conservation of Twin-miRs

The functionality of minor miRs can also be inferred from their evolutionary conservation. We divided each miRNA precursor into five parts: seed (2–8 nt) and nonseed region (the remaining 14 nt) from both the major and minor miRs, as well as the loop end (fig. 1 and supplementary fig. S1, Supplementary Material online). We concatenated each region to form five pseudomolecules. We estimated the number of substitutions in these five segments between *D. melanogaster* and either *D. pseudoobscura* or *D. persimilis* (see Materials and Methods). The percentage of site identity is the conservation score depicted in figure 3B.

As expected, the conservation of solo-miRs is higher for the major miRs than for the minor miRs, and the loop end is the least conserved region (fig. 3B). Within major miRs, the seed region is more conserved than the nonseed region (first two columns). Interestingly, the order is reversed in twin-miRs between the seed regions of minor miRs and the nonseed regions of major miRs, as indicated by double-headed red arrows in figure 3B. To assess the uncertainty of the estimates, we performed 1,000 bootstrap resamples of the sites and calculated the conservation difference between the minor seed and major nonseed regions. For *D. melanogaster* and *D. pseudoobscura*, the bootstrap confidence interval for solo miRs is [−0.067 | −0.024 | 0.019]. For twin-miRs it is [−0.053 | 0.019 | 0.109], where the numbers denote the [2.5% | 50% | 97.5%] percentiles of the bootstrap distribution. The estimation for *D. melanogaster* and *D. persimilis* is largely similar (fig. 3B). In summary, the two seed regions on the opposite arms of the twin-miRs are the most conserved regions in these loci, in accordance with the existence of two mature miR products from the same hairpin.

FIG. 3.—Continued

for twin-miRs ($n=23, 24$) and solo-miRs ($n=59, 58$) in *D. pseudoobscura* and *D. persimilis*, respectively. The degree of conservation was calculated as (total sites–substitutions)/total sites \times 100% for each of the five parts: seed and nonseed of the major miR, seed, and nonseed of the minor miR, and the loop end region. The conservation difference between the seed region of minor miR and nonseed region of major miR was determined by a bootstrap analysis with 1,000 replicates. The 2.5th, 50th, and 97.5th percentile values are shown on the top.

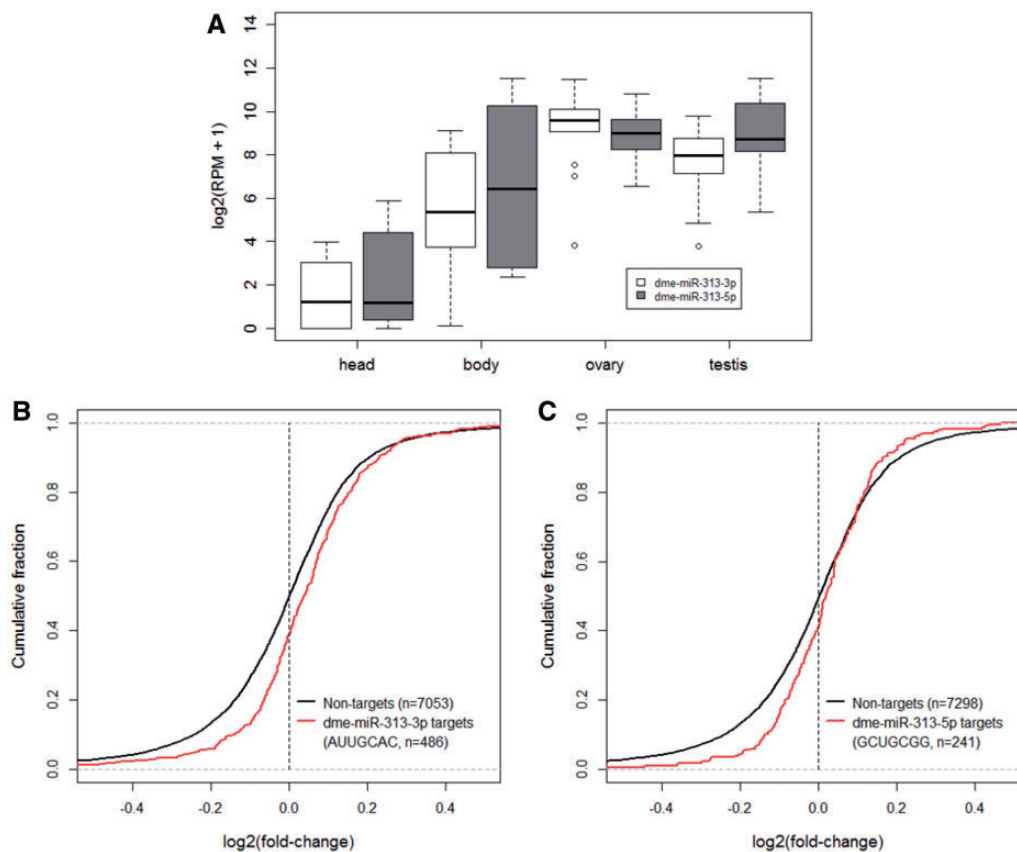


FIG. 4.—Repression effects of *dme-miR-313-3p* and *dme-miR-313-5p* in *Drosophila melanogaster*. (A) The expression levels of *dme-miR-313-3p* and *dme-miR-313-5p* are comparable among tissues. RPM, Reads Per Million. (Mann–Whitney U test, all $P > 0.05$). (B) Repression effect of *dme-miR-313-3p* (Kolmogorov–Smirnov test, $P = 1.375 \times 10^{-8}$). (C) Repression effect of *dme-miR-313-5p* (Kolmogorov–Smirnov test, $P = 0.00427$). In (B) and (C), miRNA targets are represented by red lines while nontargets by black lines. Gene numbers and seed sequences are included in parentheses.

Measurable Target Repression by a Twin-miR

Does the production of twin miRs have any effect on their targets? Previous studies have shown that both mature miRNAs from a single precursor are able to repress target gene expression (Ro et al. 2007; Okamura et al. 2008; Kuchenbauer et al. 2011; Yang et al. 2011). For example, both 5p and 3p mature miRNAs from *dme-mir-iab-4* and *dme-mir-1010* (also identified as twin-miRs here, see fig. 2) show repressive activity in reporter assays (Okamura et al. 2008). Rather than look at individual targets, we set out to test if a minor product of a twin-miR has a measurable transcriptome-wide effect.

Dme-mir-313 is a member of the miR-310s miRNA cluster in *D. melanogaster* and is a twin-miR (see fig. 2 and table 1). It expresses both arms at comparable levels in all tissue types we assayed (fig. 4A, Mann–Whitney U test, all $P > 0.05$). To detect the repression effects of miR-313 on its targets, we dissected testes from adult flies that lacked this miRNA due to a knock-out mutation (Tang et al. 2010) and performed an RNA-seq analysis (Materials and Methods). As expected, targets of both arms are significantly up-regulated in the knock-out flies (two-sample Kolmogorov–Smirnov test,

$P = 1.375 \times 10^{-8}$ for *dme-miR-313-3p* and $P = 0.00427$ for *dme-miR-313-5p*, respectively, fig. 4B and C), although the effects are not large (supplementary fig. S5A, Supplementary Material online).

These results suggest that the twin miRs of miR-313 can both repress their targets at the transcriptome level (fig. 4B and C). Collectively, previous studies (Ro et al. 2007; Okamura et al. 2008; Yang et al. 2011) and our analyses support the interpretation that twin products from a single precursor are functional and have regulatory effects on their targets.

Solo- versus Twin-miR Target Gene Number

We wanted to know how the number of targets of twin-miRs compares to the number of targets regulated by solo-miRs. We were also interested in how the targets are partitioned between the minor and major seeds. We used TargetScan (Ruby et al. 2007) to identify target loci of solo-miRNAs (A), the major products of twin-miRNAs (B), and the minor twin-miR products (B*). As a control, we used the potential targets of the minor products of solo-miRs (A*). To increase the confidence of target prediction, we focused on targets that were

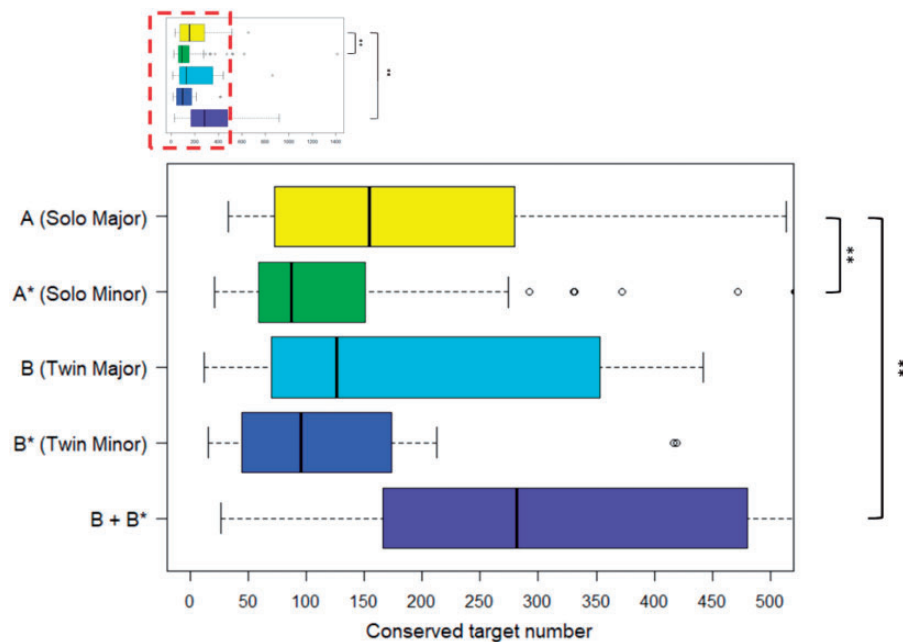


Fig. 5.—Comparisons of predicted solo- and twin-miR target numbers in *Drosophila melanogaster*. Major and minor miRs of 61 solo-miRs and 24 twin-miRs in both the *Sophophora* and *Drosophila* subgenus were used for target prediction using TargetScan. The boxplot shows the number of target genes conserved between *D. melanogaster* and *D. pseudoobscura* for the major arms of solo-miRs (A), the minor arms of solo-miRs (A*), the major arms of twin-miRs (B), and the minor arms of twin-miRs (B*); see Materials and Methods). The zoom-in layout is shown to allow a better comparison, with an inset on the top for the global distribution. Statistical significance was determined by two-sided Mann–Whitney *U* test. **, $P < 0.01$.

conserved between *D. melanogaster* and *D. pseudoobscura* (fig. 5; see Materials and Methods). As expected, the major miRs of solo-miRs target more genes than the minor miRs (fig. 5). The median number of targets for A is 154, while it is 87 for A* (two-sided Mann–Whitney *U* test, $P = 0.008$). The same pattern is evident for twin-miRs. Although the B/B^* ratio is smaller than A/A^* (1.33 vs. 1.77), the difference is not statistically significant (1,000-fold bootstrap from the values for both majors and minors, $P > 0.05$). However, since all the evidence presented above suggests that minor products of twin-miRs are active and repress their targets, the total target pool of twin miRs ($B + B^* = 281$ [332 ± 207]) is substantially larger than the target pool of solo-miRs ($A = 154$ [192 ± 141]; two-sided Mann–Whitney *U* test, $P = 0.001$). Similar trends are observed if we use a different target identification algorithm (supplementary fig. S6, Supplementary Material online). Thus, miRNAs that produce twin products target over 80% more loci than solo-miRs and the additional targets are all weakly repressed.

Discussion

It has been difficult to evaluate biological significance of the many weakly repressed miRNA targets. Since individual effects are small, experimental approaches focused on phenotypic consequences of each transcript are hard to implement. Therefore, in the absence of data to the contrary, it was reasonable to argue that only strongly repressed targets are

functionally significant and thus the real regulation pool is small (Ecsedi et al. 2015; Pinzon et al. 2017). However, the results of the present study focusing on twin-miRs provide direct evidence that at least a significant subset of the weakly regulated transcripts is functionally important and selectively nonneutral. The phenomenon can be found in both *Drosophila* and vertebrate species (Yang et al. 2011). We show that twin-miRs can be maintained over millions of years (fig. 2), produce two evolutionarily conserved seeds (fig. 3B) that are capable of affecting target repression (fig. 4B and C, see also supplementary text), and are not uncommon. There is no measurable evolutionary trend toward elimination of minor arms, as would be the case if they were purely responsible for (potentially deleterious) regulatory noise. Furthermore, minor arms of twin-miRs act exclusively through small-effect repression, suggesting that such repression does indeed carry functional importance.

If weak repression of many transcripts is an important regulatory feature of miRNAs, it is hard to imagine that any individual interaction is subject to purifying selection. Instead, it seems more likely that the effects are important to overall regulatory network structure. Selection for network stability may be the driving force giving rise to the distributed regulation mediated by miRNAs. This idea is supported by studies of RNA cross-talk networks (Salmena et al. 2011; Tay et al. 2014; Thomson and Dinger 2016). Furthermore, applying the May–Wigner theory of food webs and ecosystems (May 1972; Allesina and

Tang 2012) to transcriptional regulatory networks (Chen et al. 2017) indicates that network stability is enhanced by many weak repression events. Our results suggest that experiments perturbing overall regulatory network architecture are more likely to accurately reflect the fraction of functional miRNA regulatory events than approaches focused on individual transcripts.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by the National Science Foundation of China (31730046, 31770246, 91231117), the 985 Project (33000-18831107), the National key Research and Development Program of China (2017YFC0506100-1), the Science and Technology Program of Guangzhou (201707020035) and the fundamental Research Funds for the central Universities (16lgjc75).

Literature Cited

- Allesina S, Tang S. 2012. Stability criteria for complex ecosystems. *Nature* 483(7388):205–208.
- Attrill H, et al. 2016. FlyBase: establishing a Gene Group resource for *Drosophila melanogaster*. *Nucleic Acids Res.* 44(D1):D786–D792.
- Baek D, et al. 2008. The impact of microRNAs on protein output. *Nature* 455(7209):64–71.
- Barbash S, Shifman S, Soreq H. 2014. Global coevolution of human microRNAs and their target genes. *Mol Biol Evol.* 31(5):1237–1247.
- Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116(2):281–297.
- Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* 136(2):215–233.
- Bushati N, Cohen SM. 2007. MicroRNA functions. *Annu Rev Cell Dev Biol.* 23:175–205.
- Chen K, Rajewsky N. 2007. The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet.* 8(2):93–103.
- Chen Y, Allesina S, Shen Y, Wu C-I. 2017. From foodwebs to gene regulatory networks (GRNs) – weak repressions by microRNAs confer system stability. *bioRxiv*. <https://doi.org/10.1101/176701>.
- Ebert MS, Sharp PA. 2012. Roles for microRNAs in conferring robustness to biological processes. *Cell* 149(3):515–524.
- Ecsedi M, Rausch M, Grosshans H. 2015. The let-7 microRNA directs vulval development through a single target. *Dev Cell* 32(3):335–344.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792–1797.
- Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N. 2012. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40(1):37–52.
- Haijun Wen, Hurng-i Wang, Xionglei He, Chung-I Wu. 2018. On the low reproducibility of cancer studies. *Natl Sci Rev.* doi: 10.1093/nsr/nwy021.
- Hausser J, Zavolan M. 2014. Identification and consequences of miRNA-target interactions—beyond repression of gene expression. *Nat Rev Genet.* 15(9):599–612.
- Karres JS, Hilgers V, Carrera I, Treisman J, Cohen SM. 2007. The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in *Drosophila*. *Cell* 131(1):136–145.
- Khvorova A, Reynolds A, Jayasena SD. 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115(2):209–216.
- Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42(Database issue):D68–D73.
- Kuchenbauer F, et al. 2011. Comprehensive analysis of mammalian miRNA* species and their role in myeloid cells. *Blood* 118(12):3350–3358.
- Lu J, et al. 2008. The birth and death of microRNA genes in *Drosophila*. *Nat Genet.* 40(3):351–355.
- Lyu Y, et al. 2014. New microRNAs in *Drosophila*—birth, death and cycles of adaptive evolution. *PLoS Genet.* 10(1):e1004096.
- May RM. 1972. Will a large complex system be stable? *Nature* 238(5364):413–414.
- Meunier J, et al. 2013. Birth and expression evolution of mammalian microRNA genes. *Genome Res.* 23(1):34–45.
- Mohammed J, Flynt AS, Siepel A, Lai EC. 2013. The impact of age, biogenesis, and genomic clustering on *Drosophila* microRNA evolution. *RNA* 19(9):1295–1308.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621–628.
- Nozawa M, et al. 2016. Evolutionary transitions of MicroRNA-target pairs. *Genome Biol Evol.* 8(5):1621–1633.
- Okamura K, et al. 2008. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat Struct Mol Biol.* 15(4):354–363.
- Osella M, Bosia C, Cora D, Caselle M. 2011. The role of incoherent microRNA-mediated feedforward loops in noise buffering. *PLoS Comput Biol.* 7(3):e1001101.
- Paraskevopoulou MD, et al. 2013. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Research* 41:W169–173.
- Pelaez N, Carthew RW. 2012. Biological robustness and the role of microRNAs: a network perspective. *Curr Top Dev Biol.* 99:237–255.
- Pinzon N, et al. 2017. microRNA target prediction programs predict many false positives. *Genome Res* 27:234–245.
- Posadas DM, Carthew RW. 2014. MicroRNAs and their roles in developmental canalization. *Curr Opin Genet Dev.* 27:1–6.
- Reczko M, Maragkakis M, Alexiou P, Grosse I, Hatzigeorgiou AG. 2012. Functional microRNA targets in protein coding sequences. *Bioinformatics* 28:771–776.
- Ro S, Park C, Young D, Sanders KM, Yan W. 2007. Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res.* 35(17):5944–5953.
- Roux J, Gonzalez-Porta M, Robinson-Rechavi M. 2012. Comparative analysis of human and mouse expression data illuminates tissue-specific evolutionary patterns of miRNAs. *Nucleic Acids Res.* 40(13):5890–5900.
- Ruby JG, et al. 2007. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res.* 17(12):1850–1864.
- Salmela L, Poliseno L, Tay Y, Kats L, Pandolfi PP. 2011. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146(3):353–358.
- Schwarz DS, et al. 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115(2):199–208.
- Selbach M, et al. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature* 455(7209):58–63.
- Shen Y, et al. 2011. Testing hypotheses on the rate of molecular evolution in relation to gene expression using microRNAs. *Proc Natl Acad Sci U S A.* 108(38):15942–15947.

- Shomron N, Golan D, Hornstein E. 2009. An evolutionary perspective of animal microRNAs and their targets. *J Biomed Biotechnol.* 2009:594738.
- Tang T, et al. 2010. Adverse interactions between micro-RNAs and target genes from different species. *Proc Natl Acad Sci U S A.* 107(29):12935–12940.
- Tay Y, Rinn J, Pandolfi PP. 2014. The multilayered complexity of ceRNA crosstalk and competition. *Nature* 505(7483):344–352.
- Team RC, 2016. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
- Thomson DW, Dinger ME. 2016. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet.* 17(5):272–283.
- Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
- Varghese J, Lim SF, Cohen SM. 2010. Drosophila miR-14 regulates insulin production and metabolism through its target, sugarbabe. *Genes Dev.* 24(24):2748–2753.
- Wang H-Y, et al. 2017. Is the evolution in tumors Darwinian or non-Darwinian? *Natl Sci Rev.* 5:15–17.
- Wu CI, Shen Y, Tang T. 2009. Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome Res.* 19(5):734–743.
- Yang JS, et al. 2011. Widespread regulatory activity of vertebrate microRNA* species. *RNA* 17(2):312–326.

Associate editor: Bill Martin