

# Evolutionary Transitions of MicroRNA-Target Pairs

Masafumi Nozawa<sup>1,2,\*</sup>, Mai Fujimi<sup>1</sup>, Chie Iwamoto<sup>1</sup>, Kanako Onizuka<sup>1</sup>, Nana Fukuda<sup>1</sup>, Kazuho Ikeo<sup>1,2</sup>, and Takashi Gojobori<sup>1,3</sup>

<sup>1</sup>Center for Information Biology, National Institute of Genetics, Shizuoka, Japan

<sup>2</sup>Department of Genetics, SOKENDAI, Shizuoka, Japan

<sup>3</sup>King Abdullah University of Science and Technology, Computational Bioscience Research Center, Biological and Environmental Science and Engineering, Thuwal, Kingdom of Saudi Arabia

Data deposition: This project has been deposited at DDBJ Sequence Read Archive (DRA) under the accession DRA001729 and DRA002241

\*Corresponding author: E-mail: manozawa@tmu.ac.jp.

Accepted: April 16, 2016

## Abstract

How newly generated microRNA (miRNA) genes are integrated into gene regulatory networks during evolution is fundamental in understanding the molecular and evolutionary bases of robustness and plasticity in gene regulation. A recent model proposed that after the birth of a miRNA, the miRNA is generally integrated into the network by decreasing the number of target genes during evolution. However, this decreasing model remains to be carefully examined by considering *in vivo* conditions. In this study, we therefore compared the number of target genes among miRNAs with different ages, combining experiments with bioinformatics predictions. First, we focused on three *Drosophila* miRNAs with different ages. As a result, we found that an older miRNA has a greater number of target genes than a younger miRNA, suggesting the increasing number of targets for each miRNA during evolution (increasing model). To further confirm our results, we also predicted all target genes for all miRNAs in *D. melanogaster*, considering co-expression of miRNAs and mRNAs *in vivo*. The results obtained also do not support the decreasing model but are reasonably consistent with the increasing model of miRNA-target pairs. Furthermore, our large-scale analyses of currently available experimental data of miRNA-target pairs also showed a weak but the same trend in humans. These results indicate that the current decreasing model of miRNA-target pairs should be reconsidered and the increasing model may be more appropriate to explain the evolutionary transitions of miRNA-target pairs in many organisms.

**Key words:** bioinformatics prediction, gene regulation, RNA-seq, small RNA, transgenic fly.

## Introduction

MicroRNAs (miRNAs) provide an important layer of gene regulation by operating at the post-transcriptional stage (Ambros 2004; Bartel 2004, 2009). Although miRNAs were only discovered ~20 years ago (Lee et al. 1993), advancements in sequencing technologies enabled their extensive identification in various organisms (Kozomara and Griffiths-Jones 2014). These efforts revealed that each of a number of eukaryotic genomes contains hundreds or thousands of miRNA loci (Berezikov et al. 2006; Axtell et al. 2007; Axtell and Bowman 2008; Meunier et al. 2013). For animals, many miRNA genes originated from hairpin structures within introns or intergenic regions (Lu et al. 2008; Nozawa et al. 2010), whereas for plants a substantial number of miRNA genes appear to have been generated from the duplication of

pre-existing miRNA genes, inverted duplicates of protein-coding genes, or transposable elements (Allen et al. 2004; Fahlgren et al. 2007; Piriyaopongsa and Jordan 2008; Nozawa et al. 2012).

In spite of these observations, little is currently known about the way in which miRNAs are integrated into gene regulatory networks. When the evolution of miRNA-involved networks is considered, it is essential to clarify the number of connections and the degree of connective strength between miRNAs and other network components such as the miRNA target genes. Old miRNAs, which emerged at an earlier stage of evolution, are known to be expressed at higher levels than young miRNAs, which emerged at a later stage (Berezikov et al. 2006; Lu et al. 2008; Ma et al. 2010; Meunier et al. 2013). It therefore seems that old miRNAs are able to suppress

their targets more efficiently than young miRNAs so that a connection between a miRNA and its target gene becomes stronger over time. However, how the number of connections between a miRNA and its targets has changed over evolutionary time remains controversial. In the decreasing model proposed by Chen and Rajewsky (2007), they claimed that although young miRNAs have a large number of target genes at the time of their emergence, most of these miRNA-target pairs are eliminated by long-term natural selection, because these pairs are deleterious. This process results in reduction in the number of targets when miRNAs become older (see also Roux et al. 2012). However, Tang et al. (2010) contended that the miRNAs would be quite difficult to be fixed into a population if newly emerged miRNAs have many targets with deleterious effects. Chen and Rajewsky (2007) also argued that because young miRNAs are normally transcribed very weakly only in a specific tissue (but see also Roux et al. 2012 for an opposite conclusion), their effects on fitness could be negligible and the number of connections between young miRNAs and targets can still be large. Yet, if the number of miRNA molecules produced is so small in a cell at first, the young miRNAs binding to their potential targets may be very few in reality. Indeed, Chen and Rajewsky (2007) did not show any clear-cut evidence to directly support their decreasing model.

To tackle this question, several studies have used bioinformatics predictions and compared the number of target genes for evolutionarily old and young miRNAs (Shomron et al. 2009; Roux et al. 2012; Meunier et al. 2013; Barbash et al. 2014). However, any consistent conclusions have not been obtained so far possibly due to differences in the prediction tools and the data sets used. *In silico* predictions are useful methods of identifying the potential targets of miRNAs, but there are strong limitations: (1) as different prediction tools use different factors in predicting the targets (Bartel 2009), the results obtained are often inconsistent; (2) some factors that are not currently considered for target prediction might be essential for the formation of authentic pairs *in vivo*, resulting in the generation of false positives as well as false negatives; (3) even if a mRNA is predicted as a target of a specific miRNA *in silico*, they cannot form an authentic pair when they are not coexpressed *in vivo*; and (4) it is almost impossible for *in silico* methods to assess the effects of miRNAs on downstream indirect targets (or network components) in the regulatory networks. For these reasons, it is fundamental to combine experimental approaches with bioinformatics predictions to compensate for deficiencies of each method and reliably investigate the evolutionary transitions in the number of target genes for miRNAs.

In this context, *Drosophila* is a useful model system for studying the evolution of miRNA-target pairs, because the origin of each miRNA has already been clarified (Lu et al. 2008; Berezikov et al. 2010; Nozawa et al. 2010). In addition, dozens of genome sequences from various species with

different genetic divergence are already available (Adams et al. 2000; Richards et al. 2005; Clark et al. 2007; Vicoso and Bachtrog 2013), and a number of tools for genome editing or transgenic experiments have been established (Duffy 2002; Gratz et al. 2013). Here, we first combined *in vivo* experiments with several *in silico* predictions to investigate the target gene repertoires of three *Drosophila* miRNAs (miR277, miR982, and miR954) with different ages. We then further examined the relationship between the age and the number of target genes of all miRNAs in *D. melanogaster* using several bioinformatics predictions with the consideration of an *in vivo* situation by testing the co-expression of miRNAs and mRNAs. We finally analyzed the large-scale human data in which target genes of miRNAs were experimentally identified. We here report that all results do not support the decreasing model of miRNA-target pairs but are reasonably well consistent with the increasing model in which each miRNA is integrated into the gene regulatory network by increasing the number of target genes during evolution.

## Materials and Methods

### Constructs for Overexpression of miRNAs

The *pUAST-DsRed2-miRNA* constructs for miR982 and miR954 were obtained from the *Drosophila* RNAi Screening Center (<http://www.flyrnai.org/>, last accessed May 3, 2016). The sequences of the miRNAs and their flanking regions in the constructs were confirmed by sequencing using an ABI3730xl sequencer (Life Technologies, Carlsbad, CA) with the following primers: forward, AGCGCAGCTGAACAAGCTA, and reverse, TGCCAATTATGTACACCACA.

### Flies

The transgenic *D. melanogaster* flies with the *pUAST-DsRed2-miRNA* constructs for miR982 and miR954 were generated in the  $w^{1118}$  background. The transgenic lines with balancer chromosomes were generated by BestGene Inc. (Chino Hills, CA). The transgenic line containing the miR277 construct (*UAS-DsRed2-miR277*) was obtained from the Cohen group (Szuplewski et al. 2012). To generate flies overexpressing miR277, miR982, or miR954, we crossed virgin females of each of the lines with *Tub-Gal4/TM6C,Sb,Tb* males. Overexpression of each miRNA in the offspring was confirmed by detection of DsRed2 under a fluorescence microscope as well as qPCR. As a control,  $w^{1118}$  females were crossed with the *Tub-Gal4/TM6C,Sb,Tb* males.

### RNA Extraction

Total RNA was extracted from five female or male control flies ( $w^{1118}/Tub-GAL4$ ) or flies overexpressing one of the miRNAs (*UAS-DsRed2-miRNA/Tub-GAL4*) using a standard acid phenol-guanidinium thiocyanate-chloroform extraction method (Sambrook and Russell 2001) with slight modifications.

Samples were collected from third instar larvae (before wandering), pupae (48–60 h after pupation), and adults (72–96 h after eclosion). The total RNA was treated with DNase I (TaKaRa, Ohtsu, Japan) to digest genomic DNA, and then mRNA was purified from the samples using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, Ipswich, MA). To determine the expression levels of miRNAs in other species, small RNAs were extracted from five female or male *D. simulans* and *D. yakuba* using the miRNeasy Mini Kit (Qiagen, Hilden, Germany). Samples were collected from third instar larvae (before wandering), pupae (48–60 h after pupation), and adults (72–96 h after eclosion).

### RNA-seq

The cDNA libraries were generated from the mRNA samples using the NEBNext mRNA Library Prep Master Mix Set for Illumina (NEB). Single-end sequencing of 100 bp was performed by TaKaRa with a HiSeq 2000 sequencer (Illumina, San Diego, CA). The sequencing data were processed following the method described by Nozawa et al. (2014). To account for variance within a condition, two biological replicates were performed for each condition, starting from different individuals. [Supplementary table S1, Supplementary Material](#) online, provides detailed information for each mRNA sequencing (mRNA-seq) run. The Ion Total RNA-seq Kit v2 (Life Technologies) was used to generate small RNA libraries, and the Ion PGM system (Life Technologies) was used for sequencing with Ion 318 Chip v2 ([supplementary table S2, Supplementary Material](#) online).

### Identification of Target Genes for Each miRNA

To examine the expression levels of all transcripts, all reads obtained from a sample ([supplementary table S1, Supplementary Material](#) online) were mapped to the transcriptome of *D. melanogaster* [dmel-all-transcript-r5.48.fasta from FlyBase (<http://flybase.org/>, last accessed on May 3, 2016)] following the procedure described by Nozawa et al. (2014). Next, transcripts that were differentially expressed (DETs) in the lines overexpressing a miRNA compared with the control line were identified using the tag count comparison method (Sun et al. 2013). If the false discovery rate (or  $q$  value; under the null hypothesis that the expression level of a transcript was equal between the conditions) was  $< 1\%$ , the transcript was regarded as a DET. To narrow down the candidates and identify the reliable target transcripts for the miRNA, the following criteria were used: (1) at least one target site in the transcript predicted by miRanda (Enright et al. 2003), PITA (Kertesz et al. 2007), or TargetScan (Ruby et al. 2007) (see table 1, [supplementary tables S3 and S4, Supplementary Material](#) online, respectively, for detailed settings) and (2)  $\min(N^{\text{cont}})/\max(N^{\text{ox}}) \geq 2$ , where  $N^{\text{cont}}$  is the normalized expression of the transcript in the control line based on the tag count comparison method,  $N^{\text{ox}}$  is the normalized expression of the

**Table 1**

The Numbers of Target Genes of miR277, miR982, and miR954 in *D. melanogaster*

| miRNA                     | Number of Target Genes |      |       |                             |       |
|---------------------------|------------------------|------|-------|-----------------------------|-------|
|                           | Larva                  | Pupa | Adult | Stage-specific <sup>a</sup> | Total |
| <b>miR277</b>             |                        |      |       |                             |       |
| Female                    | 58                     | 35   | 5     | 91                          | 94    |
| Male                      | 55                     | 75   | 38    | 117                         | 139   |
| Sex-specific <sup>b</sup> | 57                     | 58   | 37    | 108 <sup>c</sup>            | 115   |
| Total                     | 85                     | 84   | 40    | 147                         | 174   |
| <b>miR982</b>             |                        |      |       |                             |       |
| Female                    | 31                     | 5    | 2     | 34                          | 36    |
| Male                      | 22                     | 2    | 17    | 39                          | 40    |
| Sex-specific              | 17                     | 7    | 15    | 34                          | 34    |
| Total                     | 35                     | 7    | 17    | 51                          | 55    |
| <b>miR954</b>             |                        |      |       |                             |       |
| Female                    | 3                      | 7    | 7     | 17                          | 17    |
| Male                      | 5                      | 3    | 9     | 17                          | 17    |
| Sex-specific              | 6                      | 10   | 14    | 28                          | 28    |
| Total                     | 7                      | 10   | 15    | 30                          | 31    |

NOTE.—To narrow down the candidate target genes, miRanda software was used with the default settings.

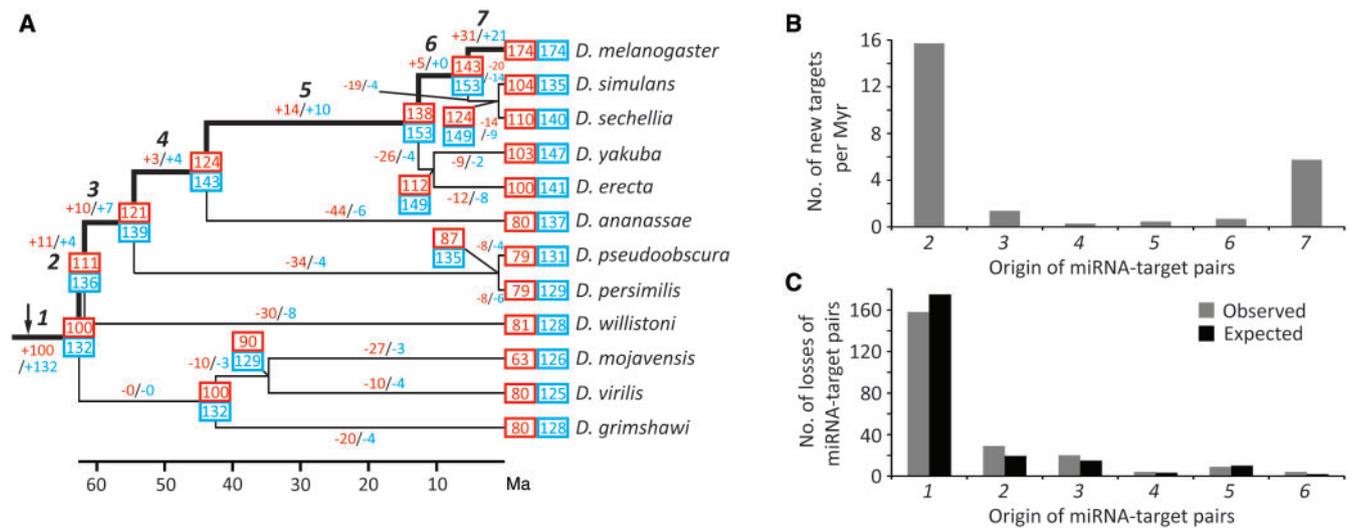
<sup>a</sup>The number of target genes that were regulated by the miRNA at only one of the developmental stages examined.

<sup>b</sup>The number of target genes that were regulated by the miRNA in only one of the sexes.

<sup>c</sup>The number of target genes that were regulated by the miRNA at only one developmental stage in one sex.

transcript in one of the overexpression lines, min is the minimum value of the two biological replicates, and max is the maximum value of the replicates. We regarded a gene as a target if at least one of its transcripts was a DET containing at least one miRNA binding site. We also counted the number of DEGs by merging the DETs derived from a single gene.

To determine whether the orthologs of target genes for a miRNA identified in *D. melanogaster* are also targets in other *Drosophila* species, transcript sequences were downloaded for the following species: *D. simulans* (dsim-all-transcript-r1.4.fasta), *D. sechellia* (dsec-all-transcript-r1.3.fasta), *D. yakuba* (dyak-all-transcript-r1.3.fasta), *D. erecta* (dere-all-transcript-r1.3.fasta), *D. ananassae* (dana-all-transcript-r1.3.fasta), *D. pseudoobscura* (dpse-all-transcript-r3.2.fasta), *D. persimilis* (dper-all-transcript-r1.3.fasta), *D. willistoni* (dwil-all-transcript-r1.3.fasta), *D. mojavensis* (dmoj-all-transcript-r1.3.fasta), *D. virilis* (dvir-all-transcript-r1.2.fasta), and *D. grimshawi* (dgri-all-transcript-r1.3.fasta). Most of the transcript sequences in these species only contained CDSs without UTRs; therefore, a gene consisting of a CDS and its 5' and 3' flanking sequences from each of the *Drosophila* species examined was aligned with the orthologous transcripts in *D. melanogaster*. We regarded an aligned region as putative transcript sequences in other species. Finally, we searched for miRNA binding sites in these putative transcripts using miRanda software with the default settings; if at least one miRNA binding site was



**FIG. 1.**—Gains and losses of miR277 target genes during *Drosophila* evolution. (A) Evolutionary changes in the number of miR277 target genes under the parsimony principle. The blue boxes show the numbers of orthologs of the genes identified as direct targets of miR277 in *D. melanogaster*. The red boxes show the numbers of orthologs with miR277 target sites at each node. The numbers along each branch indicate the gains (+) and losses (–) of these orthologs during evolution. The arrow indicates the origin of miR277. (B) The rate of gain of miR277 targets in each branch leading to *D. melanogaster*. The numbers on the x axis correspond to those shown on the branches in (A). Myr, million years. (C) The number of losses of miR277 target genes in each branch leading to *D. melanogaster* at which they were generated [branches 1–6 in (A)]. The expected numbers based on random losses were also computed (see Materials and Methods for detailed procedures).

detected in a transcript from another species, the transcript (or gene) in the species was regarded as a miRNA target.

qPCR

To examine the level of upregulation of a miRNA in overexpression flies compared with control flies, we conducted qPCR experiments. Each of the three miRNAs in the total RNA extracted with the miRNeasy Mini Kit was reverse-transcribed using TaqMan MicroRNA Reverse Transcription Kit (Life Technologies). qPCR was then conducted with the Chromo4 thermal cycler (Bio-Rad, Hercules, CA) using TaqMan Universal PCR Master Mix (Life Technologies). To examine the expression level of a control gene, *robl*, that has often been used as an internal control in the comparative threshold cycle method (Scheffe et al. 2006), PrimeScript RT reagent Kit (TaKaRa) and SYBR Premix Ex Taq (TaKaRa) were used for reverse transcription and qPCR, respectively. We made two biological replicates for each sample. In each replicate, three technical replicates were made for qPCR.

Estimation of Target Genes in Ancestors

Based on the target genes of the miRNA identified in other species, the gains and losses of miRNA-target pairs during *Drosophila* evolution were estimated using the parsimony principle. For example, if an orthologous gene existed in all 12 *Drosophila* species, but contained target sites of the miRNA only in *D. melanogaster*, *D. sechellia*, and *D. yakuba*, we concluded that the orthologous gene existed before the

divergence of the 12 *Drosophila* species, became the target of the miRNA in the lineage after splitting from *D. ananassae*, and returned to a non-target status in the lineages ancestral to *D. simulans* after splitting from *D. sechellia* and to *D. erecta* after splitting from *D. yakuba*. Here, note that independent emergences of the miRNA-target pairs in the *D. melanogaster*, *D. sechellia*, and *D. yakuba* lineages are equally parsimonious (i.e., three events). However, independent emergences (gains) are less likely than independent losses, because a loss of pairing can occur through any mutation whereas a gain of pairing would require a specific mutation. In addition, note that with the exception of *D. melanogaster*, we were unable to experimentally verify the miRNA target genes in any of the other species. Therefore, the losses of target genes in the lineages leading to *D. melanogaster* and the gains of target genes in other lineages were unable to be identified.

The gain rate of miRNA targets on each branch was estimated by dividing the number of gains of miRNA targets by the evolutionary time of the branch (Myr). For example, the number of miRNA targets newly generated on the branch 3 is 10 and the evolutionary time of the branch 3 is 62.2–54.9 = 7.3 Myr (fig. 1A). Therefore, the gain rate of miRNA targets on the branch 3 is 10/7.3 = 1.4 per Myr.

The number of losses of target genes under random loss was computed in the following way. As an example, let us focus on the branch leading to *D. willistoni* in fig. 1A. Here, the number of target gene losses is 30. These losses are classified into groups based on the timings when the gene became targets of the miRNA. If these losses occurred

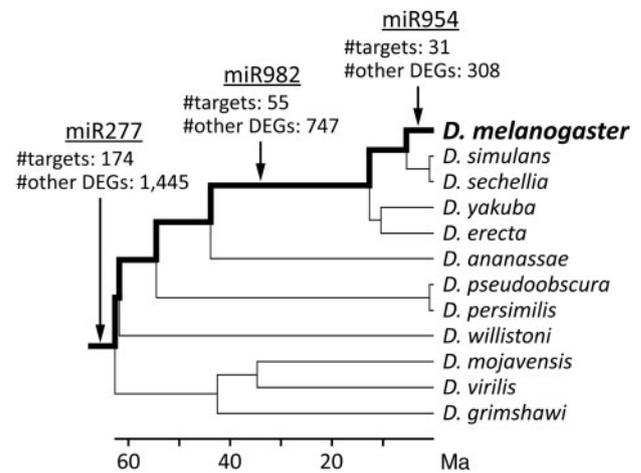
randomly, the number of losses in each group is expected to be proportional to the number of gains in each group. Therefore, the expected number of losses in group 1 becomes 27 ( $=30 \times 100/111$ ) whereas the number in group 2 is three ( $=30 \times 11/111$ ). In other words, 27 and three targets that originated in branches 1 and 2, respectively, are expected to be lost in the *D. willistoni* lineage under random loss. Similarly, the expected numbers under random loss can be estimated for all other branches. The numbers in each of the branches are then summed up for each group. Finally, a  $\chi^2$  test is conducted by comparing these expected numbers computed above and the observed number of losses in each group to calculate a statistical significance.

## Results

### An Old miRNA Has a Larger Number of Targets Than a Young miRNA

First, we examined the expression levels and breadths of two groups of miRNAs in *D. melanogaster* (see [supplementary table S5, Supplementary Material](#) online, for data sets used). One group was named “old,” because the miRNA genes in this group were generated before *Drosophila* radiation. The remaining group was referred to as “young,” because the miRNA genes included in this group emerged after the *Drosophila* radiation in the lineage to *D. melanogaster*. Consistent with the results of previous studies (Berezikov et al. 2006; Lu et al. 2008; Ma et al. 2010; Meunier et al. 2013), old miRNAs were expressed in a larger number of tissues at significantly higher levels of expression than young miRNAs ([supplementary fig. S1, Supplementary Material](#) online), indicating that the strength of connections between miRNAs and their targets in a gene regulatory network becomes stronger with time. In other words, old miRNAs are likely to suppress their targets more efficiently than young miRNAs.

Next, we focused on three miRNAs with different ages after origination, namely, miR277, miR982, and miR954. Whereas miR277 was generated before the divergence of *Drosophila* species, miR982 originated in the lineage after splitting from *D. ananassae* ([supplementary fig. S2 and table S2, Supplementary Material](#) online). By contrast, miR954 expression was detected only in *D. melanogaster*, suggesting that it was generated (or at least has become expressed) in the lineage after the divergence from *D. simulans* and *D. sechellia*. To identify the target genes of these three miRNAs, we conducted mRNA-seq and compared the expression levels of protein-coding genes between the control *D. melanogaster* flies and the transgenic flies constitutively overexpressing one of the miRNAs ([supplementary fig. S3 and table S1, Supplementary Material](#) online). Genes that were down-regulated in the flies overexpressing a miRNA and that had at least one target (or binding) site of the miRNA



**Fig. 2.**—The *Drosophila* phylogeny showing the relationships between the ages of the three miRNAs examined and the numbers of targets and of other differentially expressed genes (DEGs). The arrows indicate the branches in which the miRNA genes were generated. The time scale is based on Tamura et al. (2004). Ma, million years ago.

predicted by miRanda (Enright et al. 2003) were regarded as authentic targets (see Materials and Methods for more details). Because miRNA-target pairs can be stage- and/or sex-specific, mRNA-seq was performed in males and females at the larval, pupal, and adult stages separately. As a result, 174, 55, and 31 target genes were identified for miR277, miR982, and miR954, respectively (table 1). Many of these target genes were sex- as well as developmental stage-specific, particularly those of miR954 [90% ( $=28/31$ ) vs. 62% ( $=108/174$ ) for miR277 and 62% ( $=34/55$ ) for miR982]. Mapping the number of target genes onto the *Drosophila* phylogeny revealed that miR277, the oldest miRNA examined, had the largest number of target genes, whereas miR954, the youngest miRNA examined, had the smallest number of targets (fig. 2). The trends remained the same when different tools such as PITA (Kertesz et al. 2007) and TargetScan (Ruby et al. 2007) were applied for target prediction with the same mRNA-seq data set ([supplementary tables S3 and S4, respectively, Supplementary Material](#) online). These results are inconsistent with the decreasing model but support the increasing model of miRNA-target pairs, in which the number of target genes of a miRNA increases over time.

A similar trend was also observed between the age of the miRNA and the number of other differentially expressed genes (DEGs) that did not contain a target site of the miRNA (fig. 2). More than 1,400 genes were differentially expressed due to overexpression of miR277, whereas only 308 genes were differentially expressed following overexpression of miR954. Because miR277 is older than miR954, these observations are also consistent with the increasing model of miRNA-target pairs in which old miRNAs have larger numbers of indirect targets and affect the expression levels of more genes in

the network than young miRNAs. However, note that there was no clear trend between the age of the miRNA and the number of indirect targets per direct target [8.3 (=1,445/174) for miR277; 13.6 (=747/55) for miR982; and 9.9 (=308/31) for miR954]. This result implies that old miRNAs do not necessarily regulate the upstream genes in a cascade or a regulatory circuit, but may be involved in a greater number of cascades than young miRNAs.

We also examined the functional categories of the target genes based on gene ontology using GOrilla software (Eden et al. 2009). We did not find any enrichment for the functions of the target genes of miR982 and miR954 probably due to the small numbers of target genes. On the other hand, we found that the target genes of miR277 were often associated with acid metabolism (supplementary table S6, Supplementary Material online), being consistent with a previous experimental study of miR277 target genes (Esslinger et al. 2013). This consistency of our results with the previous study indicates that our approach is reliable to capture the authentic miRNA targets.

#### MiRNA-Target Pairs Have Been Frequently Turned over During Evolution

To get insights into the turnover of miRNA-target pairs during evolution, we next examined whether the target genes of miR277 identified in *D. melanogaster* are also targets of this miRNA in other *Drosophila* species by using the miRanda software. Based on a trait-based reconstruction of ancestral status under a parsimony framework with the assumption of a single origin of each miRNA-target pair (see Materials and Methods for details), we found that 132 (76%) orthologs of the 174 miR277 target genes identified in *D. melanogaster* were present in the ancestor of the 12 *Drosophila* species (fig. 1A); however, only 100 (76%) of these 132 orthologs possessed miR277 target sites. It follows that only 57% (=100/174) of the genes identified as targets of miR277 in *D. melanogaster* were also targets in the ancestor of the *Drosophila* species, and the remaining 43% (=74/174) must have been added to the target repertoire of miR277 during the last 63 Myr of the lineage leading to *D. melanogaster* (branches 2–7 in fig. 1A). Moreover, the conservation of targets between *D. melanogaster* and other *Drosophila* species was even lower for miR982. For miR277, 59% (=103/174) of the targets were shared between *D. melanogaster* and *D. yakuba*, whereas only 22% (=12/55) of the miR982 targets were shared between these species (supplementary fig. S4, Supplementary Material online).

The observations described above suggest that although each miRNA frequently gains new target genes during evolution, the gain rate of target genes decreases with time. To further examine this hypothesis, we investigated the patterns of gain and loss of miR277-target pairs during *Drosophila* evolution. We found that the gain rate of new targets became slower after the origination of miR277 (fig. 1B), supporting the

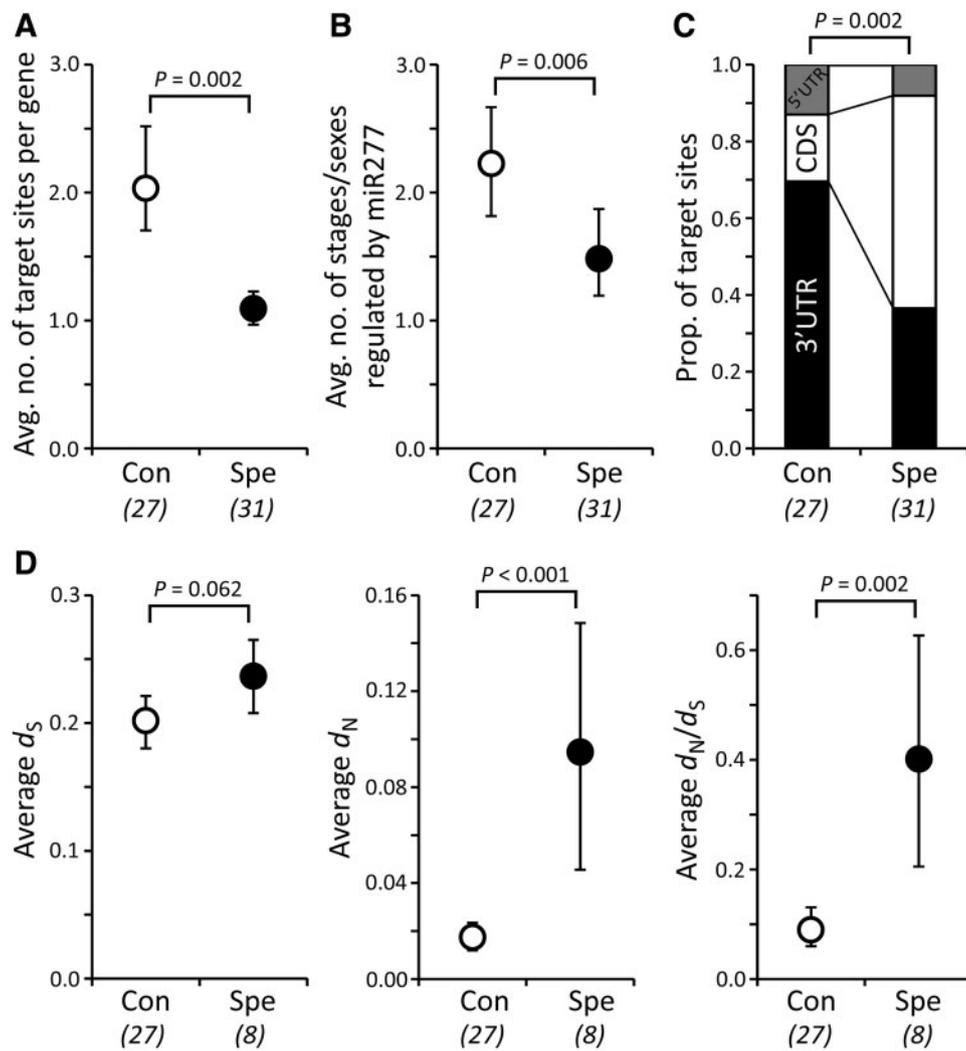
idea that the gain rate of target genes decreases with time and the number of target genes of a given miRNA eventually reaches a plateau. The gain rate was higher at the latest lineage of evolution (compare the branch 7 with branches 3–6 in fig. 1B) apparently due to the inclusion of polymorphic (or strain-specific) targets. In other words, it is possible that many of these targets may not be the targets of miR277 in other strains of *D. melanogaster* (i.e., fortuitous miRNA-target pairs that are not fixed into a population), although this argument may be too speculative.

To examine the pattern of losses of target genes, we next classified all losses of miR277-target pairs into groups based on the times at which they were established (see branches 1–6 in fig. 1A). If the losses occurred randomly without any bias irrespective of the ages of miRNA-target pairs, the number of lost pairs in each group should have been proportional to the number of gains at each branch (see Materials and Methods for details). Compared with the expectation based on random loss (black bars in fig. 1C), however, the miRNA-target pairs established before the *Drosophila* radiation were less likely to be lost (gray bars in fig. 1C), although the statistical significance of this difference was marginal ( $P=0.048$  by a  $\chi^2$  test). This result suggests that old miRNA-target pairs are more important than young miRNA-target pairs. This interpretation is also consistent with the finding that the older miR277 showed a higher proportion of conserved targets between species (e.g., *D. melanogaster* and *D. yakuba*) than the younger miR982.

It should be mentioned that our parsimony approach may be potentially biased toward the systematic increase in the number of miRNA-target pairs over time (Simkin et al. 2014). Therefore, we cannot evaluate whether the decreasing or increasing model is more appropriate based on this analysis. Nevertheless, the parsimony approach tends to underestimate the ancestral numbers of target genes and overestimate the number of gains of target genes on descendent branches, which makes our analysis conservative to support our hypothesis that the gain rate of target genes decreases with time. In relation to this argument, we used a trait-based reconstruction of ancestral binding status of miRNA-target pairs, although a nucleotide-based reconstruction of ancestral sequences to estimate ancestral status is likely to be more reliable (Simkin et al. 2014). Indeed, we also tried the nucleotide-based reconstruction. However, we decided not to use this approach, because sequence alignment of each of orthologs in 12 *Drosophila* species was unreliable in the predicted UTRs, which consequently made the estimation of ancestral sequences also untrustworthy (see Materials and Methods for details).

#### Conserved Targets Have a Larger Number of miRNA Binding Sites in Their 3'-UTRs, Evolving at a Slower Rate Than *D. melanogaster*-Specific Targets

The analyses described above revealed that some genes have been targets of miR277 throughout *Drosophila* evolution



**Fig. 3.**—Characteristics of target genes regulated by miR277 that are specific to *D. melanogaster* (Spe) or conserved in all 12 species (Con). (A) The average numbers of miR277 binding sites per target gene. (B) The average number of developmental stages/sexes in which each target gene was regulated by miR277. Note that the maximum number is six because RNA-seq was performed using male and female flies at the larval, pupal, and adult stages. (C) The locations of the miR277 binding sites in the target transcripts. (D) The evolutionary rates of the miR277 target genes. The modified Nei–Gojobori method (Zhang et al. 1998) with a transition-transversion ratio of 2 was used to estimate the synonymous distance ( $d_s$ ), non-synonymous distance ( $d_n$ ), and  $d_n/d_s$  ratio. The numbers of genes analyzed are shown in parentheses. The error bars indicate 95% confidence intervals based on the 1,000 bootstrap resampling. A Monte Carlo simulation with 1,000 replications was used to determine statistical significance. The data shown in (A–C) are based on *D. melanogaster*, and the data shown in (D) are based on the comparison of *D. melanogaster* and *D. yakuba*.

whereas others are targets only in *D. melanogaster*. (Note that the conserved targets among 12 *Drosophila* species and the *D. melanogaster*-specific targets must be robust even if the nucleotide-based reconstruction method was able to be applied.) Therefore, we examined whether there are any discriminating features between the two groups of targets. Our analyses revealed that the miR277 target genes that were conserved in all 12 *Drosophila* species have on an average a greater number of miR277 binding sites than the *D. melanogaster*-specific targets (fig. 3A). Consistent with this finding, the conserved targets tended to be regulated by miR277 at multiple

developmental stages in both sexes, whereas many of the *D. melanogaster*-specific targets were regulated at a particular stage of development in only one of the sexes (fig. 3B). Furthermore, the conserved target genes had a higher frequency of miR277 binding sites in their 3'-UTRs than the *D. melanogaster*-specific targets (70% and 36%, respectively) (fig. 3C). Therefore, although there is increasing evidence that protein-coding sequences (CDSs) as well as 5'-UTRs can also become target sites of miRNAs (Tay et al. 2008; Moretti et al. 2010; Schnall-Levin et al. 2010; Guo et al. 2015; Qian et al. 2016), 3'-UTRs may have a greater potential to maintain

target sites than CDSs and 5'-UTRs consistent with the current idea that animal conserved miRNAs normally bind to 3'-UTRs of target mRNAs (Bartel 2009). In addition, it should be mentioned that the increasing model of miRNA-target pairs remained supported even if we focused on 3'-UTRs for miRNA targeting (see [supplementary table S4](#), where only 3'-UTRs were considered as potential target sites with the usage of TargetScan, [Supplementary Material](#) online).

A difference in the evolutionary rates of the CDSs between the conserved and non-conserved miR277 targets was also observed. The non-synonymous distance ( $d_N$ ) as well as the ratio of non-synonymous to synonymous distances ( $d_N/d_S$ ) were significantly lower for the conserved targets than the *D. melanogaster*-specific targets, whereas the difference in the synonymous distances ( $d_S$ ) between these two targets was not significant (fig. 3D). These observations suggest that genes under stronger functional constraints at the protein level are also regulated at the mRNA level in a more stringent manner. Hence, these genes may have been consistently regulated by miR277 for fine-tuning of gene expression during long-term evolution.

#### Other *Drosophila* miRNAs Also Do Not Support the Decreasing Model

The analyses above do not support the decreasing model. However, drawing conclusions based on only three miRNAs might be unreliable, although we chose these miRNAs just based on the availability of flies and vectors without any intention. Therefore, it is important to analyze a greater number of miRNAs and see whether the observations obtained above are robust. As it was impractical to apply the above procedures to all *Drosophila* miRNAs, we designed the following approach. First, bioinformatics prediction was applied to identify the candidate targets for each of all miRNAs in *D. melanogaster*. Second, using the available data of small RNA-seq and mRNA-seq ([supplementary table S7](#), [Supplementary Material](#) online), we examined whether the candidate miRNA-target pairs identified *in silico* were co-expressed *in vivo*. If a pair was not co-expressed in any of the tissues examined, we removed the pair from the list as a false positive. In this way, we inferred the *in vivo* conditions of miRNAs and mRNAs, to some extent, from the data analysis. As mentioned above, we classified all miRNA genes into "old" and "young" miRNAs depending on the timings when the miRNA genes were generated.

As a result, we found that miRanda and PITA give a weak trend that old miRNAs have greater numbers of target genes than young one. Although the result obtained by TargetScan was slightly different from those by miRanda and PITA (fig. 4), even in this case the decreasing model was not statistically supported. Note that in reality the probability of making pairs with mRNAs *in vivo* would be even higher for old miRNAs than young miRNAs, because the expression level

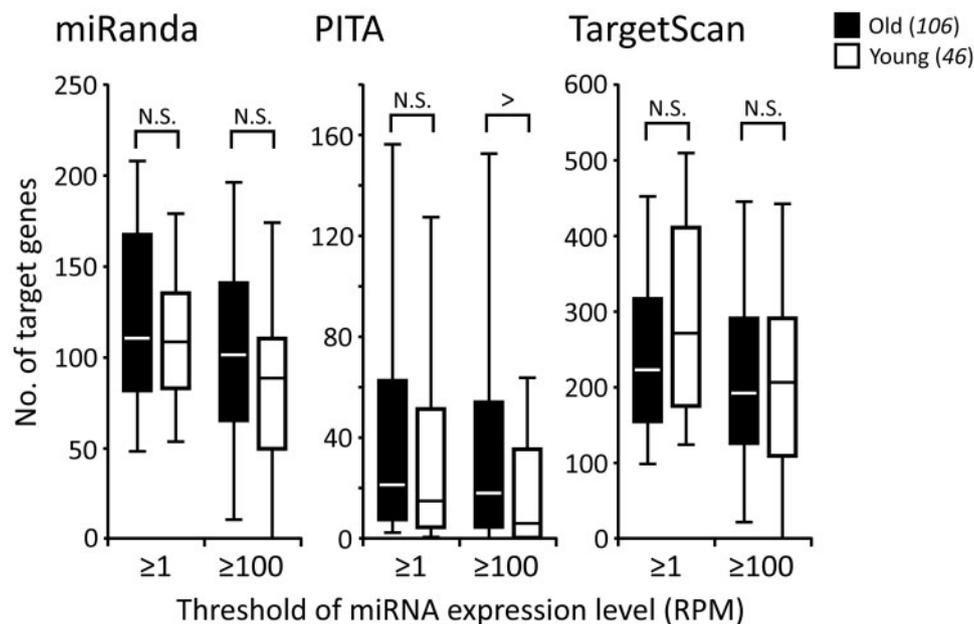
of old miRNAs is much higher than young ones ([supplementary fig. S1](#), [Supplementary Material](#) online). Indeed, the difference in the number of target genes for old and young miRNAs became greater when a more stringent threshold of miRNA expression was applied (fig. 4). In particular, old miRNAs showed a significantly larger number of target genes than young miRNAs in the PITA prediction.

#### Data in Other Organisms Again Do Not Support the Decreasing Model

To examine the evolutionary transition of miRNA-target pairs in different organisms, we analyzed human miRNA-target pairs based on cross-linking immunoprecipitation sequencing (CLIP-seq) data that have been deposited in the starBase v2.0 database (Li et al. 2014). The CLIP-seq data were generated by immuno-precipitating RNA-induced silencing complexes using an anti-Argonaute antibody and by then sequencing the miRNA and mRNA fractions derived from the samples. After sequencing, bioinformatics tools were used to predict miRNA-mRNA pairs. Our analyses showed that old miRNAs generated before the divergence from platypuses have a larger number of target genes than young miRNAs emerged after splitting from elephants and armadillos, although the absolute numbers of target genes varied depending on the prediction software used (fig. 5). The result based on TargetScan was again a bit different from those based on miRanda and PITA, but even in this case the decreasing model was not supported. It should be mentioned that Ma et al. (2010) also reported a greater number of target genes for older miRNAs than that for younger miRNAs in *Arabidopsis* in which miRNA-target pairs were determined by degradome sequencing. These results are consistently incompatible with the decreasing model and raise the possibility that the increasing model of miRNA-target pairs is applicable to many organisms.

## Discussion

In this study, we investigated the evolution of miRNA-target pairs, with a special focus on the number of target genes of each miRNA. In the analyses of three miRNAs (miR277, miR982, and miR954) in *D. melanogaster*, the number of target genes was positively correlated with the age of the miRNA; the oldest miRNA examined (miR277) had the largest number of targets, and the youngest miRNA (miR954) had the smallest number of targets. The number of predicted targets co-expressed with a miRNA also showed a weak but similar tendency in other miRNAs of *D. melanogaster*. Moreover, the same trend was observed weakly in humans and strongly in *Arabidopsis*. Therefore, the results collectively do not support the decreasing model but may be consistent with the increasing model of miRNA-target pairs. We also found that the turnover of miRNA-target pairs is high during evolution, but old miRNA-target pairs tend to be maintained with a certain extent. Indeed, the older miR277 showed a higher proportion

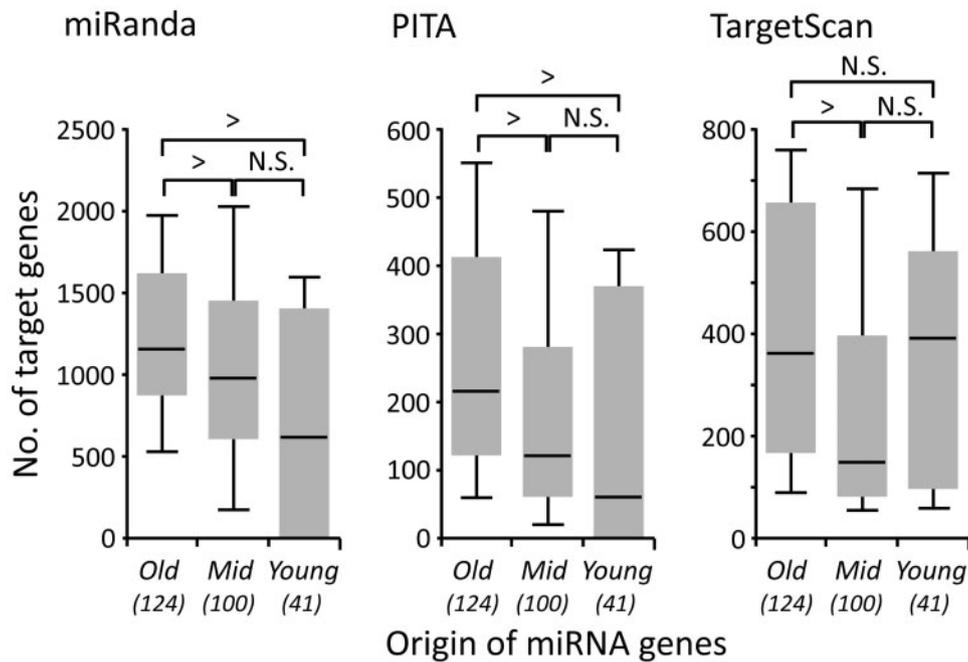


**Fig. 4.**—The numbers of target genes for each of the old and young miRNAs in *Drosophila melanogaster*. miRanda (Enright et al. 2003), PITA (Kertesz et al. 2007), and TargetScan (Ruby et al. 2007) were used for predicting miRNA targets. Only co-expressed miRNAs and mRNAs were considered to infer *in vivo* conditions of miRNA-target pairs. Two different thresholds (reads per million mapped reads or RPM  $\geq 1$  and 100) were used to determine whether a miRNA is expressed, whereas reads per kilobase of exon per million mapped reads (RPKM)  $\geq 1$  was adopted for mRNA expression. Black and white rectangles indicate the number of target genes for old and young miRNAs, respectively, and the numbers in parentheses indicate the number of miRNA genes in each group. Here, old and young miRNAs are the miRNAs generated before and after the *Drosophila* radiation, respectively. The lines in the boxes represent the median; 50% of the values are included in the boxes, and 80% of the values are included within the bars. Statistical significance was based on Monte Carlo simulation with 1,000 replications. NS, not significant ( $P > 0.05$ );  $>$ ,  $P < 0.05$ . A max score of  $\geq 160$  in miRanda (instead of the default settings in table 1) and a target score of  $\leq -15$  (in addition to the criteria in supplementary table S3, Supplementary Material online) in PITA was used to predict target genes. In TargetScan, only 8mer-1a (perfect Watson-Crick base pairing of 2–8 nucleotide positions in the seed region with target sites as well as A in the 1st position of the mature region) and 7mer-m8 (perfect Watson-Crick base pairing of 2–8 nucleotide positions in the seed region with target sites) were considered as predicted target sites (in addition to the criteria in supplementary table S4, Supplementary Material online).

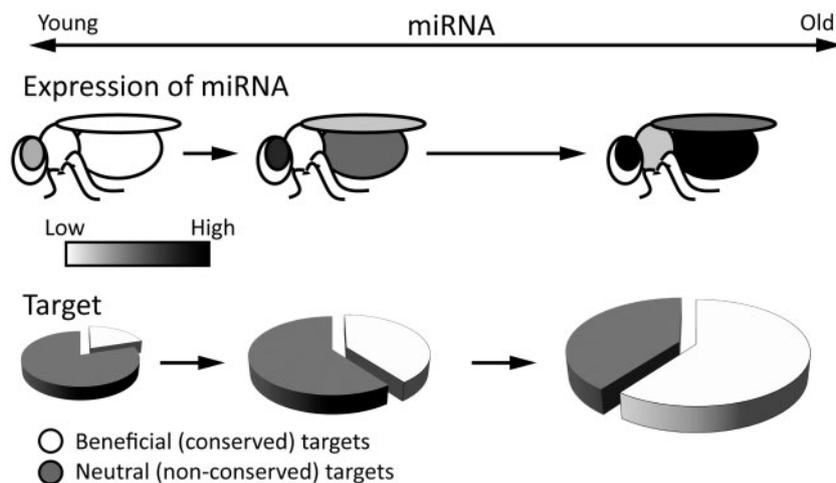
of shared target genes between *D. melanogaster* and *D. yakuba* than the younger miR982. A subsequent analysis of miR277 target genes demonstrated that those that are conserved among species tend to have multiple target sites in their 3'-UTRs and evolve at a slower rate than lineage-specific targets.

Based on the findings presented here, we tentatively propose a possible model for the evolution of miRNA-target pairs (fig. 6). In this model, a new miRNA has a small number of targets, most of which are effectively neutral, and only a small proportion of the target genes are beneficial and provide biological functions to the organism. The miRNA quickly loses most of its neutral targets due to random mutation in the targets (or the miRNA with less frequency) as well as subsequent genetic drift. (Many young miRNA genes themselves would also disappear quickly if they do not find any targets that are biologically meaningful.) By contrast, a few miRNA-target pairs with functional importance are maintained under purifying selection. During this period, the expression level and/or breadth of the miRNA may increase to enable efficient suppression of its important targets. At the same time, the

miRNA may also acquire new targets because the chances of forming pairs with mRNAs become higher as the expression level/breadth of the miRNA increases. Most of these new miRNA-target pairs are again effectively neutral and break down, but a few may become functional pairs. If this process continues, the number of conserved target genes with functional importance would increase over time, whereas most of the neutral pairs would be constantly turned over. Consequently, the number of target genes of the miRNA is expected to increase and eventually reach a plateau. In this way, each miRNA is integrated gradually into the regulatory network of the organism by increasing the number of connections during long-term evolution. It should be noted that, in this model, we do not consider deleterious miRNA-target pairs because if an emerging miRNA has a deleterious target, the miRNA is unlikely to be fixed into a population. In addition, if a miRNA that is already fixed into a population acquires a deleterious target in an individual, the individual with the deleterious pair would have a lower fitness and be eliminated from the population. In this way, the deleterious miRNA-target pairs would not contribute to long-term evolution.



**Fig. 5.**—The numbers of miRNA target genes in humans identified using CLIP-seq data (Chi et al. 2009) with miRanda (Enright et al. 2003), PITA (Kertesz et al. 2007), and TargetScan (Ruby et al. 2007). The numbers in parentheses indicate the number of miRNA genes in each group. The lines in the boxes represent the median; 50% of the values are included in the boxes, and 80% of the values are included within the bars. Statistical significance was based on Monte Carlo simulation with 1,000 replications. NS, not significant ( $P > 0.05$ );  $>$ ,  $P < 0.05$ . The miRNA genes were classified into three groups depending on their ages (Iwama et al. 2013): Old, miRNA genes generated before the divergence from platypuses; Middle, miRNA genes generated after the divergence from platypuses but before the divergence from Atlantogenata (e.g., elephants and armadillos); Young, miRNA genes generated after the divergence from Atlantogenata. The information of miRNA-target pairs based on the CLIP-seq data with predictions by several tools was retrieved from starBase v2.0 (Li et al. 2014).



**Fig. 6.**—A possible model for the evolution of miRNA-target pairs. Most of the miRNA-target pairs that are fixed in a population are effectively neutral at the initial stage. Deleterious pairs are not considered because they would not be fixed in a population and would not contribute to long-term evolution. Most of the neutral pairs are removed by random mutations and genetic drift. Only a small number of pairs that acquire solid functions are maintained under purifying selection. During this process, the expression level and breadth of the miRNA can be increased to enable fine-tuning and more efficient suppression of their functional targets. Therefore, the miRNA may find new mRNAs as targets, most of which are again effectively neutral and turned over quickly. In this way, the number of targets of the miRNA increases over time and eventually reaches a plateau. In addition, the proportion of functional targets under purifying selection also increases over time.

However, slightly deleterious pairs could be fixed if the population size is small (Ohta 1973). Therefore, it would be interesting to examine the number of target genes for miRNAs in a species with a small population size, such as *D. sechellia* (Legrand et al. 2009).

Although *in silico* analysis may predict large numbers of targets of young miRNAs, binding of these miRNAs to the predicted target sites may not occur *in vivo* due to low expression levels or restricted tissue distribution of the miRNAs. Indeed, we found a weak but general trend that the number of target genes for young miRNAs tended to be smaller than that for old miRNAs when co-expression of miRNAs and mRNAs was taken into account. In addition, note that the expression levels of the majority of the predicted target genes using an *in silico* approach (miRanda, PITA, or TargetScan) were not affected significantly by overexpression of the miRNA *in vivo* (supplementary table S8, Supplementary Material online). It is possible that some authentic targets were not detected as DEGs because they are regulated via miRNA-mediated translational repression rather than mRNA degradation. However, this large discrepancy between the predicted targets and the DEGs is quite unlikely if the prediction tools reflect the situation *in vivo*. In addition, it has been reported that many animal miRNAs induce the degradation of their target mRNAs in addition to translational repression (Huntzinger and Izaurralde 2011). Of course, bioinformatics predictions are useful to understand the regulatory capacity of miRNAs and narrow down the candidate targets for screening via experimental approaches, but a blind use of prediction tools may lead to inaccurate conclusions (Wang 2006).

The decreasing model of miRNA-target pairs was originally proposed by Chen and Rajewsky (2007) and later supported by Roux et al. (2012). Note that Roux et al. (2012) mainly used mouse data to support the decreasing model, whereas we have mainly focused on *Drosophila* and humans. Therefore, it is possible that evolutionary patterns of miRNA-target pairs are considerably different among species. Yet, if we consider that humans and mice are both mammals and share a long ancestry, this possibility is quite unlikely. It should also be noted that in the study by Roux et al. (2012) a negative correlation between the age of miRNAs and the number of their targets was not statistically significant and their data sets of miRNA-target pairs were not really experimentally identified. Therefore, our data sets are likely to be more reliable to study the evolutionary transitions of miRNA-target pairs, although further studies with larger data sets in a wider range of organisms are apparently necessary.

The mutant flies generated in this study overexpressed the particular exogenous miRNA constitutively in all tissues at all developmental stages; however, it is unlikely that all tissues spontaneously express the endogenous miRNA. Therefore, a certain fraction of the target genes identified in this study are likely false positives. However, young miRNAs are generally expressed in restricted tissues at low levels; hence, the

difference in the expression level between the wild-type and transgenic flies would be greater for young miRNAs than old miRNAs. Consequently, the number of false positives would be expected to be greater for younger miRNAs than older miRNAs. Indeed, our quantitative PCR (qPCR) experiments clearly showed that the fold change in gene expression of a miRNA due to the overexpression was much more conspicuous for the younger miR982 and miR954 than the older miR277 (supplementary fig. S5, Supplementary Material online). In addition, the absolute difference in the expression level between transgenic and the wild-type flies was also mostly greater for the younger miR982 and miR954 than the older miR277 (supplementary table S9, Supplementary Material online). With this approach, we are therefore conservative to conclude the possibility of the increasing model of miRNA-target pairs. In relation to this argument, it should be noted that the experimental design used to identify target genes of miRNAs here is somewhat analogous to bioinformatics predictions because constitutive overexpression of a miRNA theoretically enables to identify all mRNAs that have sequence complementarity to the miRNA. Despite this fact, the number of target genes was still smaller for the young miR954 than the old miR277. This observation is unlikely if only the expression level and tissue distribution of a miRNA determine the number of target genes. Therefore, other factors may also affect the formation of miRNA-target pairs *in vivo*, and further studies are required to clarify this point.

In summary, the decreasing model of miRNA-target pairs must be reconsidered based on the combination of experimental and bioinformatics approaches. We tentatively conclude that each miRNA is integrated gradually into the regulatory network of an organism by forming increasing numbers of connections during long-term evolution. This process may have provided regulatory innovations such as robustness and plasticity to the network. Yet, further studies are apparently necessary to evaluate this increasing model extensively. We are currently establishing a new experimental method, in which we do not need any bioinformatics predictions to identify all miRNA-target pairs in tissues. Hopefully, we can publish the method in the near future. This type of studies must provide significant insights into not only the evolution of miRNA-target pairs, but also the mechanism by which the miRNA-based gene regulatory network has evolved.

## Supplementary Material

Supplementary figs. S1–S5 and tables S1–S9 are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

## Acknowledgments

We thank Shu Kondo for advice regarding the *Drosophila* crossing experiments and Miu Kubota for help with the

experiments. We are also grateful to Junichi Imoto, Sonoko Kinjo, Norikazu Kitamura, Kaoru Matsumoto, Masatoshi Nei, Masa-aki Yoshida, and Ikuko Yuyama for their comments on earlier versions of the manuscript. We also thank the associate editor and the two reviewers for their constructive comments on our work. Computations were partially performed on the NIG supercomputer at National Institute of Genetics. This work was supported by grants from the National Institute of Genetics, the Center for the Promotion of Integrated Sciences (CPIS), and JSPS KAKENHI Grant Number 25711023 to M.N.

## Literature Cited

- Adams MD, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287:2185–2195.
- Allen E, et al. 2004. Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat Genet*. 36:1282–1290.
- Ambros V. 2004. The functions of animal microRNAs. *Nature* 431:350–355.
- Axtell MJ, Bowman JL. 2008. Evolution of plant microRNAs and their targets. *Trends Plant Sci*. 13:343–349.
- Axtell MJ, Snyder JA, Bartel DP. 2007. Common functions for diverse small RNAs of land plants. *Plant Cell* 19:1750–1769.
- Barbash S, Shifman S, Soreq H. 2014. Global coevolution of human microRNAs and their target genes. *Mol Biol Evol*. 31:1237–1247.
- Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297.
- Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* 136:215–233.
- Berezikov E, et al. 2006. Diversity of microRNAs in human and chimpanzee brain. *Nat Genet*. 38:1375–1377.
- Berezikov E, et al. 2010. Evolutionary flux of canonical microRNAs and mirtrons in *Drosophila*. *Nat Genet*. 42:6–9.
- Chen K, Rajewsky N. 2007. The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet*. 8:93–103.
- Chi SW, Zang JB, Mele A, Darnell RB. 2009. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460:479–486.
- Clark AG, et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203–218.
- Duffy JB. 2002. GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis* 34:1–15.
- Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48.
- Enright AJ, et al. 2003. MicroRNA targets in *Drosophila*. *Genome Biol*. 5:R1.
- Esslinger SM, et al. 2013. *Drosophila* miR-277 controls branched-chain amino acid catabolism and affects lifespan. *RNA Biol*. 10:1042–1056.
- Fahlgren N, et al. 2007. High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One* 2:e219.
- Gratz SJ, et al. 2013. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194:1029–1035.
- Guo ZW, et al. 2015. MtiBase: a database for decoding microRNA target sites located within CDS and 5'-UTR regions from CLIP-Seq and expression profile datasets. *Database* 2015:bav102.
- Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet*. 12:99–110.
- Iwama H, Kato K, Imachi H, Murao K, Masaki T. 2013. Human microRNAs originated from two periods at accelerated rates in mammalian evolution. *Mol Biol Evol*. 30:613–626.
- Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. 2007. The role of site accessibility in microRNA target recognition. *Nat Genet*. 39:1278–1284.
- Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 42:D68–D73.
- Lee RC, Feinbaum RL, Ambros V. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75:843–854.
- Legrand D, et al. 2009. Species-wide genetic variation and demographic history of *Drosophila sechellia*, a species lacking population structure. *Genetics* 182:1197–1206.
- Li JH, Liu S, Zhou H, Qu LH, Yang JH. 2014. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 42:D92–D97.
- Lu J, et al. 2008. The birth and death of microRNA genes in *Drosophila*. *Nat Genet*. 40:351–355.
- Ma Z, Coruh C, Axtell MJ. 2010. *Arabidopsis lyrata* small RNAs: transient MIRNA and small interfering RNA loci within the *Arabidopsis* genus. *Plant Cell* 22:1090–1103.
- Meunier J, et al. 2013. Birth and expression evolution of mammalian microRNA genes. *Genome Res*. 23:34–45.
- Moretti F, Thermann R, Hentze MW. 2010. Mechanism of translational regulation by miR-2 from sites in the 5' untranslated region or the open reading frame. *RNA* 16:2493–2502.
- Nozawa M, Fukuda N, Ikeo K, Gojobori T. 2014. Tissue- and Stage-Dependent Dosage Compensation on the Neo-X Chromosome in *Drosophila pseudoobscura*. *Mol Biol Evol*. 31:614–624.
- Nozawa M, Miura S, Nei M. 2010. Origins and evolution of microRNA genes in *Drosophila* species. *Genome Biol Evol*. 2:180–189.
- Nozawa M, Miura S, Nei M. 2012. Origins and evolution of microRNA genes in plant species. *Genome Biol Evol*. 4:230–239.
- Ohta T. 1973. Slightly deleterious mutant substitutions in evolution. *Nature* 246:96–98.
- Piriyaopongsa J, Jordan IK. 2008. Dual coding of siRNAs and miRNAs by plant transposable elements. *RNA* 14:814–821.
- Qian J, Tu R, Yuan L, Xie W. Forthcoming 2016. Intronic miR-932 targets the coding region of its host gene, *Drosophila neuroigin2*. *Exp Cell Res*.
- Richards S, et al. 2005. Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res*. 15:1–18.
- 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*. 40:5890–5900.
- Ruby JG, et al. 2007. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res*. 17:1850–1864.
- Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
- Scheffe JH, Lehmann KE, Buschmann IR, Unger T, Funke-Kaiser H. 2006. Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. *J Mol Med*. 84:901–910.
- Schnall-Levin M, Zhao Y, Perrimon N, Berger B. 2010. Conserved microRNA targeting in *Drosophila* is as widespread in coding regions as in 3'-UTRs. *Proc Natl Acad Sci U S A*. 107:15751–15756.
- Shomron N, Golan D, Hornstein E. 2009. An evolutionary perspective of animal microRNAs and their targets. *J Biomed Biotechnol*. 2009:594738.
- Simkin AT, Bailey JA, Gao FB, Jensen JD. 2014. Inferring the evolutionary history of primate microRNA binding sites: overcoming motif counting biases. *Mol Biol Evol*. 31:1894–1901.

- Sun J, Nishiyama T, Shimizu K, Kadota K. 2013. TCC: an R package for comparing tag count data with robust normalization strategies. *BMC Bioinformatics* 14:219.
- Szuplewski S, et al. 2012. MicroRNA transgene overexpression complements deficiency-based modifier screens in *Drosophila*. *Genetics* 190:617–626.
- Tamura K, Subramanian S, Kumar S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol Biol Evol.* 21:36–44.
- Tang T, et al. 2010. Adverse interactions between micro-RNAs and target genes from different species. *Proc Natl Acad Sci U S A.* 107:12935–12940.
- Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455:1124–1128.
- Vicoso B, Bachtrog D. 2013. Reversal of an ancient sex chromosome to an autosome in *Drosophila*. *Nature* 499:332–335.
- Wang X. 2006. Systematic identification of microRNA functions by combining target prediction and expression profiling. *Nucleic Acids Res.* 34:1646–1652.
- Zhang J, Rosenberg HF, Nei M. 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc Natl Acad Sci U S A.* 95:3708–3713.