Maternal microRNAs in *Drosophila* eggs: selection against target sites in maternal protein-coding transcripts

Antonio Marco

School of Biological Sciences, University of Essex, Colchester, United Kingdom

Address for correspondence:

School of Biological Sciences

University of Essex

Wivenhoe Park, Colchester CO4 3SQ

United Kingdom

Email: amarco.bio@gmail.com

Telephone: +44 (0) 120 687 3339

ABSTRACT

In animals, before the zygotic genome is expressed, the egg already contains gene products deposited by the mother. These maternal products are crucial during the initial steps of development. In *Drosophila melanogaster* a large number of maternal products are found in the oocyte, some of which are indispensable. Many of these products are RNA molecules, such as gene transcripts and ribosomal RNAs. Recently, microRNAs – small RNA gene regulators – have been detected early during development and are important in these initial steps. The presence of some microRNAs in unfertilized eggs has been reported, but whether they have a functional impact in the egg or early embryo has not being explored. To characterize a maternal microRNA set, I have extracted and sequenced small RNAs from *Drosophila* unfertilized eggs. The unfertilized egg is rich in small RNAs, particularly in ribosomal RNAs, and contains multiple microRNA products. I further validated two of these microRNAs by qPCR and also showed that these are not present in eggs from mothers without Dicer-1 activity. Maternal microRNAs are often encoded within the intron of maternal genes, suggesting that many maternal microRNAs are the product of transcriptional hitch-hiking. Comparative genomics and population data suggest that maternally deposited transcripts tend to avoid target sites for maternally deposited microRNAs. A potential role of the maternal microRNA mir-9c in maternal-to-zygotic transition is also discussed. In conclusion, maternal microRNAs in *Drosophila* have a functional impact in maternal protein-coding transcripts.

INTRODUCTION

In animals, the initial steps of embryonic development are driven by the gene products deposited by the mother in the egg. For instance, in *Drosophila melanogaster*, the anteroposterior axis is determined by the presence of maternal transcripts from genes such as *bicoid* and *nanos* [1]. Traditionally, maternal genes have been identified by genetic analysis [1]. More recently, the roles of microRNAs during development have become a major research area. MicroRNAs are small RNA molecules that regulate gene expression by targeting gene transcripts by sequence complementarity. MicroRNAs are expressed during early development [2,3], and they target other embryonic expressed gene transcripts [4,5]. As a matter of fact, a number of homeotic genes detected by genetic analysis were later shown to be microRNA encoding genes (reviewed in [6]). The characterization of maternal microRNAs is particularly difficult as they are too short for standard genetic analyses. However, it is now possible to isolate small RNAs from egg extracts and characterize their products with RNAseq or microarray technologies. For instance, the microRNA content of mouse [7] and cow [8] oocytes have been characterized with this high-throughput approach. In other cases, such as in zebrafish [9] and *Xenopus* [10], microRNAs appear to have a minor presence in oocytes.

Several lines of evidence suggested that, in *Drosophila melanogaster*, maternally transmitted microRNAs are important. First, some microRNAs are highly abundant during early development, and then their expression levels drop as development progresses [11]. Also, the enzymes responsible for microRNA biogenesis are highly expressed in the ovaries [12] and microRNAs may have a role in oocyte maturation [13]. As a matter of fact, some mature microRNAs have been identified in *Drosophila* unfertilized eggs [14–16]. Likewise, other small RNAs are maternally transmitted, mainly, piRNAs involved in the response against transposable elements [17,18]. Recently, it has been shown that maternally transmitted microRNAs are adenylated during the maternal-to-zygotic transition [16] .Whether maternal microRNAs have a

functional impact in *Drosophila* eggs is still unknown.

To identify which microRNAs are maternally transmitted I extracted and sequenced small RNAs from *Drosophila* unfertilized eggs. To explore their potential function I predicted their targets in maternal and non-maternal gene products. The evolutionary impact of maternal microRNAs was estimated by using comparative genomics and population data.

RESULTS

Mature microRNAs are maternally deposited in the egg

To characterize the maternal microRNA set of *Drosophila melanogaster* I first extracted RNA from unfertilized eggs (see Materials and Methods) and prepared small cDNA libraries for high-throughput sequencing. MicroRNAs detected in unfertilized eggs must have come from the mother, i.e. they are maternal microRNAs. The most abundant microRNAs in unfertilized eggs were mir-92b, mir-184, the mir-310/mir-311/mir-312/mir-313 cluster and bantam, which account for over a half of the microRNA reads. Table 1 shows microRNA loci producing more than 13 reads (1‰ of the microRNA-associated reads). A full list of detected microRNAs with their read counts is in Supplementary Table 1. The dataset was screened for new microRNAs as previously described [19,20], but no new microRNAs were found. This tell us that maternal microRNAs are already known in *Drosophila*.

Figure 1 compares the relative expression of maternal microRNAs in the ovary, unfertilized eggs and early stages of development. From this comparison three types of maternal microRNAs can be distinguished. First, some maternal microRNAs are highly expressed in the ovary. A second class consists on microRNAs that are found primarily in the unfertilized egg. Third, a large proportion of maternal microRNAs is also transcribed later on during development. These groups are referred hereinafter as 'high in ovary', 'high in egg' and 'high in zygote' maternal microRNAs. A

similar classification of maternal products has been proposed elsewhere [15]. Some of these microRNAs were detected at very low levels, and whether they are bona fide maternal microRNAs may need further evidence. For instance, only one read for mir-34 was detected (Supplementary Table 1).

To further confirm the presence of maternal microRNAs in unfertilized eggs, I validated the presence of mir-311 and mir-995 mature products by qPCR (see Materials and Methods). Both products were present at a significant level in unfertilized eggs (Figure 2). I also generated females that are mutants for *Dicer-1* in the germ line (see Materials and Methods). The levels of microRNAs in *Dicer-1* mutant eggs is undetectable (Figure 2A), reinforcing the idea that microRNAs in unfertilized oocytes are produced by the mother and maternally deposited during oogenesis. Additionally, I measured the microRNA levels in fertilized embryos (2-6 hours) showing that, as suggested in Figure 1, mir-311 and mir-995 are abundant in the egg, and their level is reduced during early development, although the difference was only significant for mir-311 (Figure 2A).

In a recent report, Narry Kim and collaborators identified maternally transmitted microRNAs in *Drosophila* and demonstrated that they are targeted to degradation during maternal-to-zygotic transition (MZT) by adenylation via Wispy [16]. Their set of maternal microRNAs is virtually identical to the set here described. Their top five loci producing maternal microRNAs are the mir-310/311/312/313 cluster, mir-995, bantam, mir-184 and mir-14 (Supplementary Table 1 in [16]). All these are among the highest expressed microRNAs in Table 1. With the exception of mir-994/318, all the maternal microRNAs from Table 1 are also at a high copy number in their dataset. Overall, both datasets are highly correlated ($R=0.79$; $p<0.001$). This further supports that the microRNAs here studied are *bona fide* maternally transmitted gene products.

Intronic maternal microRNAs hosted in maternal protein-coding genes

In a previous work I observed that female biased microRNAs tend to be produced from introns of

female biased protein coding transcripts [21]. For instance, mir-92a is highly expressed in females, and it is encoded within the *jigr1* gene, which is maternally deposited in the egg. Here I show that mir-92a is also maternal. To further explore the relationship between maternal microRNAs and the maternal deposition of overlapping genes, I compared the expression pattern of intronic maternal microRNAs and the host protein coding gene. Table 2 lists 12 maternal microRNA clusters hosted in protein coding genes. For nine of these host genes there are in situ hybridization experiments [22,23], and eight of them are maternally loaded. Since 55.8% of genes in this dataset are shown to produce maternally deposited transcripts, our set of host genes is statistically enriched for maternal products ($p \sim 0.044$; binomial test). For a tenth host gene, *grp*, there is no information from high-throughput in situ hybridization analyses, but it is known to be present in unfertilized oocytes [24]. The other two host genes have no expression information available at FlyBase. From this section I conclude that intronic maternal microRNAs are frequently produced from introns of maternally deposited gene transcripts.

Maternal microRNAs in the Maternal-to-Zygotic Transition

As shown in Figure 1, a significant fraction of maternal microRNAs have a lower expression when zygotic transcription starts. One possibility of that some of these maternal microRNAs have a role in destabilizing maternal transcripts during the MZT. A similar role has been described for early expressed zygotic microRNAs in *Drosophila* [25] and other species such as zebrafish [26]. I predicted target sites for each maternal microRNAs in stable and unstable maternal transcripts during MZT [27]. Table 3 shows maternal microRNAs targeting more unstable maternal transcripts than expected by chance (FDR < 10%). Two of the microRNAs, mir-283 and mir-277, were detected at very low levels in unfertilized eggs (Suppl. Table 1) and have a higher expression level later on during embryonic development (Figure 1). It is possible that these microRNAs contribute to the destabilization of maternal transcripts, but probably as zygotic microRNAs. A third set of

microRNAs which may contribute to transcript clearance during MZT are the mir-310 and mir-92 families. They both share the same seed sequence (which defines the targeted transcripts). Last, the microRNA-9 family also target unstable maternal transcripts. These microRNAs are abundant in the *Drosophila* unfertilized egg, and may be contributing to the MZT. In summary, some maternally deposited microRNAs have a potential role in destabilizing maternal transcripts.

Maternal protein-coding transcripts avoid target sites for maternal microRNAs

If maternal microRNAs have a functional impact on maternal transcripts, these transcripts should have a different target repertoire compared to non-maternal (zygotic) transcripts. Thus, I estimated how many maternal and zygotic transcripts are targeted by maternal microRNAs. Overall, 73% of maternal transcripts and 63% of zygotic transcripts have canonical seed target sites for maternal microRNAs. However, transcripts from genes with different evolutionary conservation have different target profiles. Figure 2B shows the proportion of targets for maternal microRNAs among maternal and zygotic transcripts with different degrees of evolutionary conservation. Strikingly, for evolutionary young genes (*D. melanogaster*-specific) maternal transcripts were less likely to be targeted by maternal microRNAs than zygotic transcripts. This suggests that, at the evolutionary short-term, maternal microRNA target sites may be avoided by maternal transcripts.

To test whether there is a selection against maternal microRNA target sites we have to evaluate population data. To do so, I first constructed a model of microRNA target mutation as follows (see Figure 3A): 1) a target site is defined as any 6 nucleotide sequence (sixmer) in a 3'UTR complementary to the seed region [28] of a microRNA; 2) any target site has 18 mutant neighbours, which are one nucleotide mutation apart from the canonical target, and are not themselves targets; 3) only polymorphic sites in which one of the alleles is a target site and the other a non-target are further considered in this analysis. Allele frequency is defined as the proportion of the target allele (p in Figure 3B). For instance, an allele frequency of 0.8 means that 80% of the sampled individuals

have the target site at a given position and 20% have a non-target mutant neighbour. Conversely, an allele frequency of 0.3 will show that the non-target neighbour is more frequent (70%) than the target allele (30%). Population genetics theory [29,30] predicts that, in a finite population, two alleles neutral to each other will have a U-shaped distribution, that is, most individuals will be homozygous for one of the alleles. However, if there is a selection pressure to conserve a target site, the distribution will be shifted to the right. Conversely, if selection is against a target site allele, the distribution will be shifted to the left (see Figure 3C). By comparing the allele frequency distribution of target sites between maternal and zygotic transcripts we can estimate the relative selection pressure on microRNA target sites in maternal with respect to zygotic transcripts.

To do so, polymorphic sites for target/non-target seed sequences in *Drosophila* 3'UTRs are identified and then the allele frequency is calculated. The distribution of allele frequencies between maternal and zygotic gene transcripts for each microRNA can then be compared. Figure 3D shows the case for mir-995 microRNA products. One of them, mir-995-3p is abundant in unfertilized eggs whilst the alternate arm, mir-995-5p, is virtually absent in eggs. By comparing the allele frequency distributions for both arms we observe that the degree of non-target conservation in maternal transcripts with respect to zygotic transcripts is is higher in mir-995-3p than in mir-995-5p. That is, the dark grey bar is higher than the light grey bar for allele frequency less than 0.1 in mir-995-3p in Figure 3D. In other words, there is a preference for alleles that avoid being targeted by the maternal mir-995-3p, but no such a preference for non-maternal mir-995-5p. Both arms of mir-305 are present at high levels in unfertilized eggs. Figure 4A shows the allele frequency distribution for their targets, and both arms show evidence of target avoidance. As a counterexample, Figure 4B shows the allele frequency distribution of a microRNA for which none of the arms was detected in unfertilized eggs: mir-4986. Consistently, none of the microRNA products showed evidence of target avoidance by maternal transcripts.

To explore whether this pattern is a general feature of maternal microRNAs I defined a

measure of target avoidance at the population level as the log2 ratio of the number of non-target sites (allele frequency <0.1) between maternal and zygotic transcripts. Positive values indicate that targets for a specific microRNA tend to be avoided by maternal transcripts. Figure 5A is a bar plot of target avoidance log2 ratios for different levels of microRNA abundance in the egg. Maternally deposited coding transcripts tend to avoid some target sites for highly abundant maternal microRNAs (with respect to zygotic transcripts). Differences were statistically significant (Figure 5A). In a similar manner I defined a target conservation measure as the log2 ratio of the number of target sites (allele frequency >0.9) between maternal and zygotic transcripts. A positive value indicates that target-sites are preferentially conserved in maternal transcripts. In Figure 5B I plot these values for different microRNA abundances. Overall, maternal transcripts conserve some target sites, but there is no distinctive pattern between maternal and non-maternal microRNAs (Figure 5B). To further validate the observed target avoidance, we can compare the whole allele frequency distribution and evaluate whether the distribution for maternal transcripts is shifted to the left with respect to that of zygotic transcripts (as suggested in Figure 3C, right panel). Figure 5C plots the p-value for a one-tail Kolmogorov-Smirnov test, the lowest the p-value, the more shifted to the left, that is, the more evidence for target avoidance. For highly abundant maternal microRNAs, the allele frequency distribution of target/non-target sites is shifted to the left, that is, there is a preference for the non-target allele. From these analyses I conclude that maternal transcripts, compared to non-maternal zygotic transcripts, have a mild yet significant tendency to avoid maternal microRNAs target sites.

DISCUSSION

This study characterizes microRNA products from *Drosophila* unfertilized eggs. I validated two of these microRNAs by qPCR, and also show that they are not in eggs from mother lacking Dicer-1

function in the germline. This demonstrates that there are microRNAs deposited by the mother into the *Drosophila* developing egg. The presence of microRNAs in unfertilized oocytes have been described in mice [7]. However, it has been shown that microRNA activity is suppressed in mice oocytes, indicating that maternally deposited microRNAs may not have a defined function in this species [31,32]. Here I show evidence for *Drosophila* maternal microRNA activity as they have an impact in the evolution of potential target sites in maternal microRNAs (Figures 3, 4 and 5).

One of the most abundant maternal microRNAs, mir-184, has been already described in freshly laid *Drosophila* eggs [33]. Also, mir-184 has important roles during oocyte development as well as in early development [33]. On the other hand, the conserved microRNA mir-34 has been also described as a maternal microRNAs [34] but it has only one read copy in our dataset, and it has not been detected in another independent high-throughput screen [16]. This may suggest that either mir-34 is a very low copy maternal microRNA, or that it is rapidly degraded after egg deposition/activation.

Another maternal microRNA, mir-9c, has been shown to be crucial to regulate the number of germ cells [35]. Indeed, is the maternal loss of mir-9c what produces this phenotype [35]. This microRNA is hosted within a maternally deposited gene, *grapes* (Table 2). Here I show that mir-9c targets more unstable transcript during the MZT than expected by chance (Table 3), which indicates that mir-9c may have a role during maternal transcript clearance during the initial steps of development. A similar role has been described for zygotically transcribed microRNAs [25].

Other maternally deposited microRNAs derive from the mir-310/mir-313 cluster. This cluster is highly conserved in the *Drosophila* lineage [36], although it may have originated in insects [37], and is evolutionarily related with the (also maternal) mir-92a/mir-92b cluster [37,38]. Both the orthologous mir-310/311/312/313 and mir-92a/92b clusters in *Drosophila virilis* has been detected at high levels during the first two hours of development, suggesting that this cluster is also maternally deposited in this species (Supplementary Table 2 in [37]). Interestingly, some maternal

microRNAs have other functions later on during development. MicroRNAs from the mir-310/311/312/313 cluster are known to be involved in male gonads development [39]. Recently, Ranz and collaborators found that mir-310/mir-313 microRNAs show male biased expression pattern at the onset of metamorphosis [40]. On the other hand, mir-92a is expressed in the adult, and it is involved in leg morphology [41]. Some other maternal microRNAs have roles unrelated with embryonic development such as mir-14, which regulates insulin production [42]; mir-279, involved in the circadian clock [43]; or mir-8, associated to abdominal pigmentation [44], to name but a few cases. Altogether, these examples show that maternal microRNAs frequently have other functions in different developmental stages and/or tissues.

MicroRNA target avoidance has been observed in *Drosophila* [45], as well as in mice [46] and humans [47]. Here I observe a similar pattern in *Drosophila* eggs, in which maternal transcripts tend to avoid target sites for maternal microRNAs. Comparative genomics and population genetic data independently suggest that maternal transcripts tend to avoid being targeted by maternal microRNAs. This is the first time that microRNA target avoidance has been shown at the population level. The main advantage of working with microRNAs to study evolution at the population level is that we can predict the impact of single point mutations in both microRNAs and their targets, something that is not yet possible with other regulators (such as transcription factors). Further analyses, including the theoretical expectations of the proposed target-site-mutation model, will shed light on how microRNA function diversify and, more generally, how gene regulation evolves.

Alternatively, a lower number of target sites in maternal transcripts could be a bias maternal transcripts with conserved target sites may be degraded early, and therefore not detected in early embryos. However, in *Drosophila* microRNA-mediated transcript degradation happens a few hours after microRNA-mediated repression [48]. Maternal transcripts are detected from 0-2 hour old embryos, and they are unlikely to have had microRNA-mediated transcript degradation. The microRNAs studied in that paper were: mir-9b, mir-279 and bantam, all of which were detected in

this study as maternal.

Overall, this paper describes three features of maternally transmitted microRNAs: 1) they are often produced from introns of maternally deposited transcripts; 2) they can be zygotically transcribed and have other functions during development; 3) maternal transcripts tend to avoid target sites for maternal microRNAs; 4) mir-9c may be involved in maternal transcript clearance during MZT. These observations suggest that some maternal microRNAs may have a function during development, whilst other are potentially damaging to the normal function, and selective pressures favour maternal transcripts to avoid being targeted by them.

MATERIALS AND METHODS

Flies and egg collection

Fly stocks used in this study, with Bloomington reference number in square brackets, were: w^{1118} [#3605], w^{*}; P{neoFRT}82B P{ovoD1-18}3R/st^{1} βTub85D^{D} ss^{1} e^{s}/TM3, Sb^{1} [#2149] [49] and y^{d2} w^{1118} P{ey-FLP.N}2; P{neoFRT}82B Dcr-1^{Q1147X}/TM3, Sb^{1} [#32066] [14]. All flies were kept at 25ºC on cornmeal based media, with 12 hours light/dark cycles. Virgin females were sorted at the pupae stage to avoid any unwanted fertilization. (Previous attempts selecting for <6 hours females produced a small yet significant number of fertilized eggs). In a population cage I let 80-100 females to lay eggs in apple juice agar plates for 8 hours, collecting 1 hour after dawn. Eggs were collected with a sieve and washed with saline solution. Eggs from virgin females do not degenerate even several hours after laying [50].

To generate *Drosophila* female flies with no Dicer-1 activity in the germline, I first crossed P{neoFRT}82B ovoD1-18/TM3 males to P{neoFRT}82B Dcr-1^{Q1147X}/TM3 females. Resulting offspring at L2-L3 larvae stage was heat-shocked (37ºC) for two hours in two consecutive days to induce FRT/FLT mitotic recombination [49,51]. As ovoD1-18 is dominant and impairs oocyte

production, only females with Dcr-1^{Q1147X}/TM3 and Dcr-1^{Q1147X}/Dcr-1^{Q1147X} germline clones will produce eggs. Eggs laid by germline homozygotes *Dicer-1* mutant females were used as a negative control for maternally produced microRNAs in qPCR experiments.

RNA extraction, sequencing and profiling

Total RNA was extracted from eggs or early embryos with TRIzol reagent (Life Technologies), following instructions given by the manufacturer, and dissolving the RNA in RNase-free water. For RNA sequencing, cDNA libraries were generated with TruSeq Small RNA Sample Preparation Kit (Illumina). Amplified cDNA constructs were size selected in a 6% PAGE gel for 145 to 160 bp (fragments including RNA-derived sequences of size ~20-30 bp plus adapters). Size selected cDNAs were purified and precipitated with ethanol, and DNA integrity was checked with TapeStation (Agilent). Samples were sequenced with Illumina MiSeq in the Genomics Core Facility at the University of Manchester. A total of 4,507,291 reads were sequenced, most of them (95.5%) deriving from ribosomal RNAs which is expected in *Drosophila* where the majority of small RNAs are 2S rRNA [52]. 13,114 reads were identified as microRNA products. Sequence reads are available from Gene Expression Omnibus (GEO) at NCBI (accession number: GSE63488).

Illumina MiSeq produces 50 bp sequence reads. Hence, I removed adapters with Cutadapt (https://code.google.com/p/cutadapt/) and mapped the processed reads of size 18-26 bp to known microRNAs from miRBase v.20 [53], using Bowtie v.0.12.7 [54], allowing no mismatches and considering reads mapping to up to five positions. Additionally, reads were mapped to the *D. melanogaster* reference genome to detect potentially novel microRNAs as previously described [19,20], but none was found. Other RNA collections from embryos and ovaries were also analysed: 0-1h embryos, 2-6h embryos, 6-10 h embryos [55] and ovaries [56]. Expression profiling in Figure 1 was done with R, scaling the Z-scores of the heatmap across rows, and generating a hierarchical tree of microRNAs with complete linkage clustering [57].

Validations were made with TaqMan assay manufactured by Life Technologies. MicroRNA TaqMan probes were for mir-311-3p (5'-TATTGCACATTCACCGGCCTGA-3') and mir-995-3p (5'-TAGCACCACATGATTCGGCTT-3'), and the relative expression levels were measured against 2S rRNA. Fluorescent quantification was done in a LightCycler 96 Real-Time PCR System (Roche) for 40 cycles, and Cts were calculated using the software provided by the manufacturer with default parameters. Relative expression values in Figure 2 were calculated as the ratio (40 – Q)/(40 – R) were R was the average Ct value for 2S rRNA and Q the Ct value for the microRNA (the 40 was introduced to measure the expression levels relative to the 40 amplification cycles).

MicroRNA target analysis and polymorphisms

Target analysis was based on the presence of canonical seeds in the transcripts [28]. Canonical seed predictions have the advantage that only primary sequence information is used, so populations models (see below) can be easily fitted. Maternally deposited gene transcripts are listed in the Berkeley Drosophila Genome Project webpage at http://insitu.fruitfly.org [22,23]. Which transcripts are destabilized during the maternal-to-zygotic transition were identified from Tadros et al. microarray experiments (Gene Expression Omnibus accession number GSE13287), detecting probes with a >1.5 fold change in their expression level between 4-6 h embryos and oocytes [27]. To assess whether maternal microRNAs target transcripts that are destabilized during the MTZ transition I calculated the proportion of unstable transcripts targeted by each microRNA and compared it to the expected proportion (0.146) with a cumulative binomial test. False Discovery Rate was accounted by calculating q-values associated to the p-values [58,59].

For the population analyses, I first mapped the single-nucleotide polymorphisms (SNPs) from the Drosophila Genetic Reference Panel [60,61], available at http://dgrp2.gnets.ncsu.edu/, against the 3'UTR of *Drosophila melanogaster* release 5.13 (http://flybase.org). For each microRNA I defined a target sequence (sixmer) and its 18 non-target neighbours, that is, the 18 one-

nucleotide variations of the target site (Figure 3A). Every SNP that connects a target with a non-target sixmer was further considered. 3'UTRs with introns were discarded. For each polymorphic target site, the allele frequency distribution was calculated as the proportion of the target allele with respect to the total number of sampled individuals (isogenic lines).

ACKNOWLEDGEMENTS

I thank Maria Ninova and Fran Bonath for sharing their expertise on small RNA library preparations, and to Matt Ronshaugen for helpful advice on fly genetics. Greg Brooke, Elena Klenova and Adele Angel hosted me in their lab. I am also very grateful to Maria Ninova and Sam Griffiths-Jones form critical reading of the manuscript, and to my colleagues at the Junior European Drosophila Investigators (JEDI) network for useful feedback. The MiSeq sequencing was funded my the Career Developmental Award (University of Manchester). TaqMan assays and publication costs were funded by the University of Essex.

REFERENCES

1. Lawrence PA (1992) The Making of a Fly: The Genetics of Animal Design. Blackwell Scientific. 228 p.

2. Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, et al. (2003) The small RNA profile during Drosophila melanogaster development. Dev Cell 5: 337–350.

3. Aboobaker AA, Tomancak P, Patel N, Rubin GM, Lai EC (2005) Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. Proc Natl Acad Sci U S A 102: 18017–18022. doi:0508823102.

4. Enright A, John B, Gaul U, Tuschl T, Sander C, et al. (2003) MicroRNA targets in Drosophila. Genome Biol 5: R1. doi:10.1186/gb-2003-5-1-r1.

5. Lai EC, Tomancak P, Williams RW, Rubin GM (2003) Computational identification of Drosophila microRNA genes. Genome Biol 4: R42. doi:10.1186/gb-2003-4-7-r42.

6. Marco A (2012) Regulatory RNAs in the light of Drosophila genomics. Brief Funct Genomics 11: 356–365. doi:10.1093/bfgp/els033.

7. Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, et al. (2007) Maternal microRNAs are

essential for mouse zygotic development. Genes Dev 21: 644–648. doi:10.1101/gad.418707.

8. Tesfaye D, Worku D, Rings F, Phatsara C, Tholen E, et al. (2009) Identification and expression profiling of microRNAs during bovine oocyte maturation using heterologous approach. Mol Reprod Dev 76: 665–677. doi:10.1002/mrd.21005.

9. Chen PY, Manninga H, Slanchev K, Chien M, Russo JJ, et al. (2005) The developmental miRNA profiles of zebrafish as determined by small RNA cloning. Genes Dev 19: 1288–1293. doi:10.1101/gad.1310605.

10. Watanabe T, Takeda A, Mise K, Okuno T, Suzuki T, et al. (2005) Stage-specific expression of microRNAs during Xenopus development. FEBS Lett 579: 318–324. doi:10.1016/j.febslet.2004.11.067.

11. Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, et al. (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res 17: 1850–1864. doi:10.1101/gr.6597907.

12. Robinson SW, Herzyk P, Dow JAT, Leader DP (2013) FlyAtlas: database of gene expression in the tissues of Drosophila melanogaster. Nucleic Acids Res 41: D744–D750. doi:10.1093/nar/gks1141.

13. Nakahara K, Kim K, Sciulli C, Dowd SR, Minden JS, et al. (2005) Targets of microRNA regulation in the Drosophila oocyte proteome. Proc Natl Acad Sci U S A 102: 12023–12028. doi:10.1073/pnas.0500053102.

14. Lee YS, Nakahara K, Pham JW, Kim K, He Z, et al. (2004) Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways. Cell 117: 69–81. doi:10.1016/S0092-8674(04)00261-2.

15. Votruba SM (2009) MicroRNAS in the Drosophila Egg and Early Embryo. University of Toronto. 152 p.

16. Lee M, Choi Y, Kim K, Jin H, Lim J, et al. (2014) Adenylation of Maternally Inherited MicroRNAs by Wispy. Mol Cell 56: 696–707. doi:10.1016/j.molcel.2014.10.011.

17. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, et al. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322: 1387–1392. doi:10.1126/science.1165171.

18. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, et al. (2009) Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137: 522–535. doi:10.1016/j.cell.2009.03.040.

19. Marco A, Griffiths-Jones S (2012) Detection of microRNAs in color space. Bioinformatics 28: 318–323. doi:10.1093/bioinformatics/btr686.

20. Marco A, Kozomara A, Hui JHL, Emery AM, Rollinson D, et al. (2013) Sex-Biased Expression of MicroRNAs in Schistosoma mansoni. PLoS Negl Trop Dis 7: e2402. doi:10.1371/journal.pntd.0002402.

21. Marco A (2014) Sex-biased expression of microRNAs in Drosophila melanogaster. Open Biol

4: 140024. doi:10.1098/rsob.140024.

22. Tomancak P, Beaton A, Weiszmann R, Kwan E, Shu S, et al. (2002) Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol 3: RESEARCH0088.

23. Tomancak P, Berman BP, Beaton A, Weiszmann R, Kwan E, et al. (2007) Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biol 8: R145.

24. Fogarty P, Campbell SD, Abu-Shumays R, de Saint Phalle B, Yu KR, et al. (1997) The Drosophila grapes gene is related to checkpoint gene chk-1 rad27 and is required for late syncytial division fidelity. Curr Biol 7: 418–426.

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

26. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, et al. (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312: 75–79. doi:10.1126/science.1122689.

27. Tadros W, Goldman AL, Babak T, Menzies F, Vardy L, et al. (2007) SMAUG Is a Major Regulator of Maternal mRNA Destabilization in Drosophila and Its Translation Is Activated by the PAN GU Kinase. Dev Cell 12: 143–155. doi:10.1016/j.devcel.2006.10.005.

28. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215–233. doi:10.1016/j.cell.2009.01.002.

29. Crow JF, Kimura M (1970) An Introduction to Population Genetics Theory. Harper & Row. 591 p.

30. Nei M (1975) Molecular Population Genetics and Evolution. North-Holland. 288 p.

31. Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, et al. (2010) MicroRNA function is globally suppressed in mouse oocytes and early embryos. Curr Biol CB 20: 271–277. doi:10.1016/j.cub.2009.12.044.

32. Ma J, Flemr M, Stein P, Berninger P, Malik R, et al. (2010) MicroRNA activity is suppressed in mouse oocytes. Curr Biol CB 20: 265–270. doi:10.1016/j.cub.2009.12.042.

33. Iovino N, Pane A, Gaul U (2009) miR-184 Has Multiple Roles in Drosophila Female Germline Development. Dev Cell 17: 123–133. doi:10.1016/j.devcel.2009.06.008.

34. Soni K, Choudhary A, Patowary A, Singh AR, Bhatia S, et al. (2013) miR-34 is maternally inherited in Drosophila melanogaster and Danio rerio. Nucleic Acids Res 41: 4470–4480. doi:10.1093/nar/gkt139.

35. Kugler J-M, Chen Y-W, Weng R, Cohen SM (2013) Maternal Loss of miRNAs Leads to Increased Variance in Primordial Germ Cell Numbers in Drosophila melanogaster. G3 Bethesda Md. doi:10.1534/g3.113.007591.

36. Marco A, Ninova M, Ronshaugen M, Griffiths-Jones S (2013) Clusters of microRNAs emerge

by new hairpins in existing transcripts. Nucleic Acids Res 41: 7745–7752. doi:10.1093/nar/gkt534.

37. Ninova M, Ronshaugen M, Griffiths-Jones S (2014) Fast-evolving microRNAs are highly expressed in the early embryo of Drosophila virilis. RNA 20: 360–372. doi:10.1261/rna.041657.113.

38. Lu J, Fu Y, Kumar S, Shen Y, Zeng K, et al. (2008) Adaptive evolution of newly emerged micro-RNA genes in Drosophila. Mol Biol Evol 25: 929–938. doi:10.1093/molbev/msn040.

39. Pancratov R, Peng F, Smibert P, Yang J-S, Olson ER, et al. (2013) The miR-310/13 cluster antagonizes β-catenin function in the regulation of germ and somatic cell differentiation in the Drosophila testis. Dev Camb Engl 140: 2904–2916. doi:10.1242/dev.092817.

40. Yeh S-D, von Grotthuss M, Gandasetiawan KA, Jayasekera S, Xia X-Q, et al. (2014) Functional Divergence of the MiRNA Transcriptome at the Onset of Drosophila Metamorphosis. Mol Biol Evol. doi:10.1093/molbev/msu195.

41. Arif S, Murat S, Almudi I, Nunes MDS, Bortolamiol-Becet D, et al. (2013) Evolution of mir-92a Underlies Natural Morphological Variation in Drosophila melanogaster. Curr Biol 23: 523–528. doi:10.1016/j.cub.2013.02.018.

42. Xu P, Vernooy SY, Guo M, Hay BA (2003) The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. Curr Biol CB 13: 790–795.

43. Luo W, Sehgal A (2012) microRNA-279 acts through the JAK/STAT pathway to regulate circadian behavioral output in Drosophila. Cell 148: 765–779. doi:10.1016/j.cell.2011.12.024.

44. Kennell JA, Cadigan KM, Shakhmantsir I, Waldron EJ (2012) The microRNA miR-8 is a positive regulator of pigmentation and eclosion in Drosophila. Dev Dyn Off Publ Am Assoc Anat 241: 161–168. doi:10.1002/dvdy.23705.

45. Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. Cell 123: 1133–1146. doi:10.1016/j.cell.2005.11.023.

46. Farh KK-H, Grimson A, Jan C, Lewis BP, Johnston WK, et al. (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. Science 310: 1817–1821. doi:10.1126/science.1121158.

47. Sood P, Krek A, Zavolan M, Macino G, Rajewsky N (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. Proc Natl Acad Sci U S A 103: 2746–2751. doi:10.1073/pnas.0511045103.

48. Djuranovic S, Nahvi A, Green R (2012) miRNA-Mediated Gene Silencing by Translational Repression Followed by mRNA Deadenylation and Decay. Science 336: 237–240. doi:10.1126/science.1215691.

49. Chou TB, Perrimon N (1996) The autosomal FLP-DFS technique for generating germline mosaics in Drosophila melanogaster. Genetics 144: 1673–1679.

50. Tsien HC, Wattiaux JM (1971) Effect of Maternal Age on DNA and RNA Content of

Drosophila Eggs. Nature 230: 147–148. doi:10.1038/10.1038/newbio230147a0.

51. Harrison DA, Perrimon N (1993) Simple and efficient generation of marked clones in Drosophila. Curr Biol 3: 424–433. doi:10.1016/0960-9822(93)90349-S.

52. Seitz H, Ghildiyal M, Zamore PD (2008) Argonaute Loading Improves the 5′ Precision of Both MicroRNAs and Their miRNA∗ Strands in Flies. Curr Biol 18: 147–151. doi:10.1016/j.cub.2007.12.049.

53. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42: D68–D73. doi:10.1093/nar/gkt1181.

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

55. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, et al. (2006) Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in C. elegans. Cell 127: 1193–1207. doi:10.1016/j.cell.2006.10.040.

56. Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, et al. (2008) An endogenous small interfering RNA pathway in Drosophila. Nature 453: 798–802. doi:10.1038/nature07007.

57. Sneath PHA, Sokal RR (1973) Numerical Taxonomy: The Principles and Practice of Numerical Classification. W. H. Freeman. 573 p.

58. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 57: 289–300. doi:10.2307/2346101.

59. Storey JD (2002) A direct approach to false discovery rates. J R Stat Soc Ser B Stat Methodol 64: 479–498. doi:10.1111/1467-9868.00346.

60. Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, et al. (2012) The Drosophila melanogaster Genetic Reference Panel. Nature 482: 173–178. doi:10.1038/nature10811.

61. Huang W, Massouras A, Inoue Y, Peiffer J, Rámia M, et al. (2014) Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. Genome Res. doi:10.1101/gr.171546.113.

Table 1. Maternal microRNAs in *Drosophila melanogaster*.

MicroRNA transcript*	**Reads per miRNA**	**Total reads (%)****
mir-310/311/312/313	356/2012/1661/82	4111 (31.3)
mir-92a/92b	172/2109	2281 (17.4)
mir-184	1377	1377 (10.5)
mir-9c/306/79/9b	1064/154/4/132	1354 (10.3)
bantam	1204	1204 (9.2)
mir-995	624	624 (4.8)
mir-14	411	411 (3.1)
mir-275/305	90/269	359 (2.7)
mir-998/11	5/205	210 (1.6)
mir-8	209	209 (1.6)
mir-2b-2/2a-1/2a-2	85/25/19	129 (1.0)
mir-279/996	61/56	117 (0.9)
mir-2b-1	92	92 (0.7)
mir-281-2/281-1	66/13	79 (0.6)
mir-4969/999	0/69	69 (0.5)
mir-33	62	62 (0.5)
mir-263a	59	59 (0.4)
mir-10	28	28 (0.2)
mir-2c/13a/13b-1	0/0/27	27 (0.2)
mir-13b-2	26	26 (0.2)
mir-970	26	26 (0.2)
mir-1012	25	25 (0.2)
mir-31a	23	23 (0.2)
mir-9a	21	21 (0.2)
mir-309/3/286/4/5/6-1/6-2/6-3	0/1/15/1/1/1/1/1	21 (0.1)
mir-956	20	20 (0.2)
mir-276a	18	18 (0.1)
mir-994/318	2/14	16 (0.1)
mir-1010	14	14 (0.1)

* MicroRNAs clustered in the genome (<10kb).
** Percentage over total number of reads mapping to microRNAs.

Table 2. Maternal microRNA loci within protein-coding genes.

MicroRNA cluster	Host gene	Protein-coding gene maternal?
mir-995	cdc2c	yes
mir-11/998	E2f	yes
mir-92a	jigr1	yes
mir-999	CASK	yes
mir-281-1/281-2	Oda	yes
mir-970	Tomosyn	yes
mir-2b-2/2a-1/2a-2	spi	yes
mir-13b-2	CG7033	yes
mir-9c/306/79/9b	grp	yes*
mir-33	HLH106	no expression information
mir-1012	Lerp	no expression information
mir-1010	SKIP	no

* Detected in the oocyte.

Table 3. Maternal microRNAs targeting unstable transcripts during maternal-to-zygotic transition

MicroRNA	Unstable targets	Stable targets	Proportion unstable transcripts*	q-value
mir-283-5p	179	805	0.182	0.018
mir-277-3p	116	497	0.189	0.026
mir-9a-5p	50	166	0.232	0.036
mir-9b-5p	50	166	0.232	0.036
mir-9c-5p	50	166	0.232	0.036
mir-92a-3p	74	313	0.191	0.096
mir-92b-3p	74	313	0.191	0.096
mir-310-3p	74	313	0.191	0.096
mir-311-3p	74	313	0.191	0.096
mir-312-3p	74	313	0.191	0.096
mir-313-3p	74	313	0.191	0.096

* Expected proportion is 0.146

Figure legends

Figure 1. Expression profile of maternal microRNAs in *Drosophila melanogaster*. Hierarchical clustering of microRNAs detected in unfertilized eggs. Relative expression values were normalized across columns. The tree is split into three categories of microRNAs: those with a high abundance in ovaries compare to the other stages (high in ovaries); those which are mainly present in the unfertilized eggs (high in egg) and those which have a higher expression level later during development (high in zygote).

Figure 2. Quantitative PCR of maternal microRNA products and evolutionary conservation of targets. (A) Relative expression levels of mir-311 and mir-995 with TaqMan assays (see Methods). Levels were measured for unfertilized eggs, early embryos (2-6h) and for eggs from females with no Dicer-1 activity in the germ-line ($dcr1^{-}$). Asterisk shows that differences between eggs and 2-6h expression levels of mir-311 were statistically significant ($p<0.01$). (B) Proportion of maternal (dark grey) and zygotic (light grey) transcripts targeted by maternal microRNAs. Maternal and zygotic transcripts are sorted by evolutionary conservation. Asterisk shows a significant difference between *D. melanogaster*-specific maternal and zygotic transcripts targeted by maternal microRNAs.

Figure 3. Polymorphic target sites in *Drosophila* populations. (A) Each microRNA sixmer target site has 18 one-nucleotide mutant neighbours which are themselves not target sites. (B) The allele frequency for each pair of target/non-target site is calculated as the proportion of target site alleles with respect to the total number of alleles in the pair. (C) Allele frequency distribution in a finite population is U-shaped for pairs of alleles neutral to each other. If there is selection favouring target sites, distributions are expected to be shifted to the right. Conversely, if there is selection against target sites, distributions will be shifted to the left. (D) Allele frequency distribution for target sites

for maternal microRNAs in maternal (dark grey) and non-maternal zygotic (light grey) transcripts. Left and right panel shows the distributions for 5' and 3' arms of mir-995 respectively. Grey box in left panel indicates that mir-995-5p is virtually absent from the unfertilized egg.

Figure 4. Allele frequency distribution for target sites of maternal microRNAs. Distribution of targets for maternal and zygotic transcripts for mir-305 (both arms are highly present in the egg) and mir- 4986 (grey box, neither of the arms was detected in the egg).

Figure 5. Maternal microRNA target avoidance. (A) Target avoidance ratio (see main text for details) for microRNAs with differences abundances in the unfertilized egg (NULL – not detected, LOW – less than 0.1% of the set, MID – between 0.1% and 1%, HIGH – more than 1%). Error bars represent the Standard Error of the Mean. Asterisks show statistically significant differences ($p<0.01$) for t-test with unequal variances. (B) Target conservation ratio. (C) Distribution shifting to the left (see main text for details) of allele frequency distribution in maternal transcripts with respect to non-maternal zygotic transcripts.

Figure 1

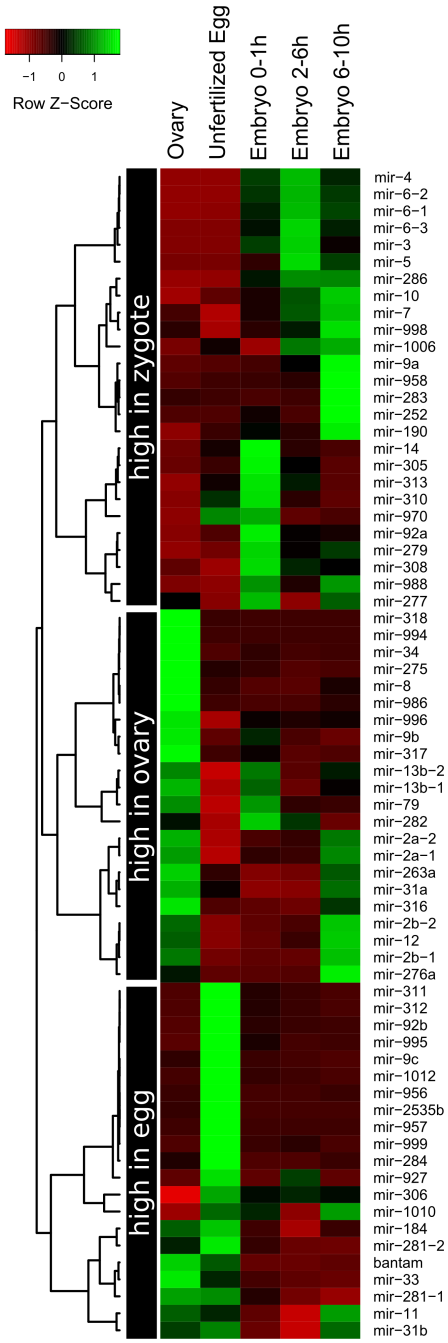

Figure 2

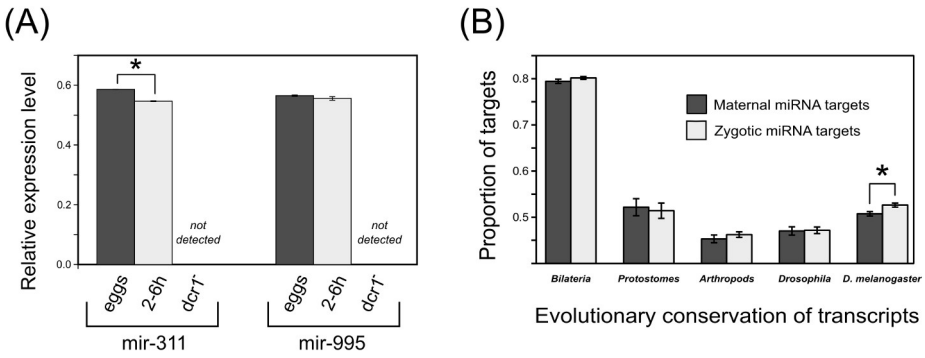

Figure 3

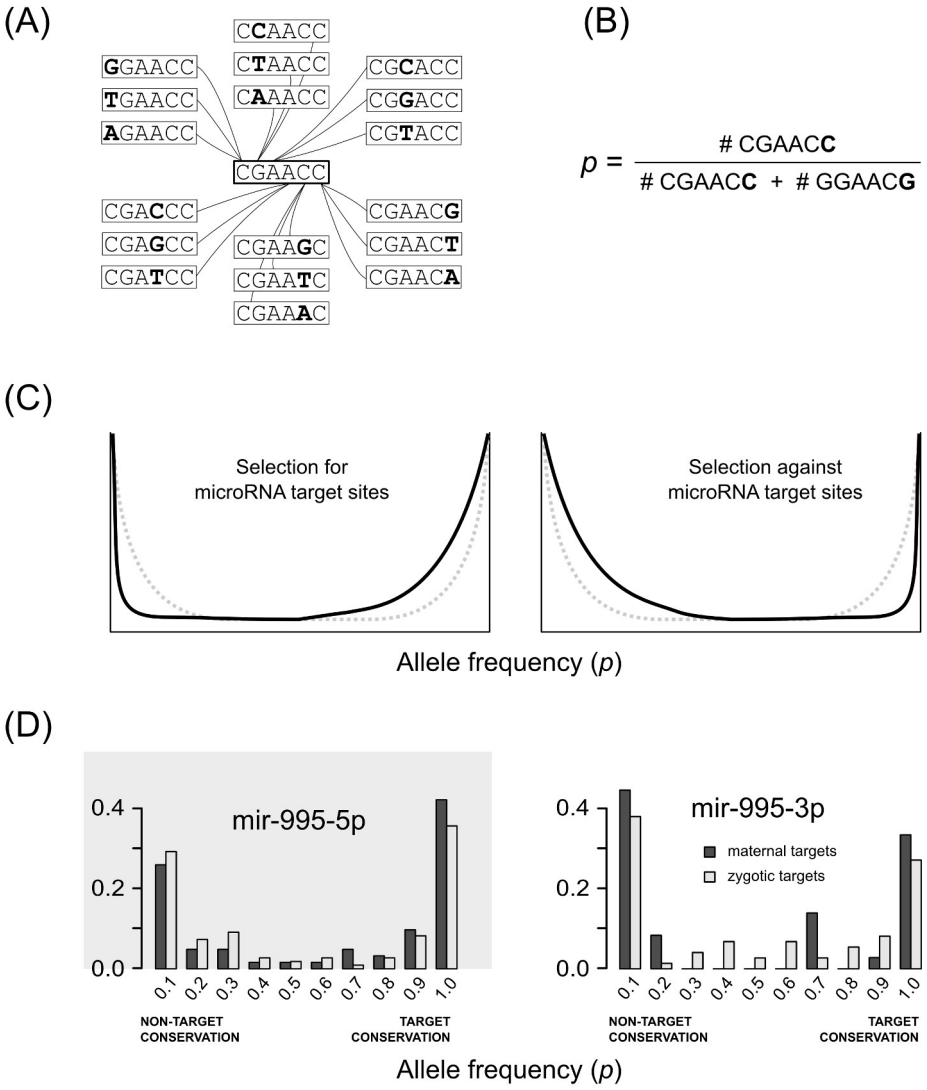

Figure 4

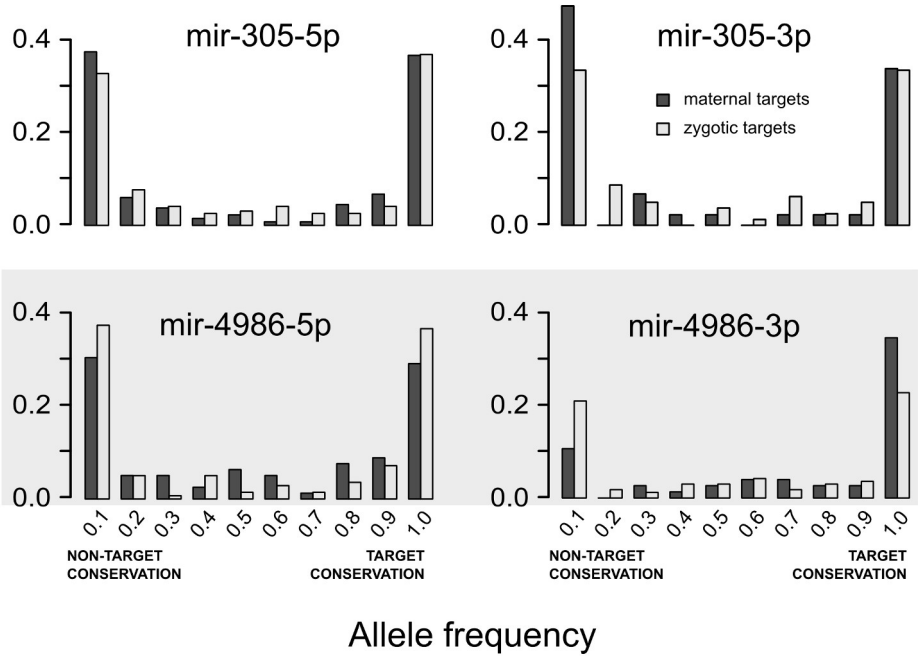

Figure 5