Supplementary Text:

Van Valen (1973) first proposed the Red Queen hypothesis (RQH) over 40 years ago. In his original famous paper [1], he stated: ”The Red Queen does not need changes in the physical environment, although she can accommodate them. Biotic forces provide the basis for a self-driving (at this level) perpetual motion of the effective environment and so of the evolution of the species affected by it.” This hypothesis revised the idea that organisms are only adapted to abiotic environment, and emphasized biotic factors are also important in driving evolutionary changes. After that, RQH has been applied to explain many biotic interactions, including co-evolution of parasite and host, prey and predator, sexual antagonism and even the evolution of sex [2-7].

Although RQH is often used to explain arm race between or within species, its application could be much broader because the essence of this metaphor is continual adaptation. Just like Red Queen advised Alice to keep running for staying in the same place, species also has to run to keep pace with a moving world. Previous adaptations do no assure the safety of a species when it is facing new challenges from environment (both biotic and abiotic). When in the adaptive landscape, newly emerged species/traits under selective pressure might evolve continually or go extinction. The same logic can also be applied to genic level. Many new genes emerged adaptively, because they have fitness advantages and outcompete “the absence of this gene” at the beginning. However, gene content remains largely constant during long period of evolution. This contradiction implies many of these new materials face elimination soon after their emergence. More importantly, it also implies the adaption is often transient rather than permanent. In this study, we test the Red Queen hypothesis by studying the evolution of *de novo* genes, which are indeed new when compared with those duplicated genes. If the adaptive landscape is as shifty as posited by the Red Queen hypothesis, we expect that these *de novo* genes will keep evolving and have different evolutionary fates in

different species. In addition, we also discuss what might be the driven force for their continual evolution.

References:

1. Van Valen, L. *A New Evolutionary Law*, 1—30 (1973).
2. Morran, L.T., Schmidt, O.G., Gelarden, I.A., Parrish, R.C., 2nd & Lively, C.M. Running with the Red Queen: host-parasite coevolution selects for biparental sex. *Science* **333**, 216-8 (2011).
3. Brockhurst, M.A. Evolution. Sex, death, and the Red Queen. *Science* **333**, 166-7 (2011).
4. King, K.C., Delph, L.F., Jokela, J. & Lively, C.M. The geographic mosaic of sex and the Red Queen. *Curr Biol* **19**, 1438-41 (2009).
5. Ebert, D. Host-parasite coevolution: Insights from the Daphnia-parasite model system. *Curr Opin Microbiol* **11**, 290-301 (2008).
6. Brockhurst, M.A. *et al.* Running with the Red Queen: the role of biotic conflicts in evolution. *Proc Biol Sci* **281**(2014).
7. Liow, L.H., Van Valen, L. & Stenseth, N.C. Red Queen: from populations to taxa and communities. *Trends Ecol Evol* **26**, 349-58 (2011).

Supplementary Figure Legends:

Figure S1 miRNA expressions in *Drosophila melanogaster* miR-972 and miR-982 cluster

All mature miRNAs from the two clusters are plotted. Major miRs from the six *de novo* miRNAs are labeled. RPM, reads per million.

Figure S2 miRNA mutant sequence, hairpin structures and mutant fly constructions in *D. melanogaster* and *D. simulans*

(A) *dme-mir-982-1/-2* mutant sequence. *dme-miR-983-5p* and *dme-miR-983-3p* are shaded in yellow and gray, respectively.

(B) Hairpin structure of *dme-mir-984* KO sequence is less stable than that of the wild type sequence. MFE, minimum free energy.

(C) Hairpin structure of *dsi-mir-983a* KO sequences are less stable than that of the wild type sequence. MFE, minimum free energy.

(D) Cross scheme for miR-983/miR-984 mutant flies generation in *D. melanogaster*. Flies in red box are used for mutation detection or sanger sequencing (methods labeled on the right) after crossing with corresponding females. After embryo injection, each emerged fly of F_0 was used for mutation detection following SURVEYOR® Mutation Detection Kit manufactory protocol. If mutation was detected in F_0, 20 of its F_1 virgin females were chosen, and each of them was crossed with FM7c balancer male. F_2 males were used for sanger sequencing to detect and confirm the mutant sequences. Fly containing mutation was crossed with its sibling female, making the mutation locus homozygous. Flies without mutation were served as control for downstream assays.

(E) Same as above except that F_0 emerged flies are females.

(F) Cross scheme for miR-983 mutant flies generation in *D. simulans*. Procedures are similar with steps shown in Fig. S2D. Note that F_3 flies were used for extra round of sanger sequencing, which ensures that homozygous mutant flies were chosen.

(G) Same as above except that F_0 emerged flies are females.

Figure S3 miRNA mutant phenotypes

(A) – (B) Phase contrast for testes of *dme-mir-984* KO fly. All the sequential cell types are present and normal.

(C) – (D) DAPI staining for testes of *dme-mir-984* mutant fly. * indicates the apical end of testis, mitotic cells are normal. Arrows indicate the normal needle shaped spermatid nuclei. T, testis; SV, seminal vesicle; AG, accessory gland.

(E) *dme-miR-983-5p* is significantly down-regulated after AMO injection. Student's t test, $P < 0.01$.

(F) – (G) Cross scheme for introducing wild type fly autosomes into miR-983 mutant or control flies. Oregon-R is used as typical wild type fly. + indicates the autosomes in the original background, while * indicates autosomes from Oregon-R.

Figure S4 miR-983 regulation in *D. melanogaster* and *D. simulans*

(A) miR-983-5p regulation strength. dme non-targets: non-targets in *D. melanogaster*, dme targets：*dme-miR-983-5p* targets in *D. melanogaster*; dsi non-targets: non-targets in *D. simulans*，dsi targets：*dsi-miR-983a-5p* targets in *D. simulans*. The small inset shows the whole distribution. Red dashed lines indicate the median expression fold changes of non-targets in the two species, respectively.

(B) Empirical cumulative distribution of log2(gene expression fold change) in *D. melanogaster*. Red and black lines represent PITA predicted *dme-miR-983-5p* targets and non-targets, respectively. Kolmogorov-Smirnov test, $P = 0.2754$.

(C) Empirical cumulative distribution of log2(gene expression fold change) in *D. simulans*. Red and black lines represent PITA predicted *dsi-miR-983a-5p* targets and non-targets, respectively. Kolmogorov-Smirnov test, $P < 0.001$.

(D) Mis-regulated genes induced by miR-983 KO in *D. melanogaster* and *D. simulans* share small overlap.

(E) Genes expressed in *D. melanogaster* and *D. simulans* testes share large overlap (~90%).

(F) Gene expressions in *D. melanogaster* and *D. simulans* testes correlate well. Pearson's product-moment correlation, r = 0.91.

Figure S1

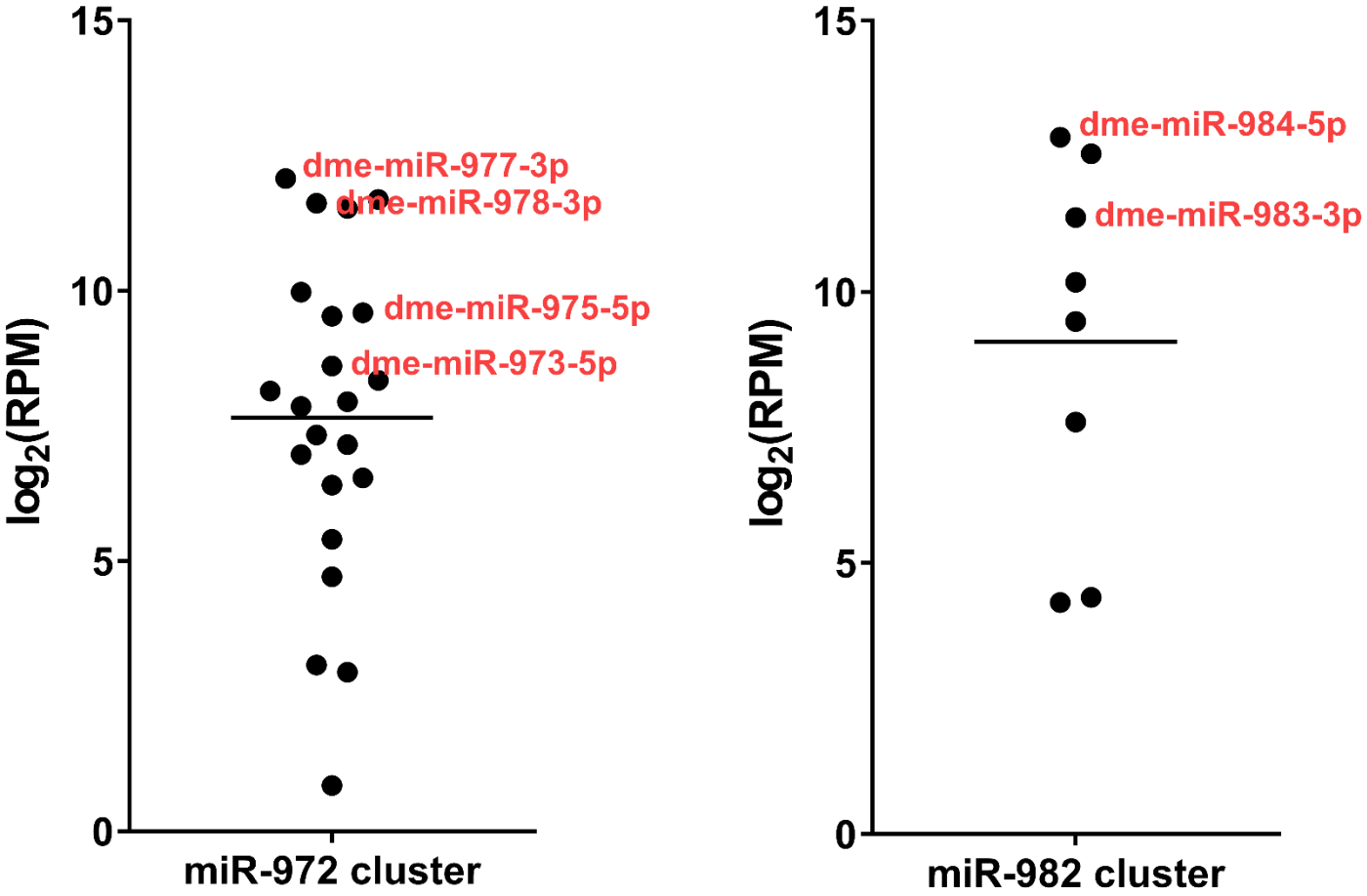

15
10
5
0
log2(RPM)
dme-miR-977-3p
dme-miR-978-3p
dme-miR-975-5p
dme-miR-973-5p
miR-972 cluster
15
10
5
0
log2(RPM)
dme-miR-984-5p
dme-miR-983-3p
miR-982 cluster

Figure S2

A

***dme-mir-983-1* and *dme-mir-983-2*:**

```
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                        5         15         25         35         45         55
Wild type           TATTATATTG CAATAATTAA ATAATACGTT TCGAACTAAT GATTTTCAGT TCATTCATTA
mir-983 KO          TATTATATTG CAATAATTAA ATAATACGTT TCGAACTAAT GATTTTCAGT TCATTCATTA

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                        65         75         85         95        105        115
Wild type           GGTAGTTACG CATTATCTAG TTGTTGTAAA CATTCAACTC GATGGCGGAT GAGAAATTAC
mir-983 KO          GGTAGTTA-- ---------- ---------- ---------- ---------- ----------

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                       125        135        145        155        165        175
Wild type           TTATGATGTA CCATTCTTGT AATAAACTGT GTGTGTAGAT TCATATTATA TGCATATGCT
mir-983 KO          ---------- ---------- ---------- ---------- ---------- ----------

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                       185        195        205        215        225        235
Wild type           ATTATATTGC AATAATTAAA TAATACGTTT CGAACTAATG ATTTTCAGTT CATTCATTAG
mir-983 KO          ---------- ---------- ---------- ---------- ---------- ----------

                    ....|....| ....|....| ....|....| ....|...
                       245        255        265        275
Wild type           GTAGTTACGC ATTATCTAGT TGTTGTAAAC ATTCAACT
mir-983 KO          ---------- ----TCTAGT TGTTGTAAAC ATTCAACT
```

Figure S2

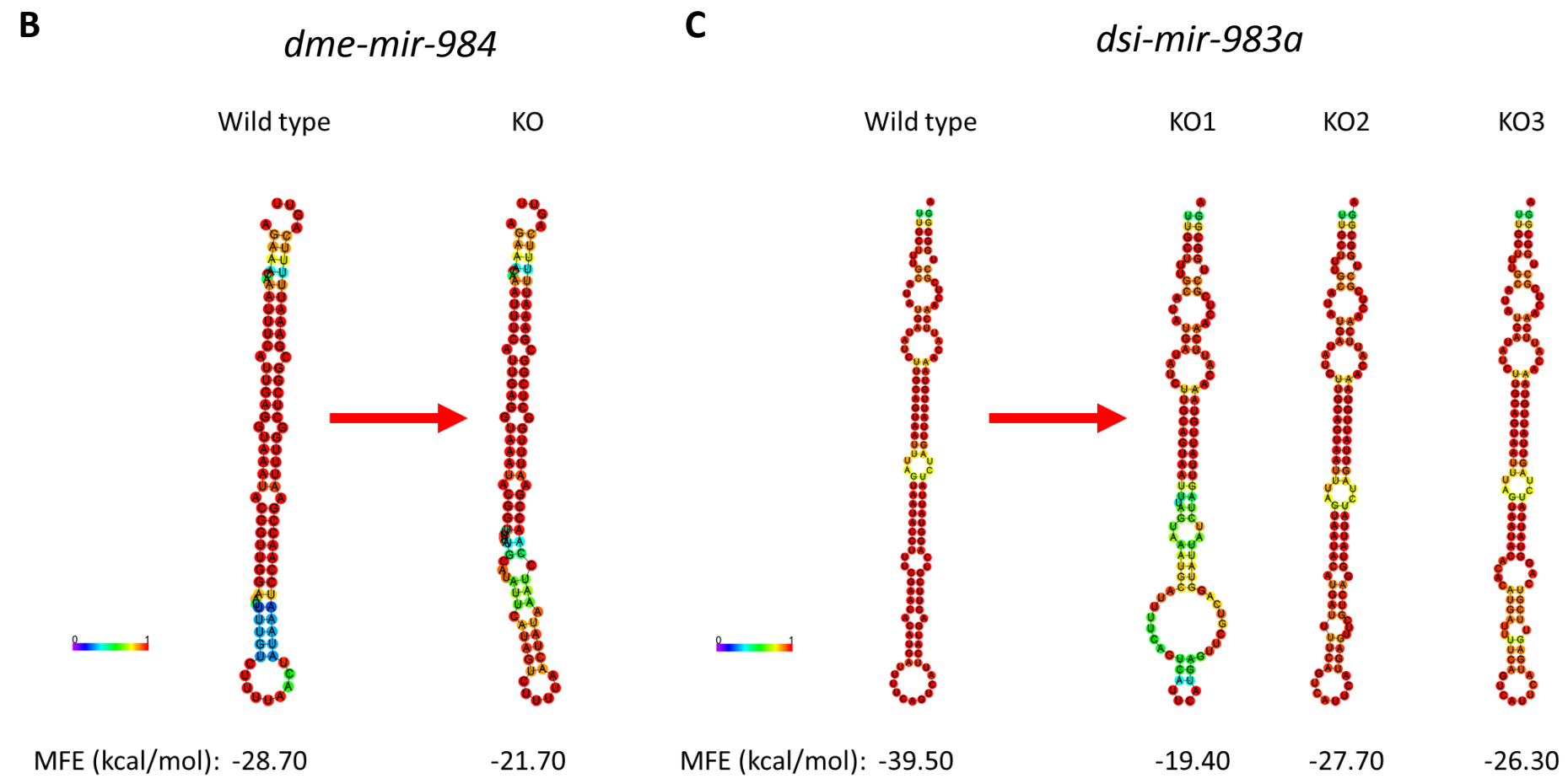

B
dme-mir-984
Wild type
KO
MFE (kcal/mol): -28.70
-21.70
C
dsi-mir-983a
Wild type
KO1
KO2
KO3
MFE (kcal/mol): -39.50
-19.40
-27.70
-26.30

Figure S2

D. melanogaster

Cross scheme for generating *dme-mir-984* or *dme-mir-983-1/-2* KO lines:

D

After injection， if the offspring is **Male**:

$\frac{\text{P\{w[+mC]=lacW\}Smox[G0348]}}{\text{FM7c}}$ ♀ (virgin) × $\frac{\text{w1118,+/*}}{\text{Y}}$ ♂ F0 SURVEYOR® Mutation Detection

↓

$\frac{\text{w1118,+/*}}{\text{FM7c}}$ ♀ (virgin) × $\frac{\text{FM7c}}{\text{Y}}$ ♂ F1 Single pair cross, 20 vials

↓

$\frac{\text{w1118,+/*}}{\text{FM7c}}$ ♀ × $\frac{\text{w1118,+/*}}{\text{Y}}$ ♂ F2 Sequencing, sibling cross

↓

$\frac{\text{w1118,*}}{\text{FM7c}}$ ♀ × $\frac{\text{w1118,*}}{\text{Y}}$ ♂ F3 Keep the stock

E

After injection， if the offspring is **Female**:

$\frac{\text{w1118,+/*}}{\text{w1118,+/*}}$ ♀ (virgin) × $\frac{\text{FM7c}}{\text{Y}}$ ♂ F0 SURVEYOR® Mutation Detection

↓

$\frac{\text{w1118,+/*}}{\text{FM7c}}$ ♀ (virgin) × $\frac{\text{FM7c}}{\text{Y}}$ ♂ F1 Single pair cross, 20 vials

↓

$\frac{\text{w1118,+/*}}{\text{FM7c}}$ ♀ × $\frac{\text{w1118,+/*}}{\text{Y}}$ ♂ F2 Sequencing, sibling cross

↓

$\frac{\text{w1118,*}}{\text{FM7c}}$ ♀ × $\frac{\text{w1118,*}}{\text{Y}}$ ♂ F3 Keep the stock

Figure S2

D. simulans

Cross scheme for generating *dsi-mir-983a* KO lines:

F

After injection， if the offspring is **Male**:

$\frac{+}{+}$ ♀ (virgin) × $\frac{w501,*/+}{Y}$ ♂ F0 SURVEYOR® Mutation Detection

↓

$\frac{w501,*/+}{+}$ ♀ (virgin) × $\frac{+}{Y}$ ♂ F1 Single pair cross, 20 vials

↓

$\frac{w501,*/+}{Y}$ ♂ × $\frac{w501,*/+}{+}$ ♀ (virgin) F2 Sequencing, sibling cross

↓

$\frac{w501,*}{w501,*}$ ♀ (virgin) × $\frac{w501,*}{Y}$ ♂ F3 Single pair cross, sequencing, confirm miRNA mutant

G

After injection， if the offspring is **Female**:

$\frac{w501,*/+}{w501,*/+}$ ♀ (virgin) × $\frac{+}{Y}$ ♂ F0 SURVEYOR® Mutation Detection

↓

$\frac{w501,*/+}{+}$ ♀ (virgin) × $\frac{+}{Y}$ ♂ F1 Single pair cross, 20 vials

↓

$\frac{w501,*/+}{Y}$ ♂ × $\frac{w501,*/+}{+}$ ♀ (virgin) F2 Sequencing, sibling cross

↓

$\frac{w501,*}{w501,*}$ ♀ (virgin) × $\frac{w501,*}{Y}$ ♂ F3 Single pair cross, sequencing, confirm miRNA mutant

Figure S3

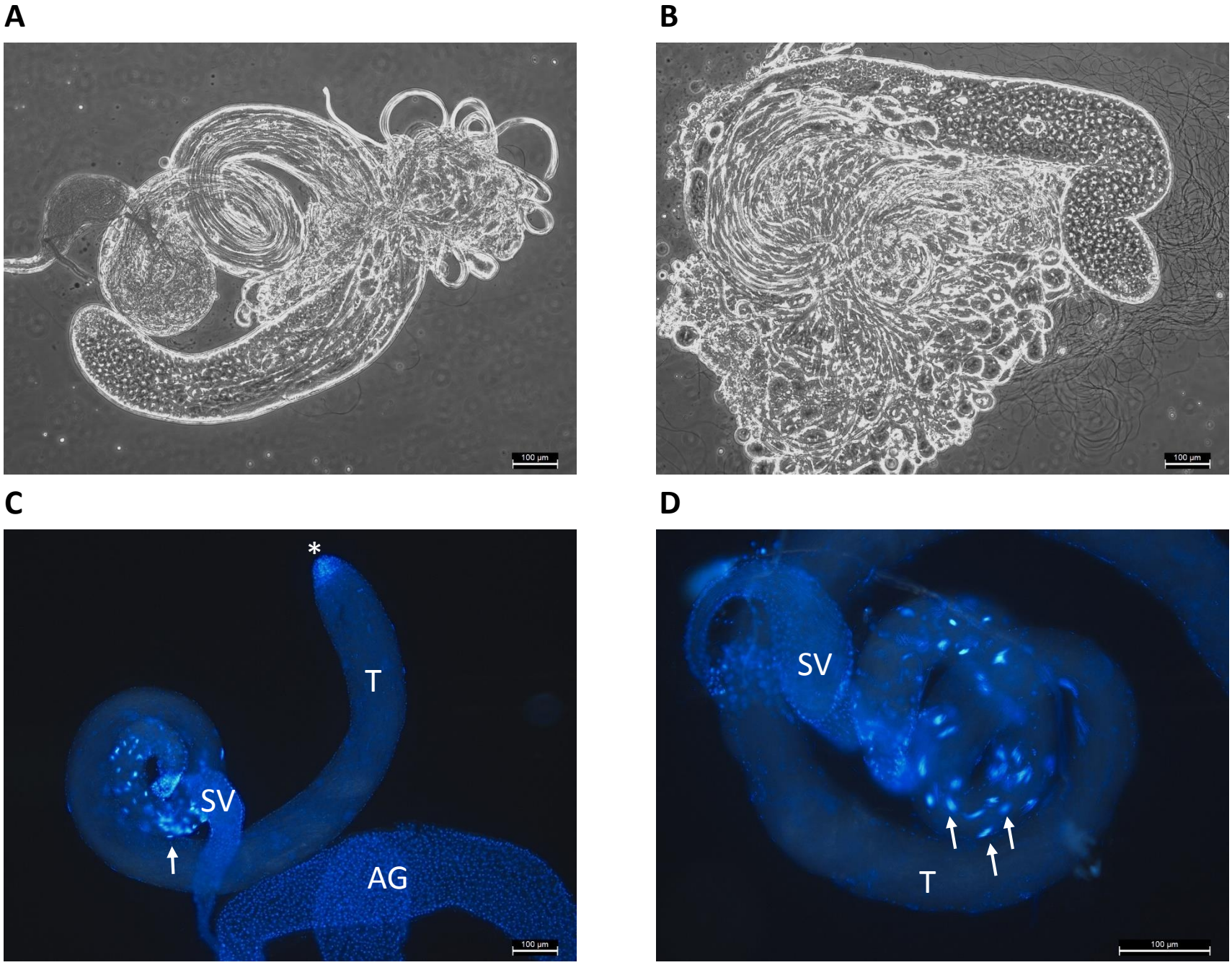
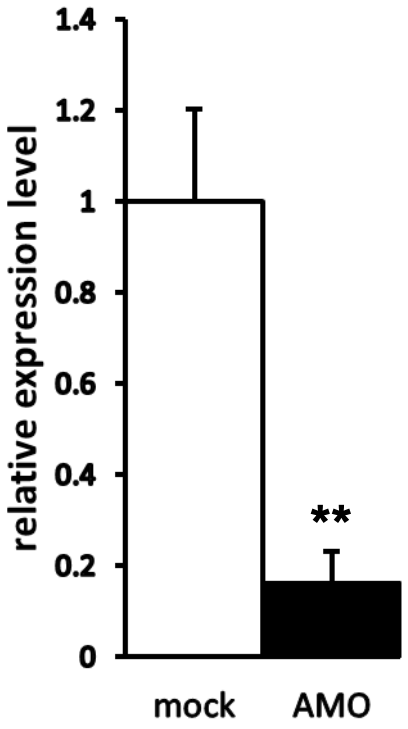


Figure S3

E



F

$$\frac{\mathbf{W1118,\ \Delta mir{-}983}}{\mathbf{Y}};\frac{+}{+};\frac{+}{+}\ ♂ \times \frac{*}{*};\frac{*}{*};\frac{*}{*}\ ♀ \quad F_0$$

$$\frac{\mathbf{W1118,\ \Delta mir{-}983}}{*};\frac{+}{*};\frac{+}{*}\ ♀ \times \frac{*}{\boldsymbol{Y}};\frac{*}{*};\frac{*}{*}\ ♂ \quad F_1$$

$$\frac{\mathbf{W1118,\ \Delta mir{-}983}}{\boldsymbol{Y}};\frac{+/*}{*};\frac{+/*}{*}\ ♂ \times \frac{*}{*};\frac{*}{*};\frac{*}{*}\ ♀ \quad F_2$$

G

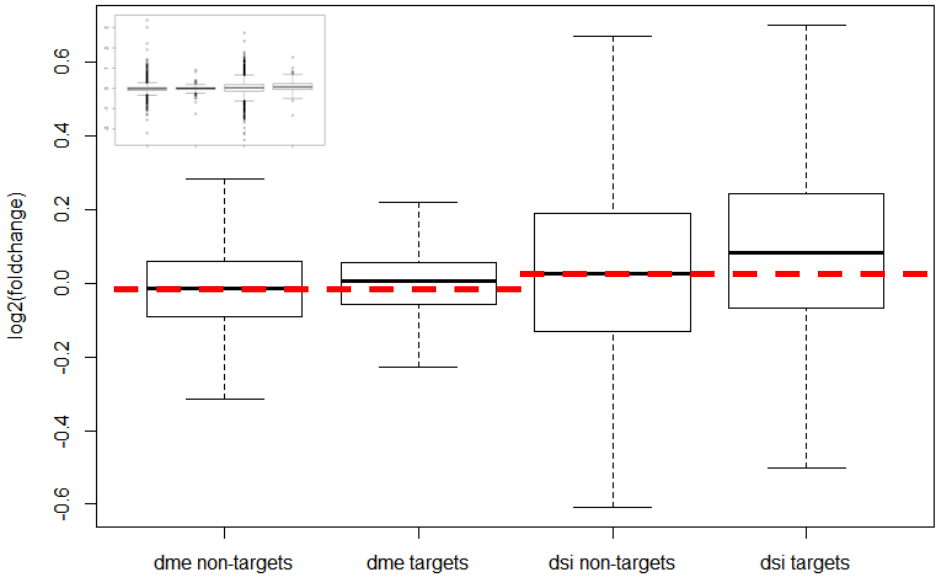
$$\frac{\mathbf{W1118}}{\mathbf{Y}};\frac{+}{+};\frac{+}{+}\ ♂ \times \frac{*}{*};\frac{*}{*};\frac{*}{*}\ ♀ \quad F_0$$

$$\frac{\mathbf{W1118}}{*};\frac{+}{*};\frac{+}{*}\ ♀ \times \frac{*}{\boldsymbol{Y}};\frac{*}{*};\frac{*}{*}\ ♂ \quad F_1$$

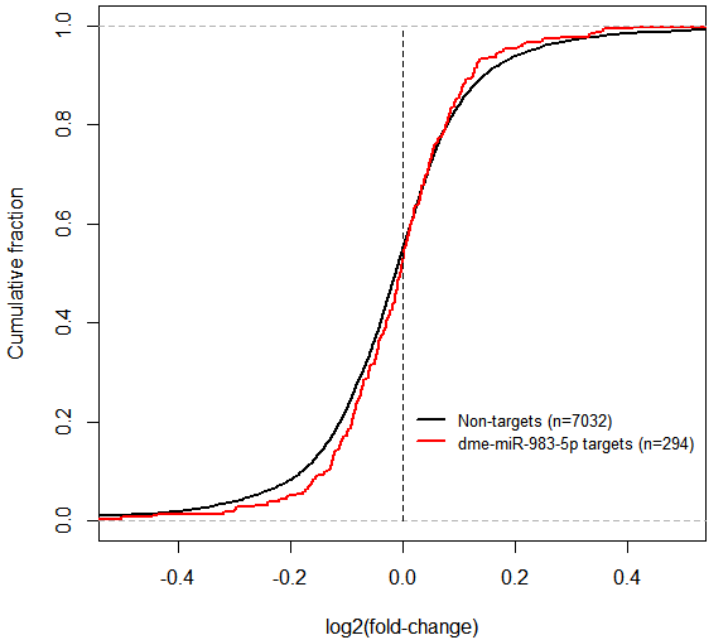
$$\frac{\mathbf{W1118}}{\boldsymbol{Y}};\frac{+/*}{*};\frac{+/*}{*}\ ♂ \times \frac{*}{*};\frac{*}{*};\frac{*}{*}\ ♀ \quad F_2$$

Figure S4

A



B



C

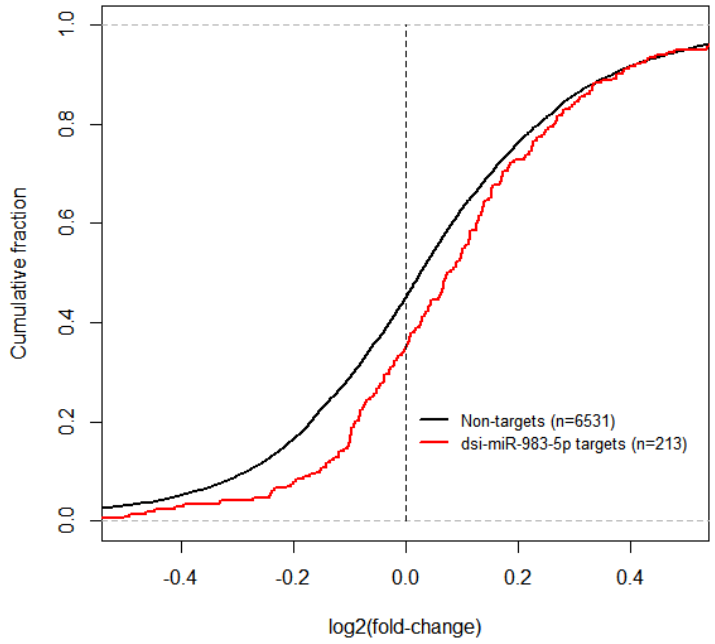
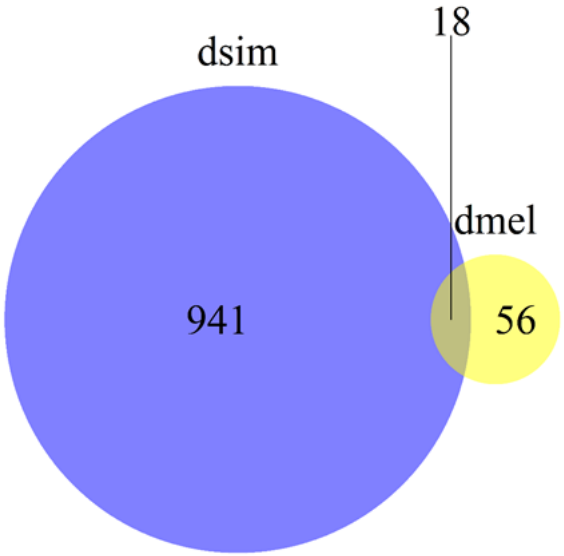
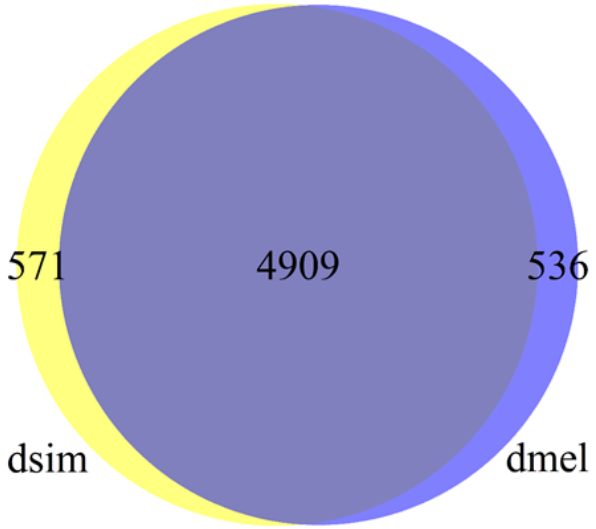


Figure S4

D



E



F

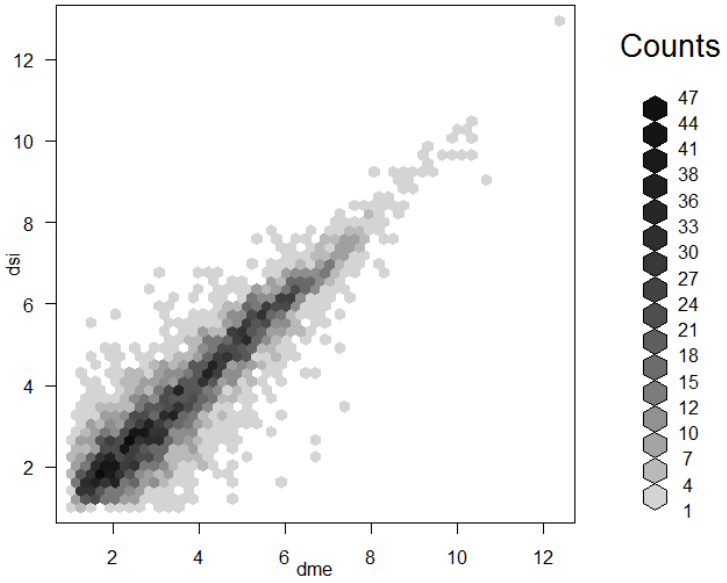


Table S1 Progeny count of control and *dsi-mir-983a* KO flies

D. simulans		♂	♀	**P value**
First round	Control	1354	1319	-
	KO-1	240	271	0.1266
	KO-2	159	181	0.1766
	KO-3	176	149	0.2334
Second round	Control	921	904	-
	KO-1	398	358	0.3134
	KO-2	314	324	0.5869
	KO-3	239	227	0.7515

Table S2 Significantly mis-regulated genes in *D. melanogaster* and *D. simulans*

	D. melanogaster		***D. simulans***	
	Up-regulated	Down-regulated	Up-regulated	Down-regulated
5p TargetScan targets	3	3	38	8
non-targets	58	57	583	577
P Value	1		< 0.0001	
5p PITA targets	2	4	27	10
non-targets	59	56	594	575
P Value	0.4392		0.0112	

Table S3 Mis-regulated genes are enriched in male biased genes in *D. simulans*

D. simulans	**misregulated genes**	**other genes**	**P value**
male biased	408	1326	< 0.0001
female biased	304	2123	

Table S4 Small RNA sequencing libraries used in this study

GEO assceesion	genetic background	tissue	species	citations
GSM909277	Oregon R	Testes	*D. melanogaster*	(Toledano et al. 2012)
GSM909278	Oregon R	Testes	*D. melanogaster*	(Toledano et al. 2012)
GSM280085	Oregon R	Testes	*D. melanogaster*	(Czech et al. 2008)
GSM548584	yw67c23(2)	Testes	*D. melanogaster*	(Rozhkov et al. 2010)
GSM2562978	Canton S	Testes	*D. melanogaster*	-
GSM2562975	Z56	Testes	*D. melanogaster*	-
GSM1165053	NC48S	Testes	*D. simulans*	(Lyu et al. 2014)
-	wild-type	Testes	*D. sechellia*	this study
-	wild-type	Testes	*D. erecta*	this study
GSM548610	strain 9	Testes	*D. virilis*	(Rozhkov et al. 2010)
GSM548623	strain 160	Testes	*D. virilis*	(Rozhkov et al. 2010)

Table S5 TALEN binding sites for different miRNAs

miRNA	TALEN binding site
dme-mir-984	TTATAGTTAAAAGACAAAATTCCAACCGTATTTACCTCAATGAA
dme-mir-983-1/-2	GAATGTTTACAACAACTAGATAATGCGTAACTACCTAATGAATGAACTG
dsi-mir-983a	GACTGAAAATCATGTGTTCGAAAGGTATTACTAAATTACTGCAAG

Mature miRNA sequences are in red, the left and right TALEN binding sites are underlined.

Table S6 Primers used for miRNAs mutant detection

miRNA	F primer	R primer	Length
dme-mir-983-1/-2 & dme-mir-984	TGCGCCATTCAATGTCTATC	TGAGTCGTCAACTCGGTGAG	566
dsi-mir-983a	GCACGGTACGTTTATGTTTCTAGG	GGTAAAGAAGGGAAAGCTTTTGGG	899

Table S7 Drosophila stocks used in this study

Species	Genotype
D. melanogaster	w[1118]
D. melanogaster	Canton-S
D. melanogaster	w[67c23] P{w[+mC]=lacW}Smox[G0348]/FM7c
D. simulans	w[501]
D. simulans	Sim6
D. erecta	wild-type (UCSD Stock Number: 14021-0224.01)
D. sechellia	wild-type (UCSD Stock Number: 14021-0248.20)

Table S8 Primers used for qRT-PCR

A

The reverse transcription primers

miRNA	RT Primers
dme-miR-983-5p	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACTCATTA
dme-miR-983-3p	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACAGATAA
dme-miR-984-5p	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACAAATTC
dme-miR-984-3p	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACCGAGCC
dsi-miR-983a-5p	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACTCATGT
dsi-miR-983a-3p	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACTAGATA
2s	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACTACAAC

Green regions show the reverse complement sequence for the last 6nt of miRNA 3' end.

B

The quantitative real-time PCR primers

miRNA	Primes
dme-miR-983-5p	GGCGATAATACGTTTCGAAC
dme-miR-983-3p	GGCGATTAGGTAGTTACGCA
dme-miR-984-5p	GGCGTGAGGTAAATACGGTTG
dme-miR-984-3p	GGCGATCCAACCGAATTT
dsi-miR-983a-5p	GGCGAGTAATACCTTTCGAAC
dsi-miR-983a-3p	GGCGATGAGTTCGTCAGGTAT
2s	GGCGTGCTTGGACTACATATGG
Universal reverse primer	GTGCAGGGTCCGAGGT

Red regions show the mature miRNA sequences (not include the last 6nt at 3' end).

Table S9 Mapping statistics for transcriptome analysis

Sample	Replicate	Left reads	Mapped	Right reads	Mapped	Read mapping rate
D. melanogaster **Control**	1	15701838	14797616	15701838	14708188	94.0%
D. melanogaster **Control**	2	16648324	15652002	16648324	15564491	93.8%
D. melanogaster **miR-983 KO**	1	15457375	14524051	15457375	14443131	93.7%
D. melanogaster **miR-983 KO**	2	16141249	15196878	16141249	15103784	93.9%
D. simulans **Control**	1	16748697	15183497	16748697	15096516	90.4%
D. simulans **Control**	2	16220169	14710465	16220169	14620237	90.4%
D. simulans **miR-983 KO-1**	1	17572804	16214515	17572804	16117784	92.0%
D. simulans **miR-983 KO-1**	2	18118800	16680233	18118800	16587830	91.8%