Supplemental Material for:

Evolutionary patterns of metazoan microRNAs reveal targeting principles in the let-7 and miR-10 families

Justin M. Wolter, Alexander Linse, Victoria A. Godlove, Hoai Huang Thi Le, Thuy-Duyen Nguyen, Kasuen Kotagama, Alissa Lynch, Alan Rawls, Marco Mangone*

* To whom correspondence should be addressed. Tel: (480) 965-7957; Email: mangone@asu.edu

This PDF includes:

Supplemental Materials and Methods

Supplemental Fig S1-S9

Table of Contents
Supporting Online Material

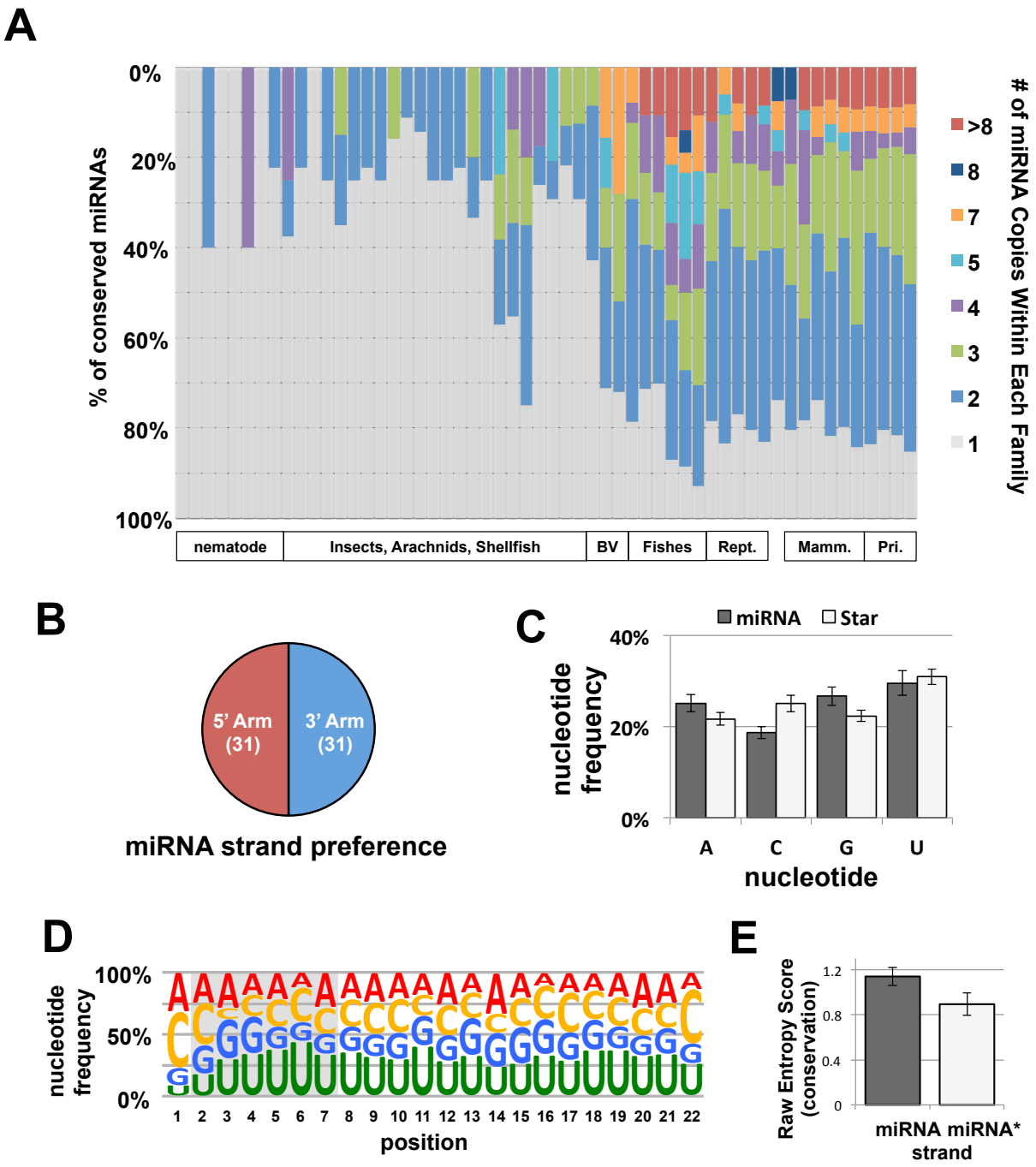
miRNA Sponge plasmid Construction:
Briefly, we replaced eGFP with dsRed2 in pLVCT-tTR-KRAB (Szulc et al. 2006) via ligation to create pLVCT-RFP-KRAB. We then cloned d2eGFP::miR-10 sponge from (Ma et al. 2010) into pLenti CMV/tet puro dest (Campeau et al. 2009) via Gateway cloning. We first integrated pLVCT-RFP-KRAB into the genome of HEK293T cells using lentivirus, picked a single cell, and clonally expanded the population. We then infected this line with lentivirus containing the CMV/tet::d2eGFP::miR-10 sponge, and selected for positive transgenics with puromycin (Sigma).

***In situ* hybridizations:**
To detect miRNA expression patterns during mouse embryogenesis, we used miRCURY LNA probes, which are specific for miRNA family members (Exiqon). E9.5 CD1 embryos were placed in MeOH and added into columns in an Intavis AG Insitu Pro. Pre-hybridization was performed at 61C for 6hrs and hybridization was performed at 61C for 10hrs. Following completion in the Insitu Pro, embryos were removed and washed with Ab buffer (1M pH9.5 Tris, 1M MgCl, 5M NaCl). Substrate was added and embryos were allowed to sit overnight at 4C. The color reaction proceeded at room temperature until desired colored was achieved. If color reaction required additional time, the substrate was removed and embryos were washed with Ab buffer. Ab buffer was then replaced with either a lower pH buffer (50mM pH7.5 Tris, 150mM NaCl, 0.1% Triton-X100) or KTBT buffer (50mM pH7.5 Tris, 150mM NaCl, 10mM KCl, 1% Tween-20) and placed at 4C overnight. Buffer was then removed and replaced with fresh substrate and embryos were allowed to resume staining at room temperature. When desired staining was achieved, substrate was replaced with PBS and embryos were stored at 4C.
Primers:

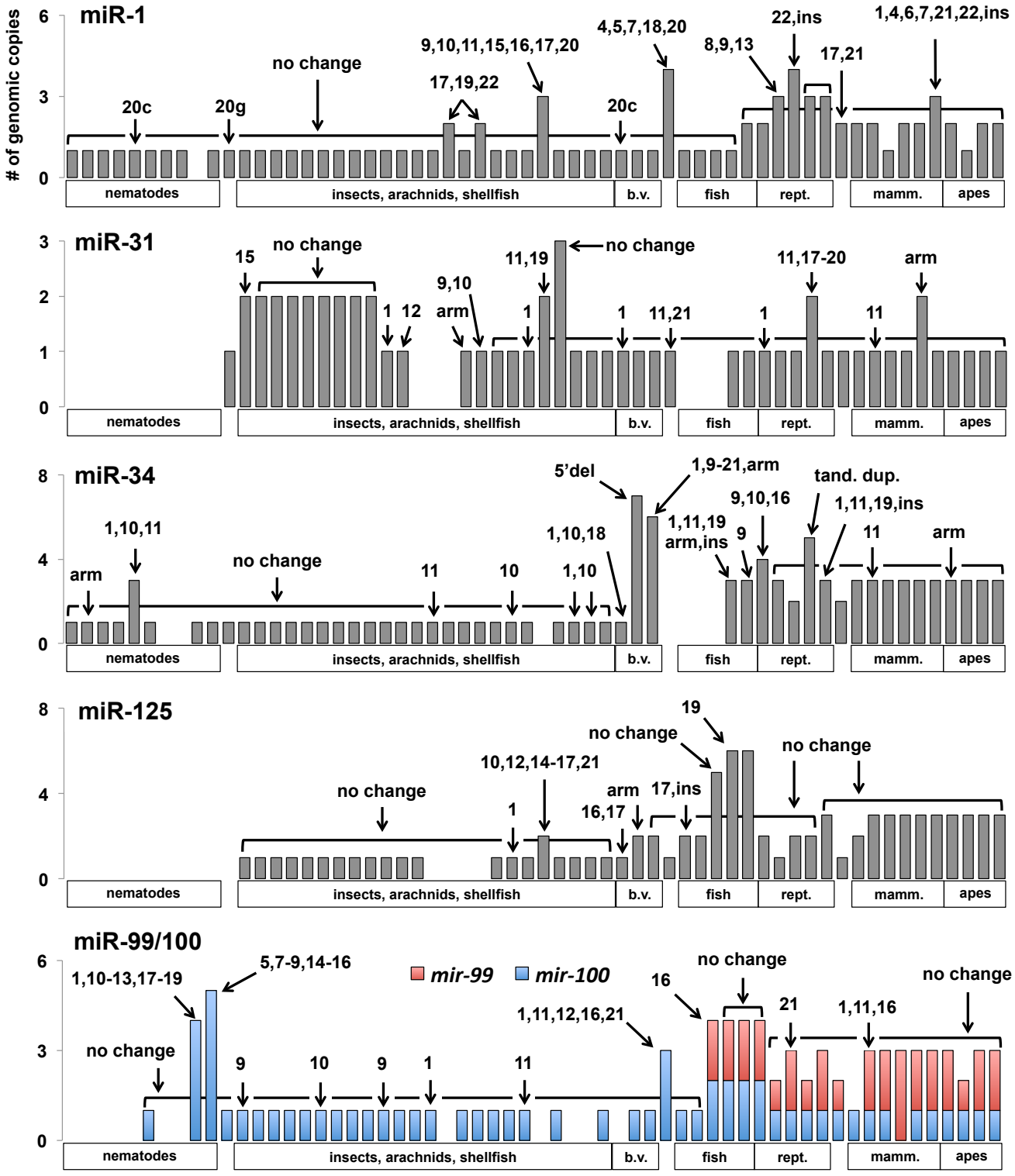
Primers:

ACTB cDNA F	TCCGCAAAGACCTGTACGCC
ACTB cDNA R	AGGGGCCGGACTCGTCATAC
ALDH1A3 qPCR F	CTCGACAAAGCCCTGAAGTTGG
ALDH1A3 qPCR R	TCAGGGGTTCTTGTCGCCAAG
CYP26A1 cDNA F	AGACCCTTCGACTGAATCCC
CYP26A1 cDNA R	GCTGAACCTGGATGCATCC
ICAM1 cDNA F	GAGATCTTGAGGGCACCCTACCC
ICAM1 cDNA R	ATAGAGGTACGTGCTGAGGCC
let-7a-2 F	GCCAGTCCCAAGTATTTGCTCCC
let-7a-2 R	GGACACCTGATATACTGAATCCCTC
let-7c F	GACATTTTACGTGACCTATGCTG
let-7c R	CCCATTAGAAATACCATTTTGACA
let-7f F	CCTTTCTCCCTTCTTACTAATCAATTGG
let-7f R	GGACAGAGTTGCAGTCAGGAAATG
miR-10a F	GGCGTCCCTTTGCGAACTGG
miR-10a R	GGGGAGAGTTCAGGTAGATG
miR-10a promoter F	TGGGCGGTGAAGGGGAGAGGC
miR-10a promoter R	GCCCAGAAATTTTCCAGTTCTCC
miR-10b F	AAGAATATTCTGGTTGTTCGCC
miR-10b R	TCTTTCTTTCTTTTCAGCACCC

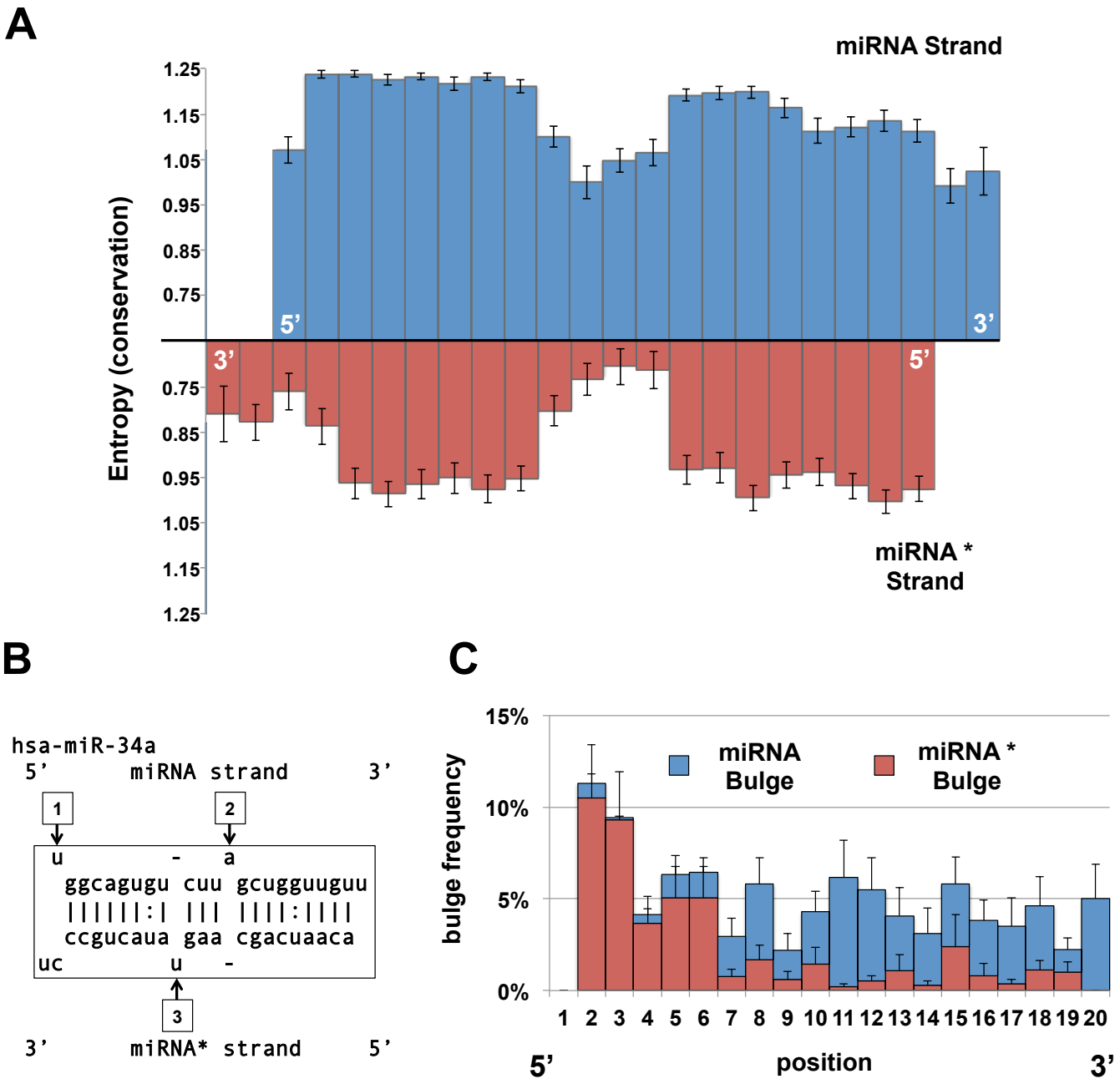
miR-10b promoter F	CCTCTGCTTCTCTGTAACTGCAG
miR-10b promoter R	GGTCCCACCCAGAGTGAGGTAC
miR-125a F	TGGCTTCTGTGTCTCTTTCACAGTGG
miR-125a R	TTGGCCAGCAGGGCCATCGTG
mir-125b-2 F	GGTCTGGAATTAGTCTATAAATGGTCG
mir-125b-2 R	CGTTATAATACAAACGCTATTTCAGTGC
miR-34a F	GGAGTCTTGCTAGTTGCCTGGG
miR-34a R	GACGTGCAAACTTCTCCCAGCC
NCOA6 cDNA F	CCCGTCGCAGAATTTAGTCTC
NCOA6 cDNA R	AGTAGAGCTGGAGGCTGAAGC
RARG cDNA F	AAAAAGTGGACAAGCTGCAGG
RARG cDNA R	CCAGCATCTCTCGGATTAAGG
RBP1 cDNA F	GTCACTCCCCGAAATGCC
RBP1 cDNA R	CCTTCCCAACCTGGAAGTCC



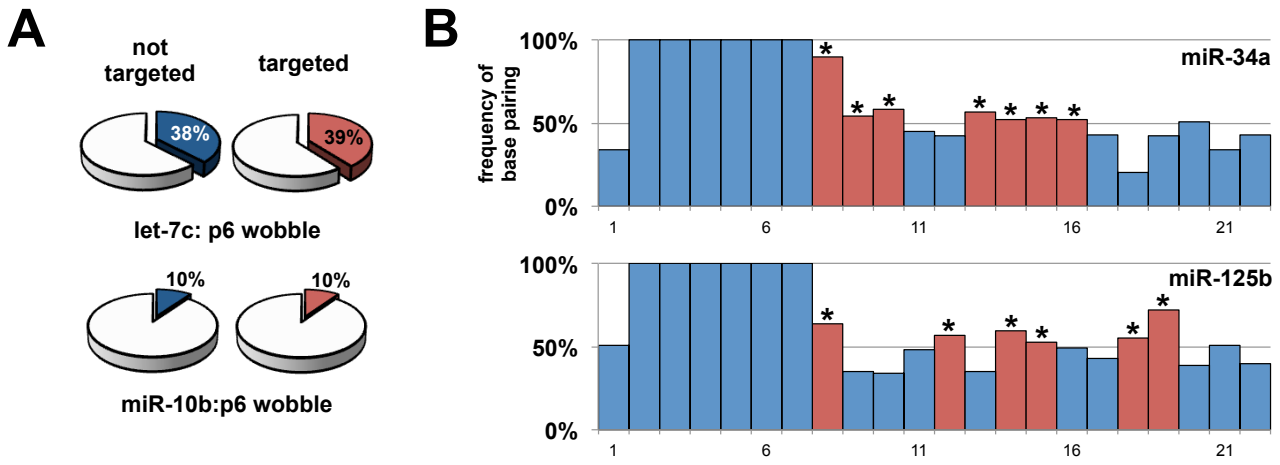
Supplemental Fig S1. Duplication rates, conservation, and nucleotide content of conserved miRNA families (A) miRNAs were grouped based on the number of members in each family, and normalized to the total number of highly conserved miRNAs in each species. Data shows that the percentage of miRNAs that belong to multicopy families is >80% in the majority of vertebrate species. **(B)** Strand bias for 62 miRNA families shows no preference for which arm is preferentially processed. **(C)** Nucleotide content of miRNA and miRNA* strands. **(D)** Nucleotide content of miRNA* strand by position. Shows a marked depletion of U residues in the first position, and no other clear patterns. **(E)** Average entropy scores from entire miRNA or miRNA* sequences.

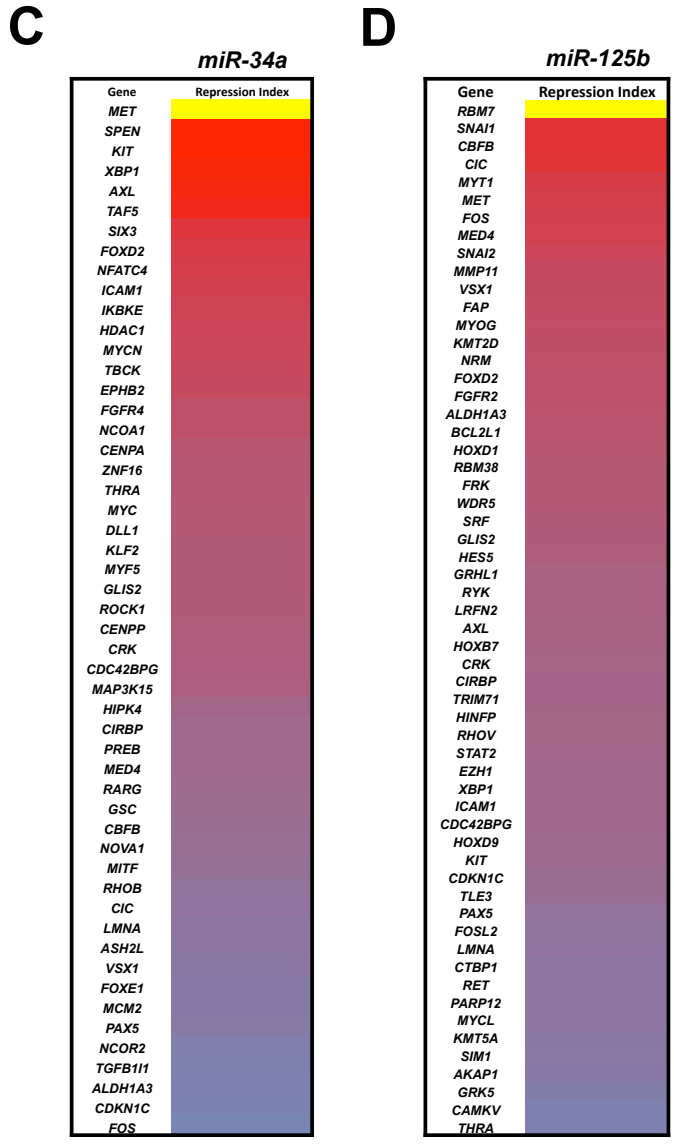


Supplemental Fig S2. History of sequence evolution in several deeply conserved metazoan miRNA families. We analyzed duplications and sequence changes of several of the most deeply conserved miRNA family members. These families were chosen based of their use our experiments, or because of unique features. For example, miR-1 is unique in that there has been a duplication without sequence divergence, and miR-31 typically has more copies in insects than in vertebrates. Regardless of the differences, the general trend is that sequence divergence most frequently occurs following duplication, and without duplication sequence changes are rare. The one consistent exception to this is the freshwater teleosts (zebrafish, catfish, and salmon), whose genomes have undergone several whole genome duplications. (arm=arm switching, ins=insertion, del=deletion, tand. dup.=tandem duplication)

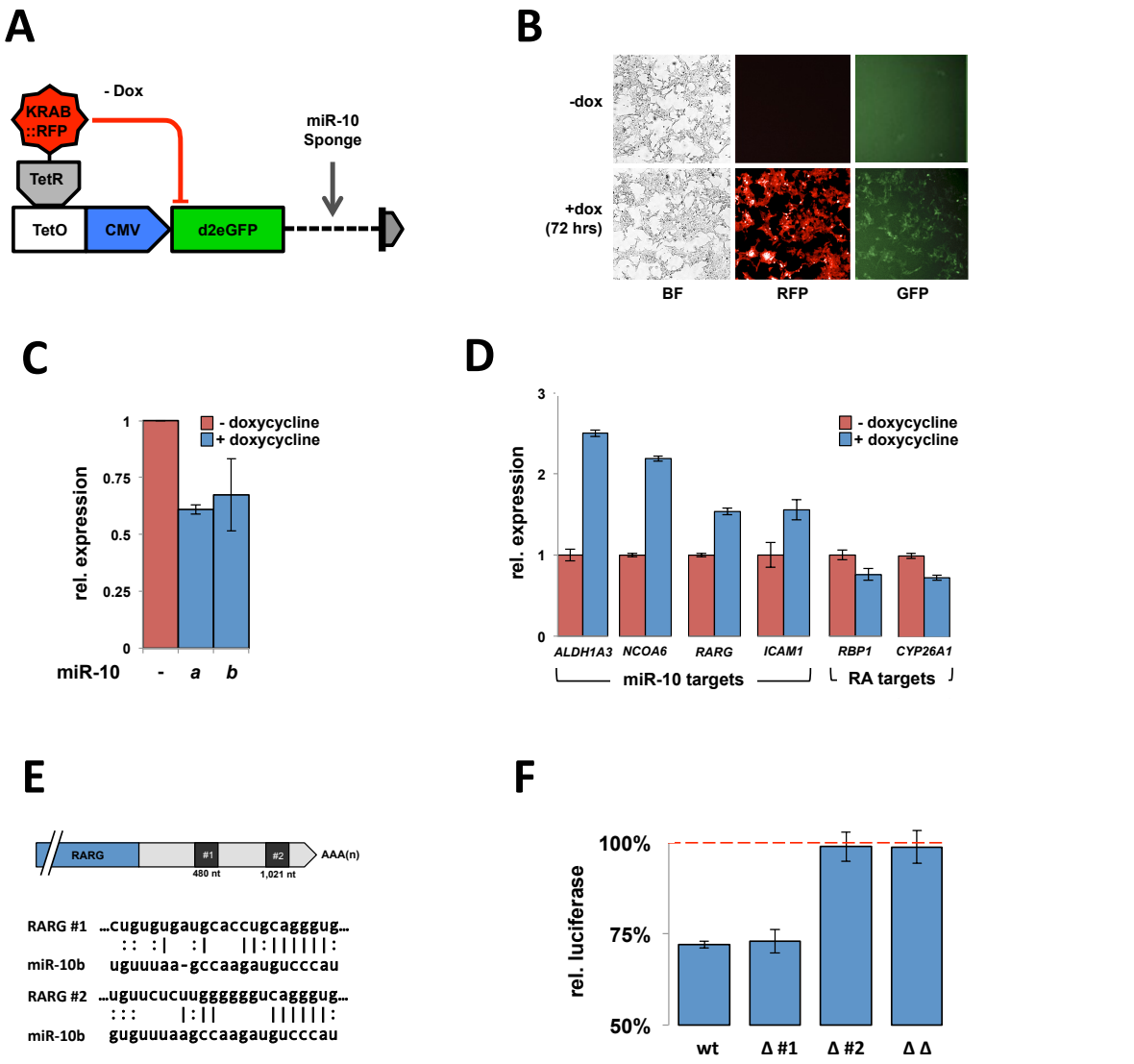


Supplemental Fig. S3: Base pairing is a conserved feature of miRNA duplexes. (A) Raw entropy scores were calculated (see materials and methods) for the miRNA or miRNA* strands for 62 miRNA families, and averaged by position. The least conserved in both strands nucleotides are also those which show the lowest frequency of base pairing (compare to Fig 1E), or are overhangs left from dicer processing (the two 3'most nucleotides). **(B)** Schematic of features analysed within miRNA duplexes. 1: mismatched bases, 2: bulge in miRNA strand, 3: bulge in miRNA* strand. Species used in this analysis were *C. elegans, D. melanogaster, S. purpuratus, C. intestinalis, P. marinus, D. rerio, M. musculus, and H. sapiens*. Data represents 590 miRNAs from 62 families. **(C)** Distribution of bulges along miRNA duplexes. Values represent mean frequency of bulge possession in 62 miRNA families. Asymetry in bulges suggests helical structure may play a role in strand discrimination.



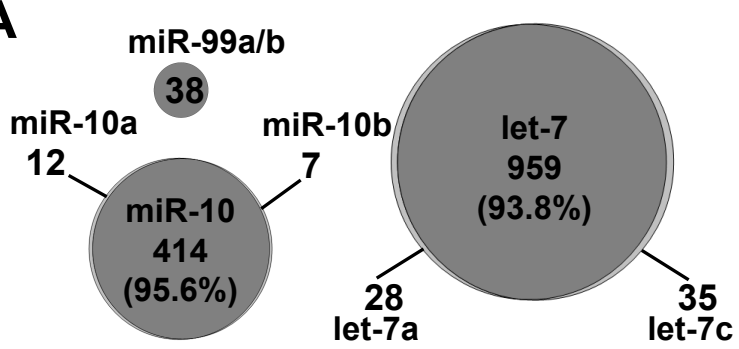


Supplemental Fig S4. miRNA targeting site analysis and target preference for miRNA family members. (A) Occurrence of p6 wobble target sites in 3'LIFE screen for let-7c and miR-10b **(B)** Targeting footprint of the top 50 target sites that lead to functional repression in the 3'LIFE screen. Heatmap of positive hits obtained for miR-34a **(C)** and miR-125b **(D)**.

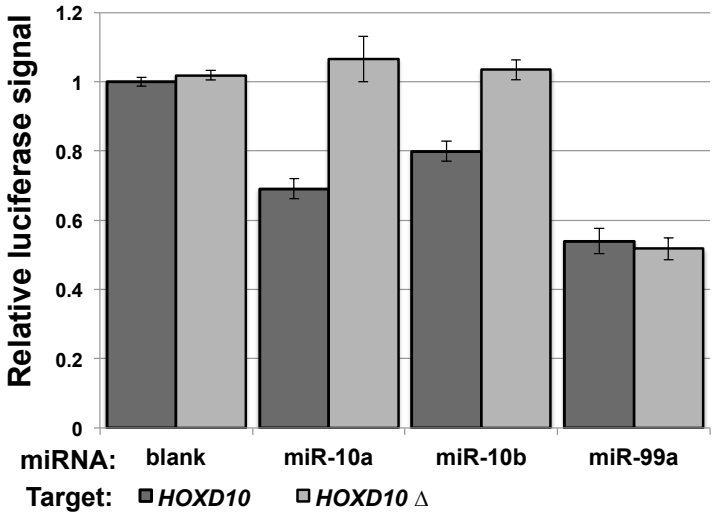


Supplemental Fig S5. *miR-10* targets several genes in the RA pathway. (A) Inducible miR-10 sponge design. **(B)** Fluorescence microscopy showing induction of RFP and GFP three days post 5 uM doxycycline treatment (5uM) in HEK293T cells. GFP images taken at 1s exposure to detect any leaky sponge expression. **(C)** RT-qPCR for mature miR-10a and miR-10b 72 hrs after dox treatment. Specificity of the qPCR primers for miR-10 family members is demonstrated in Fig. 7C. **(D)** mRNA levels of four genes in the RA pathway identified in 3'LIFE screen, as well as two RA responsive genes that were not targeted by miR-10. mRNA levels of target genes increase following doxycycline, while non-targets do not. All data normalized to ACTB, n=3. **(E)** RARG 3'UTR, showing position of two putative target sites for miR-10b. **(F)** Deletion analysis of putative miRNA target sites in RARG 3'UTR. HEK293T cells were cotransfected with miRNA and luciferase::3'UTR, and normalized to luciferase ratio of a no miRNA control (dashed line) (n=4).

A



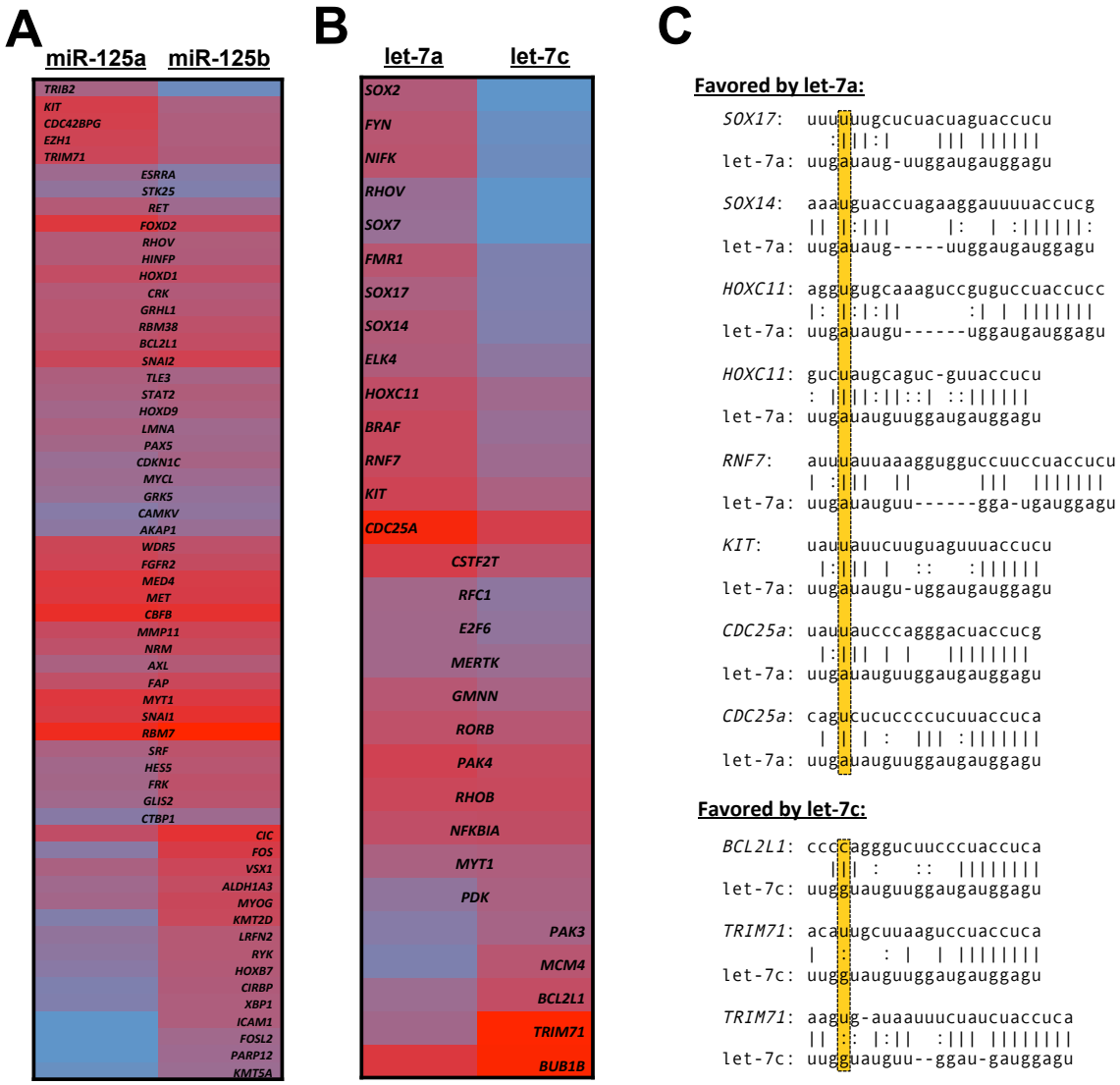
B



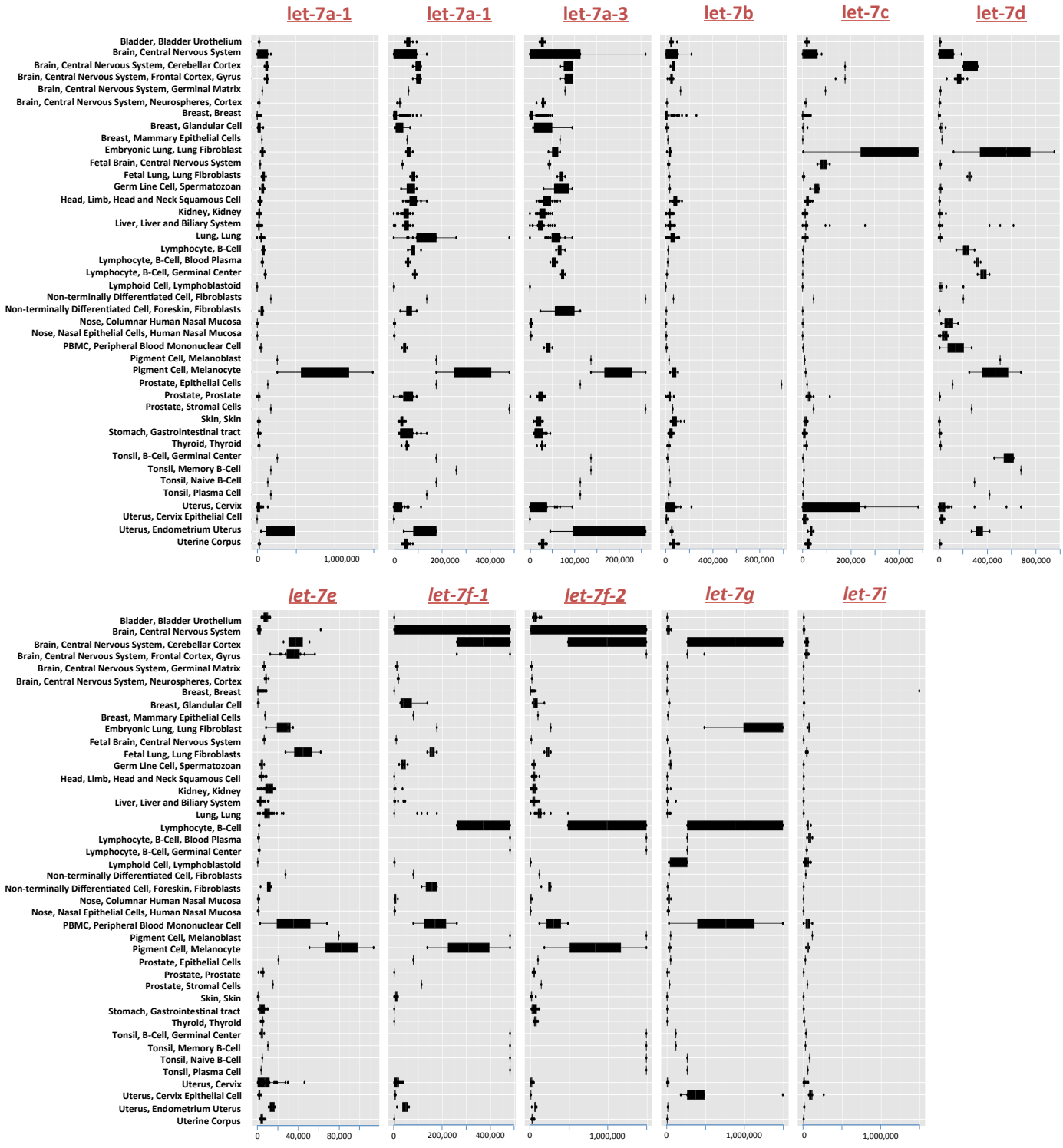
C

Gene	Repression Index	Target Scan 7.0	Diana microT CDS	PicTar
ARID3A	1.02	Poor Cons.		
ETV3	1.09	Cons.		
FOXA1	0.75			PicTar
HDAC3	0.90	Poor Cons.		
HOXA1	1.13	Cons.		PicTar
MBD3	1.20	Poor Cons.		
MTOR	0.79	Cons.	microT CDS	Pictar
MYCBP	1.27	Poor Cons.		
TRIM71	1.02	Cons.		
XRCC5	0.91	Poor Cons.		

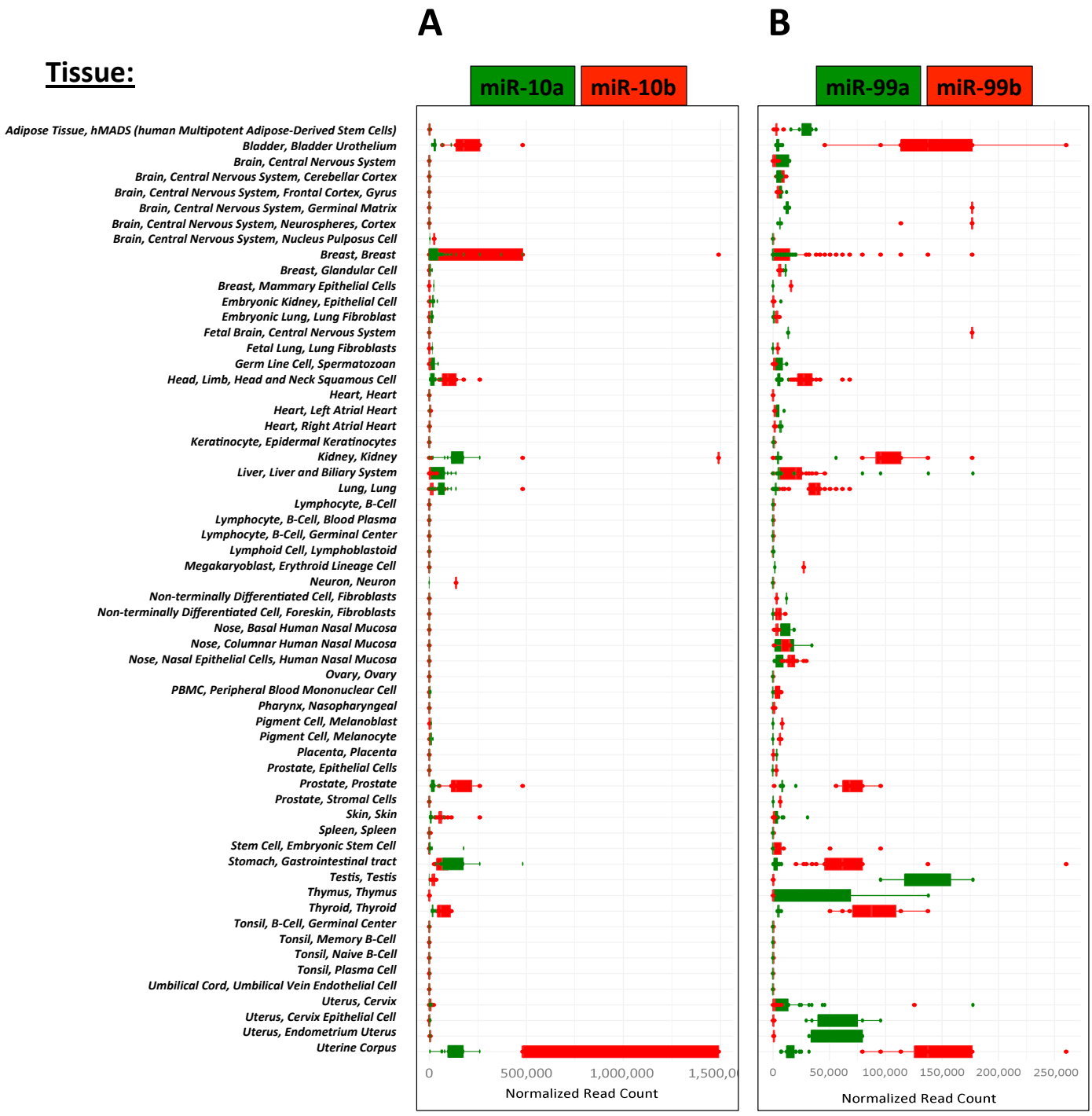
Supplemental Fig S6. miR-99a *t*argets lack canonical miRNA target sites. (A) Bioinformatic predictions for miR-99, miR-10, and let-7 family members with Diana microT-CDS, which makes independent predictions for each miRNA family member. **(B)** Mutagenesis experiments for *HOXD10*, a top target for both miR-10a/b and miR-99a. Deletion of the miR-10 target site does not rescue repression by miR-99a. **(C)** Predictions for miR-99a by three target prediction softwares. Conservation of targets predicted by TargetScan is noted. TargetScan predictions for miR-99a have 11% overlap with 3'LIFE hits compared to 35% for miR-10b and let-7c (Fig 2B).



Supplemental Fig S7. miRNA targeting site analysis and target preference for miRNA family members. **(A)** Distance analysis of targeted and non-targeting, canonical miRNA targets from screens of let-7c, miR-10b, miR-34a, and miR-125a (related to Fig. 3B). Analysis suggests that when a canonical target site is located in the last 200 nucleotides of an mRNA, it will lead to functional repression the majority of the time. However, beyond this distance, additional factors may influence the ability of a miRNA to repress the translation of a target mRNA. **(B)** Targeting footprint of the top 50 target sites that lead to functional repression in the 3'LIFE screen (related to Fig. 2B). Each miRNA prefers distinct features of target sites such as central pairing (miR-10b, miR-34a)



Supplemental Fig S8. let-7 family members have differential expression patterns in tissues: To investigate the degree of overlap of specific miRNAs in normal human tissues, we downloaded expression levels for let-7 family from miRGator v3.0 (Cho *et al.*, 2013, *Nucl. Acids Res.*), which is a compilation of high-throughput sequencing data from GEO, SRA, and TCGA.



Supplemental Fig S9: miR-10 family members have differential expression patterns in normal tissues: To investigate the degree of overlap of specific miRNAs in normal human tissues, we downloaded expression levels for miR-10 **(A)**, and miR-99 **(B)** families from miRGator v3.0 (Cho *et al.*, 2013, *Nucl. Acids Res.*), which is a compilation of high-throughput sequencing data from GEO, SRA, and TCGA. We normalized the expression levels and overlapped both miRNA family members.