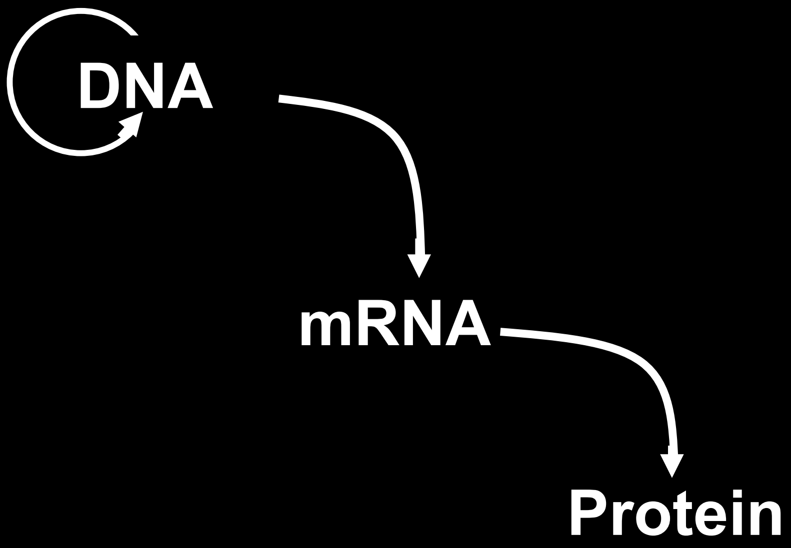
# MicroRNAs and Gene Expression

**Antonio J. Giraldez**

http://www.yale.edu/giraldezlab

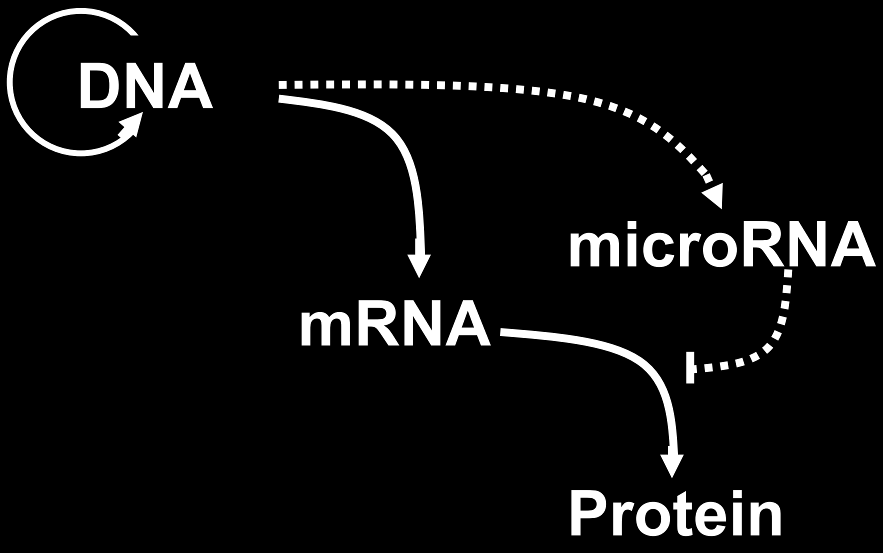


DNA

mRNA

Protein

Francis Crick 1958



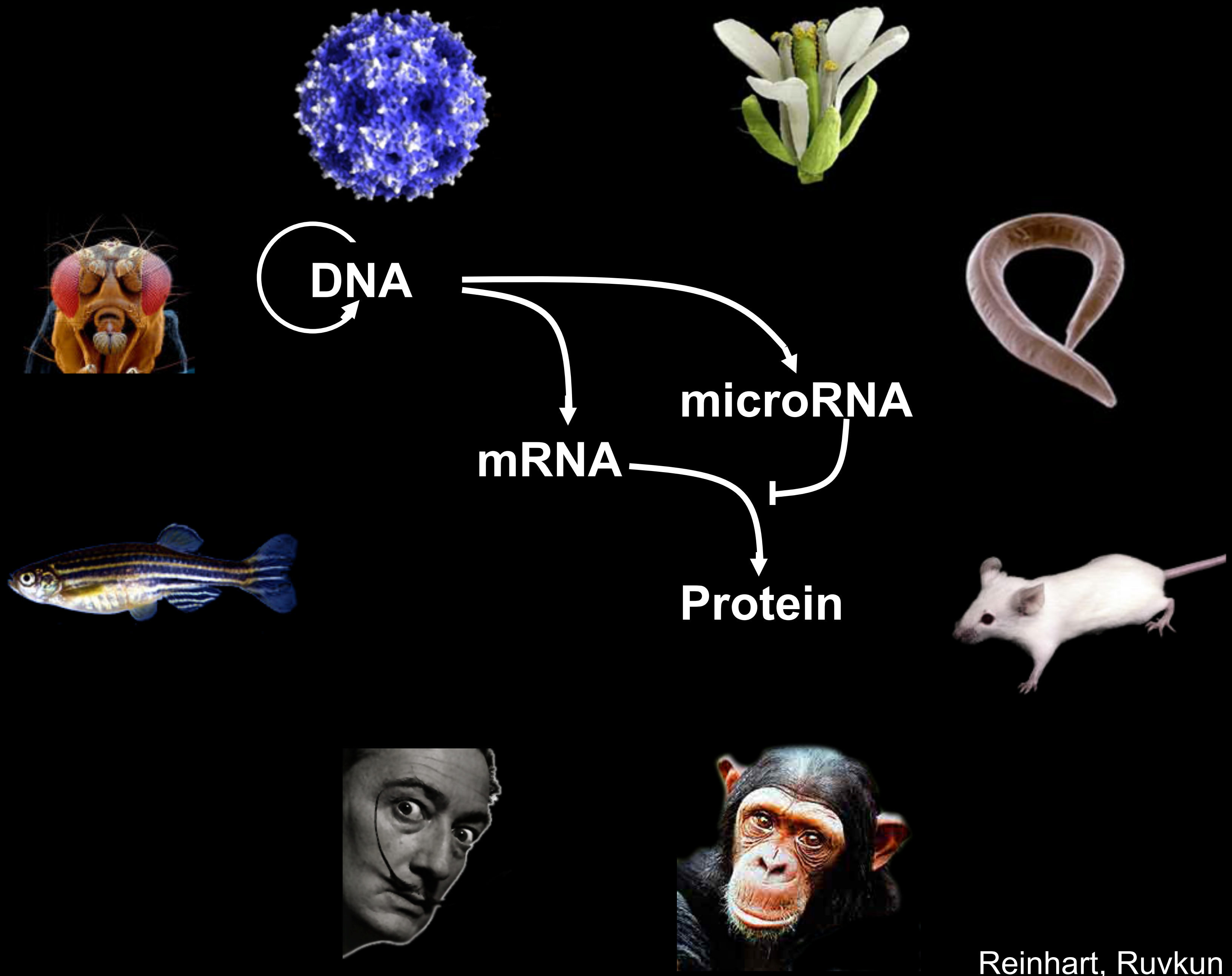
DNA

microRNA

mRNA

Protein

Ambros, Ruvkun 1993



DNA

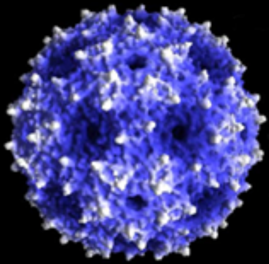
mRNA

microRNA

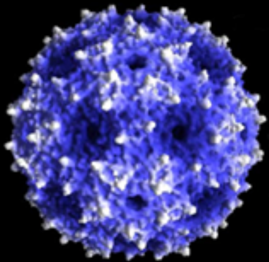
Protein

Reinhart, Ruvkun 2000

# microRNAs



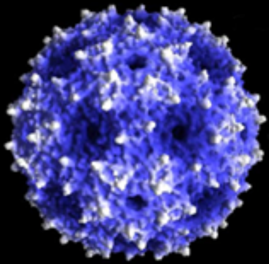
Reinhart, Ruvkun 2000




# microRNAs

1-3% of all genes



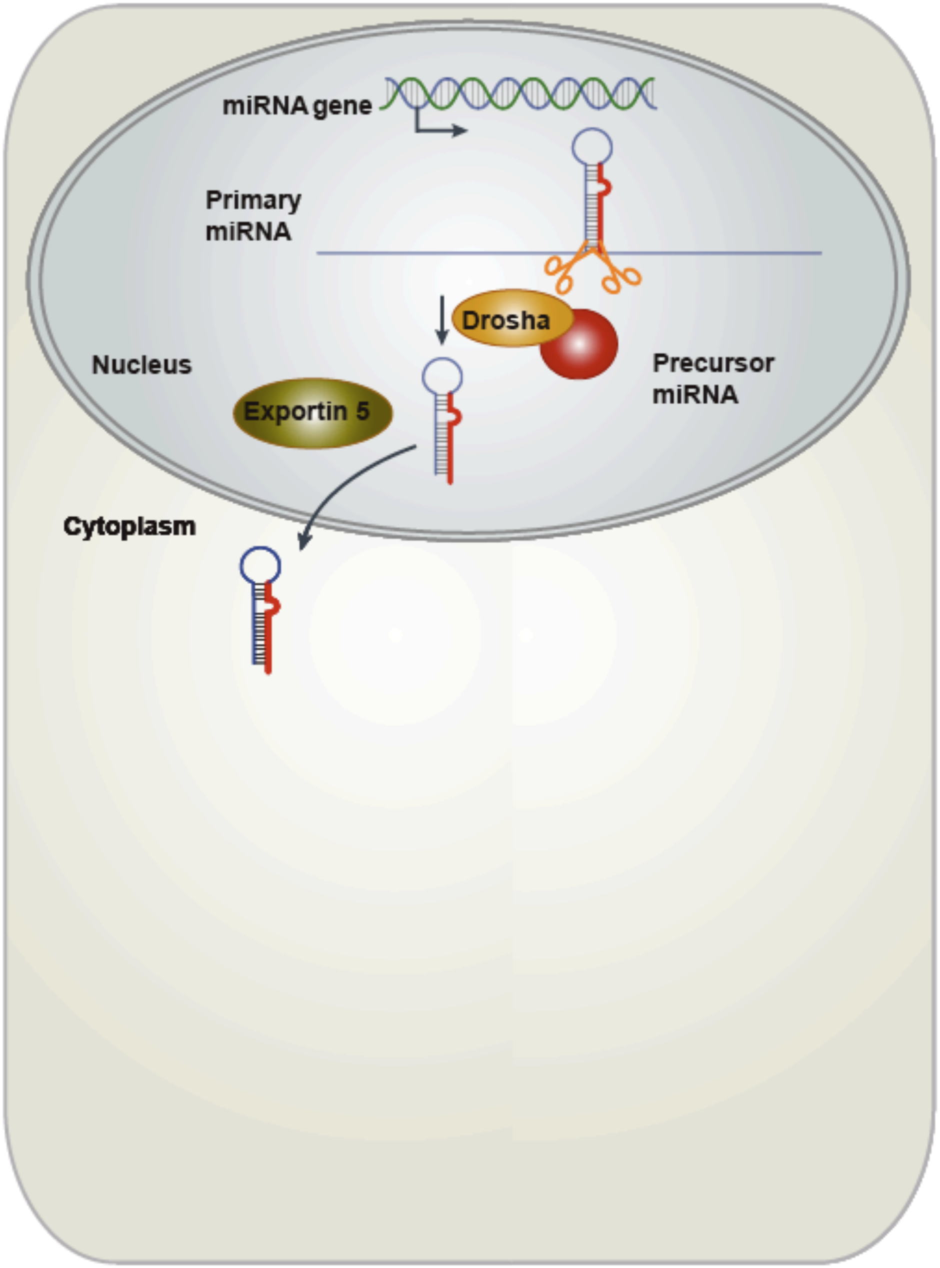



# microRNAs

1-3% of all genes
predicted to target >25% of the vertebrate genes

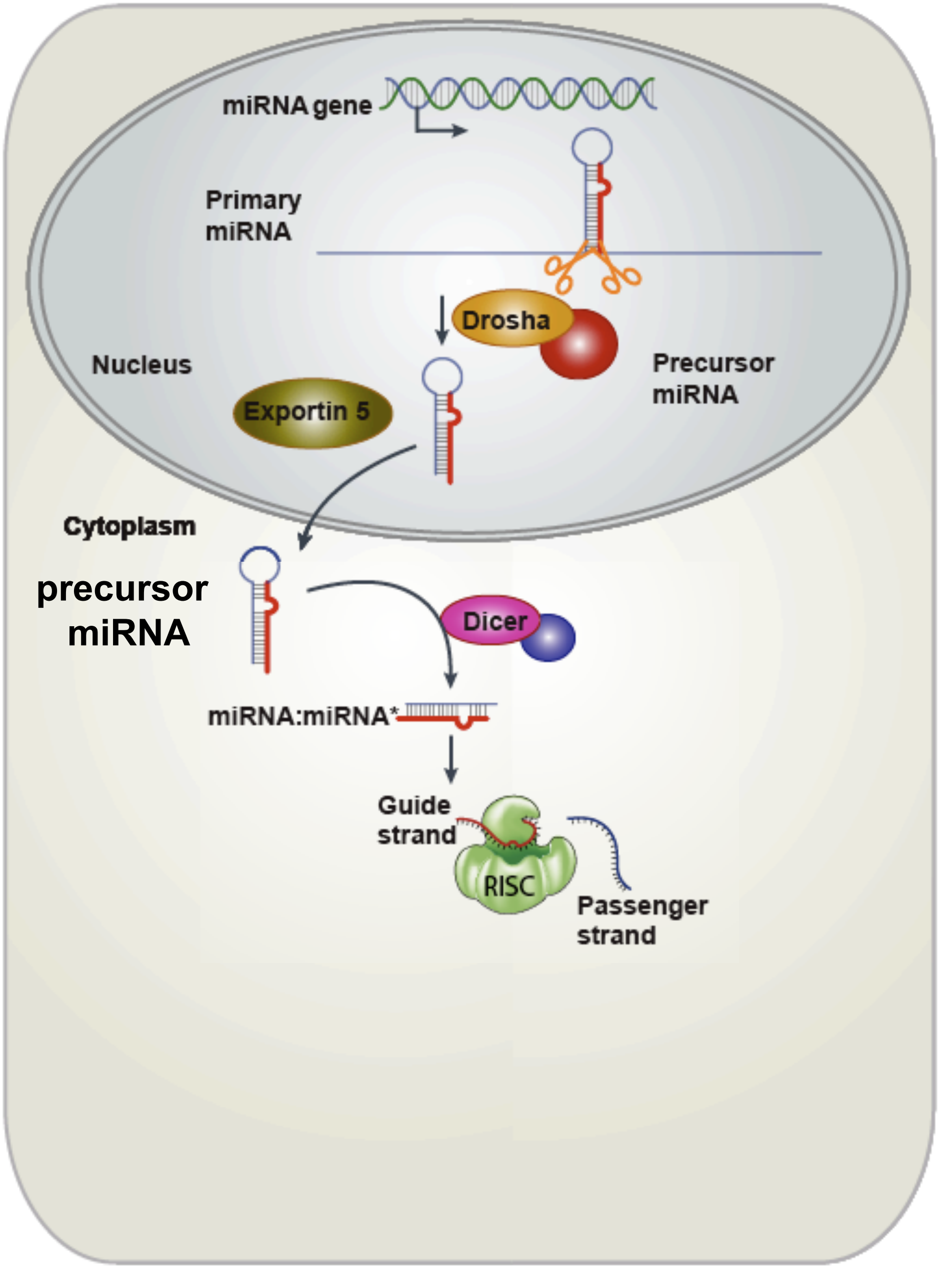
# microRNAs:

**~22nt RNAs**
**Processed by Drosha**



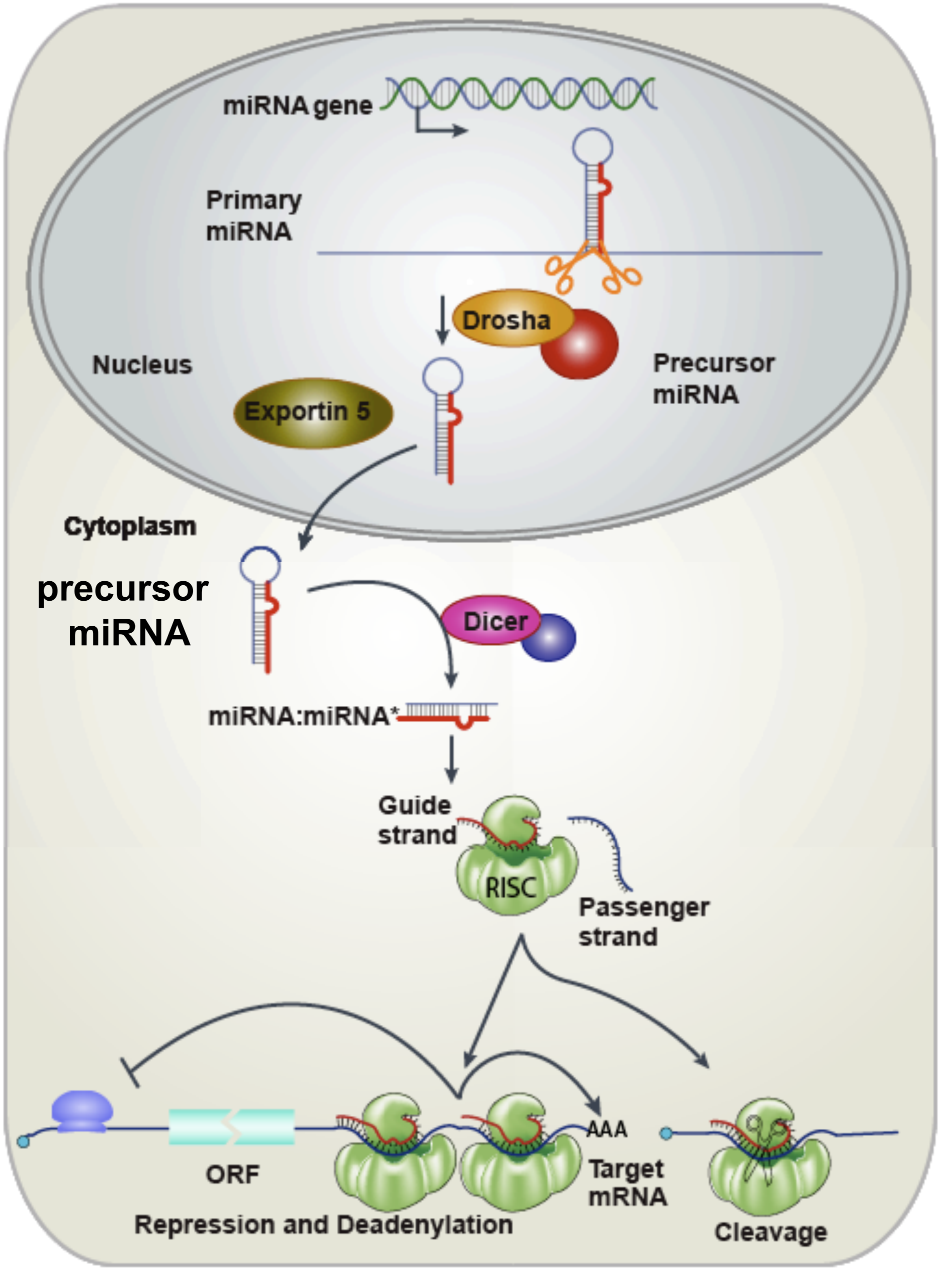
# microRNAs:

~22nt RNAs
Processed by Drosha
Processed from a hairpin by Dicer
Bound by Argonaute



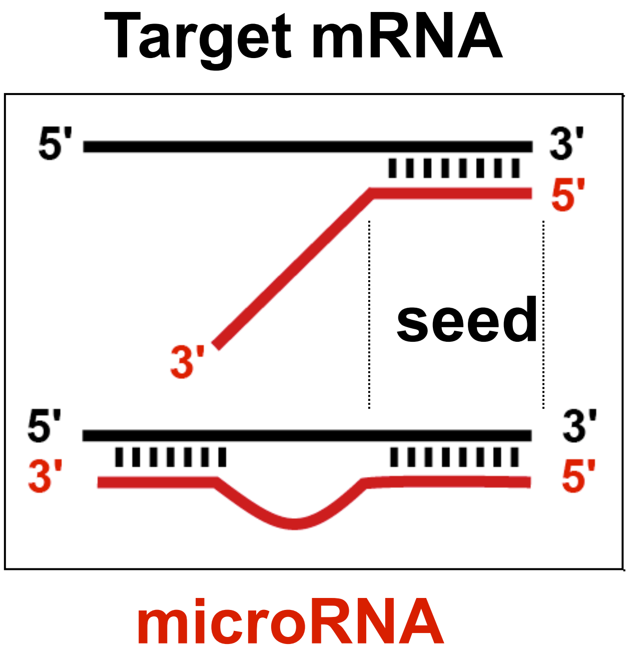
# microRNAs:

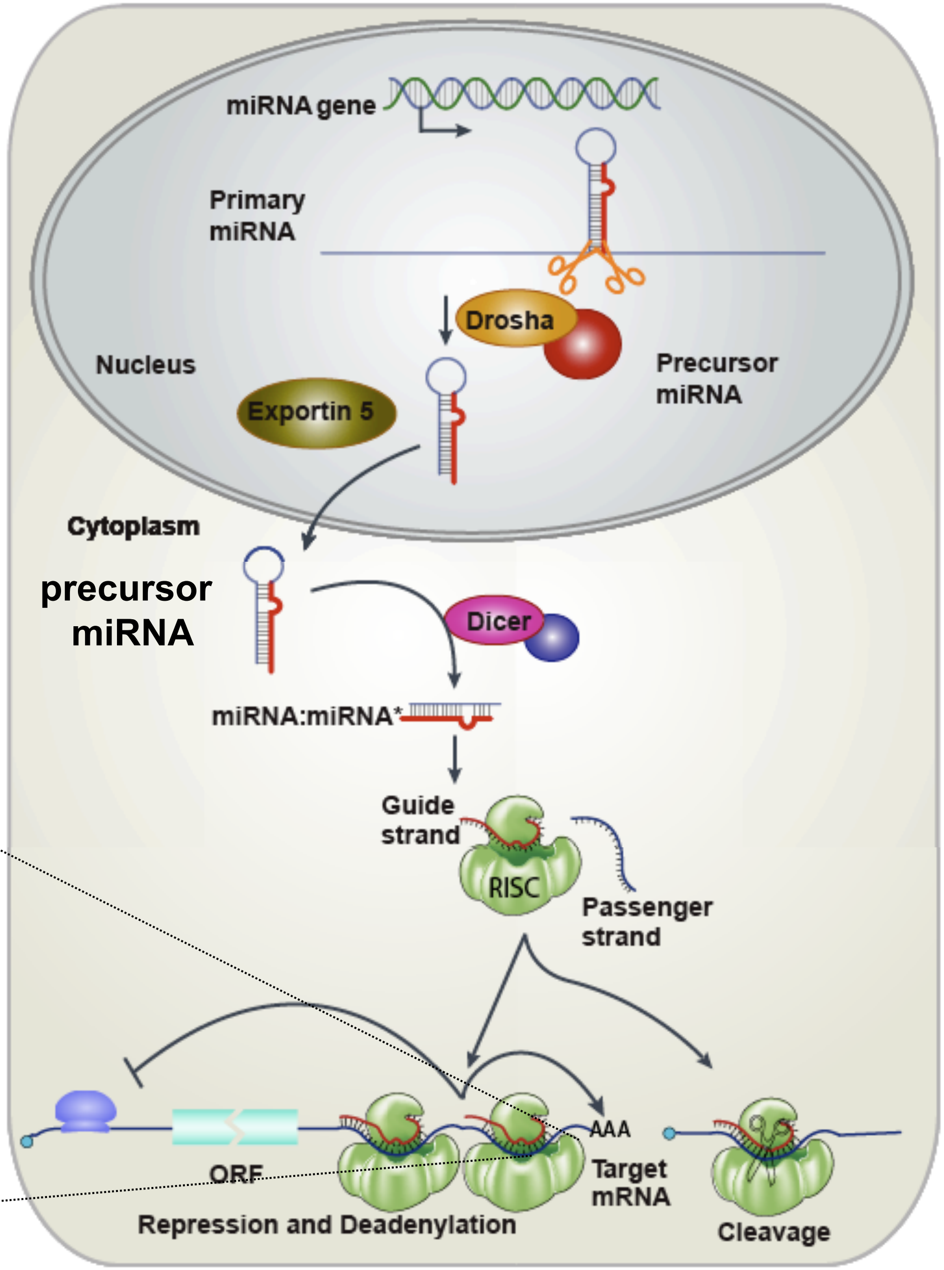
~22nt RNAs
Processed by Drosha
Processed from a hairpin by Dicer
Bound by Argonaute
Bind to the target 3’UTR
Repress translation
Induce mRNA deadenylation



# microRNAs:

~22nt RNAs
Processed by Drosha
Processed from a hairpin by Dicer
Bound by Argonaute
Bind to the target 3’UTR
Repress translation
Induce mRNA deadenylation
Seed is important for specificity

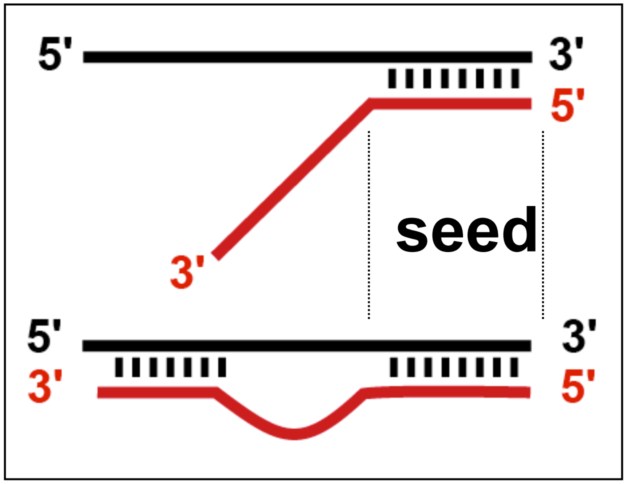




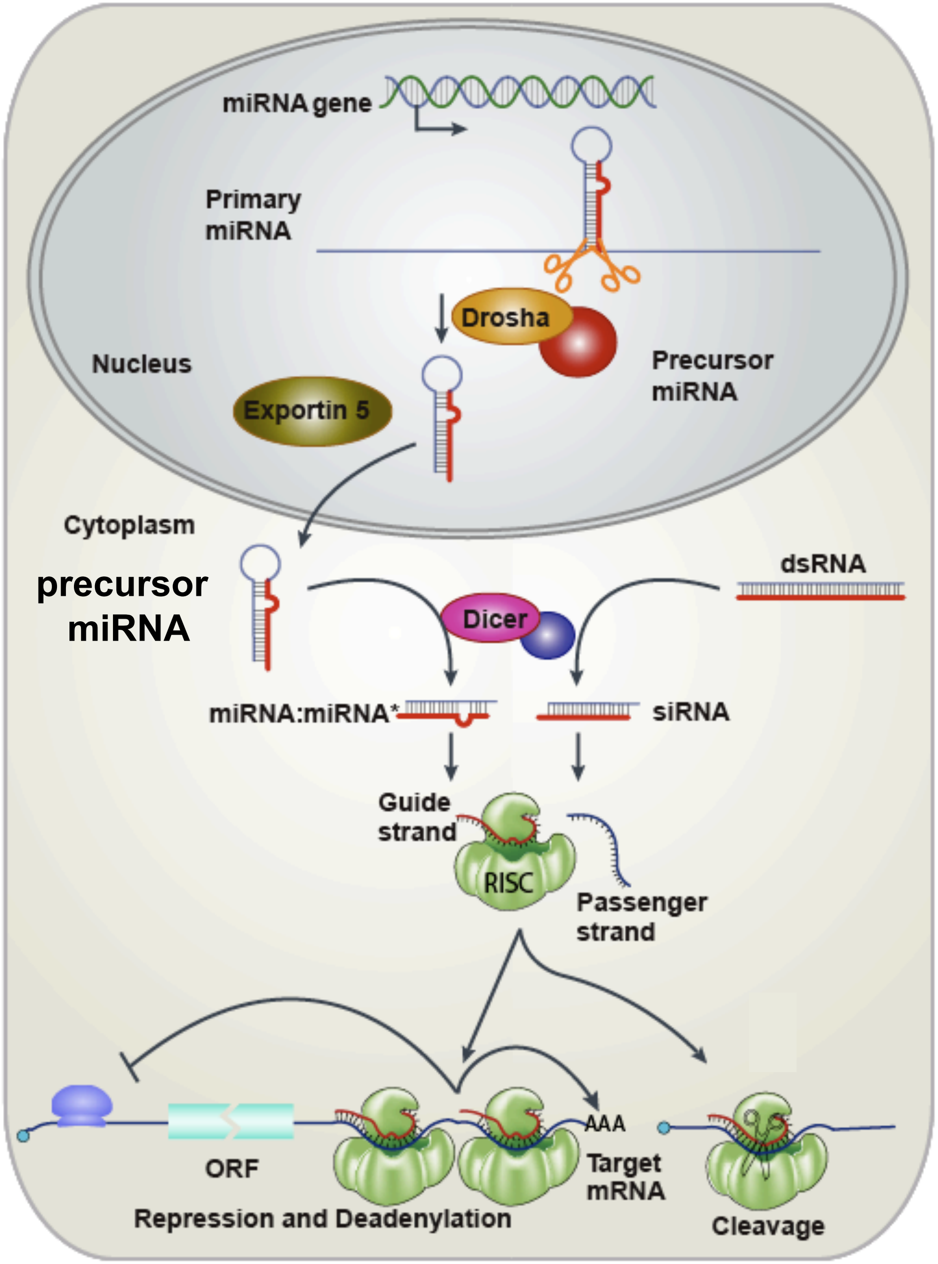
# microRNAs:

~22nt RNAs
Processed by Drosha
Processed from a hairpin by Dicer
Bound by Argonaute
Bind to the target 3’UTR
Repress translation
Induce mRNA deadenylation
Seed is important for specificity

Target mRNA

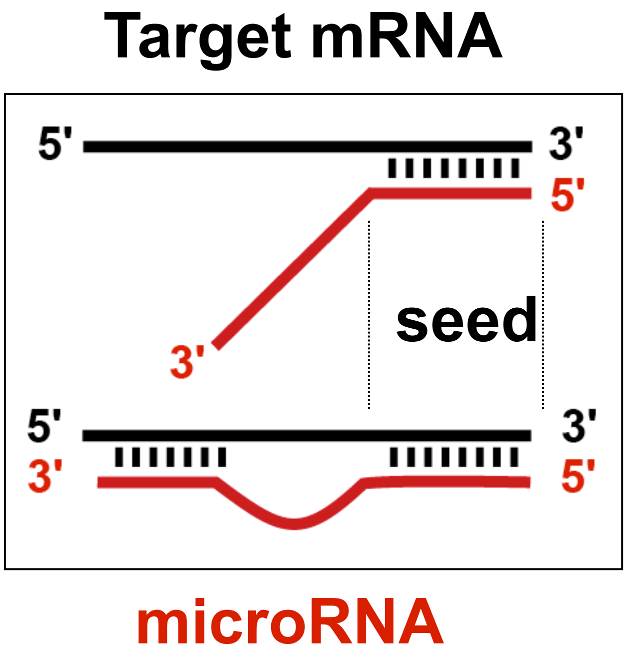


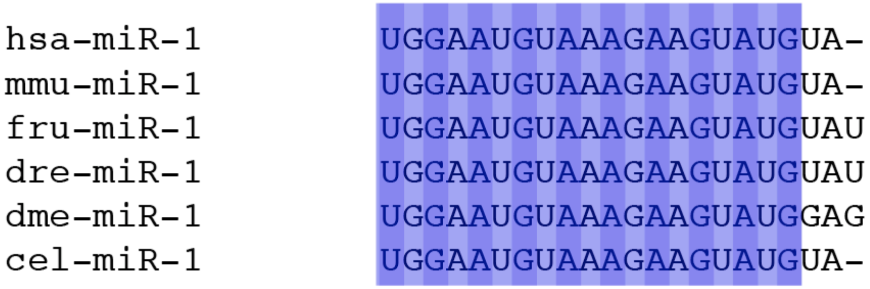
microRNA



# microRNAs:

~22nt RNAs
Processed by Drosha
Processed from a hairpin by Dicer
Bound by Argonaute
Bind to the target 3’UTR
Repress translation
Induce mRNA deadenylation
Seed is important for specificity





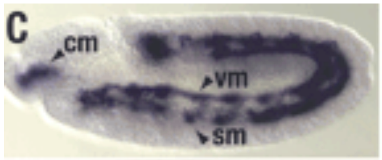
```
hsa-miR-1       UGGAAUGUAAAGAAGUAUGUA-
mmu-miR-1       UGGAAUGUAAAGAAGUAUGUA-
fru-miR-1       UGGAAUGUAAAGAAGUAUGUAU
dre-miR-1       UGGAAUGUAAAGAAGUAUGUAU
dme-miR-1       UGGAAUGUAAAGAAGUAUGGAG
cel-miR-1       UGGAAUGUAAAGAAGUAUGUA-
```

Conserved

# microRNAs:

**~22nt RNAs**
**Processed by Drosha**
**Processed from a hairpin by Dicer**
**Bound by Argonaute**
**Bind to the target 3’UTR**
**Repress translation**
**Induce mRNA deadenylation**
**Seed is important for specificity**

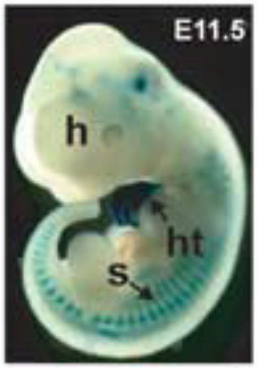
## miR-1/206



Drosophila

zebrafish

chicken



mouse

**Sokol et al., G&D 2006**

**Wienholds et al., Science 2005**

# Approaches to target identification

Genetics
Computational
Experimental

# Approaches to target identification

## Contribution of Genetics

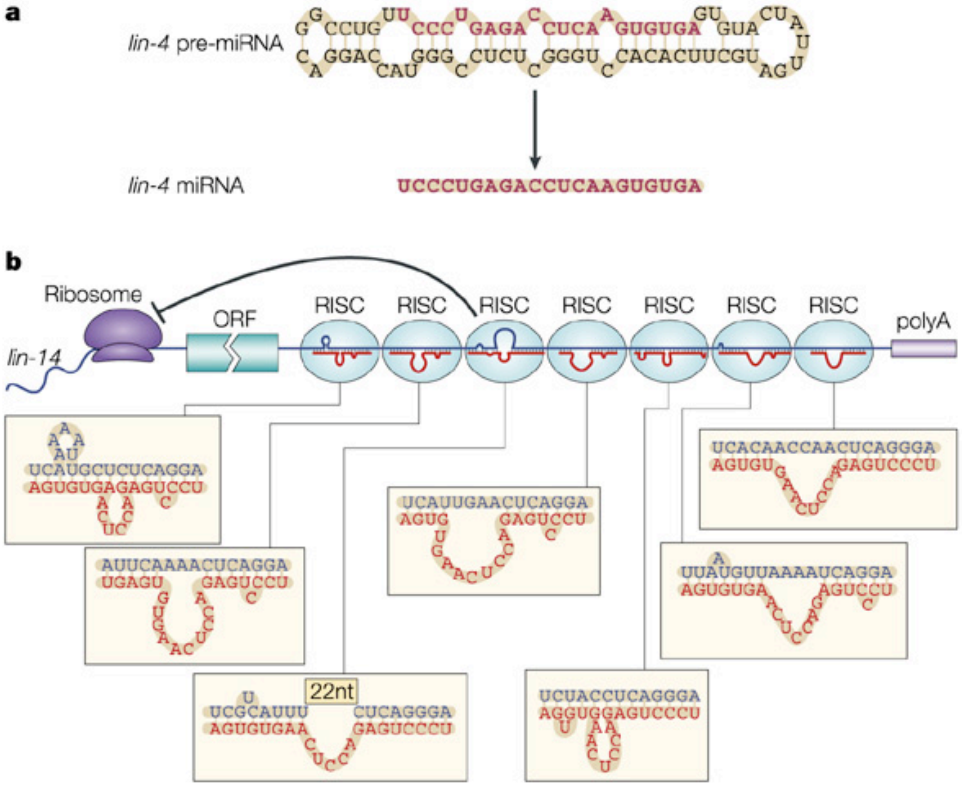
Lin-4 repressed lin-14
Heterochronic mutants in C. elegans
lin-4 encodes for two RNAs ~60 and ~22nt
Regulated lin-14 through its 3’UTR (antisense sites to lin-4)
Translational repression. mRNA levels unaffected.
GOF mutations for lin-14 delete lin-4 binding sites on it
Lin4 acted as a repressor of lin-14

**Ambros, Ruvkun, Lai and Slack**

# Approaches to target identification

## Contribution of Genetics

Lin-4 repressed lin-14



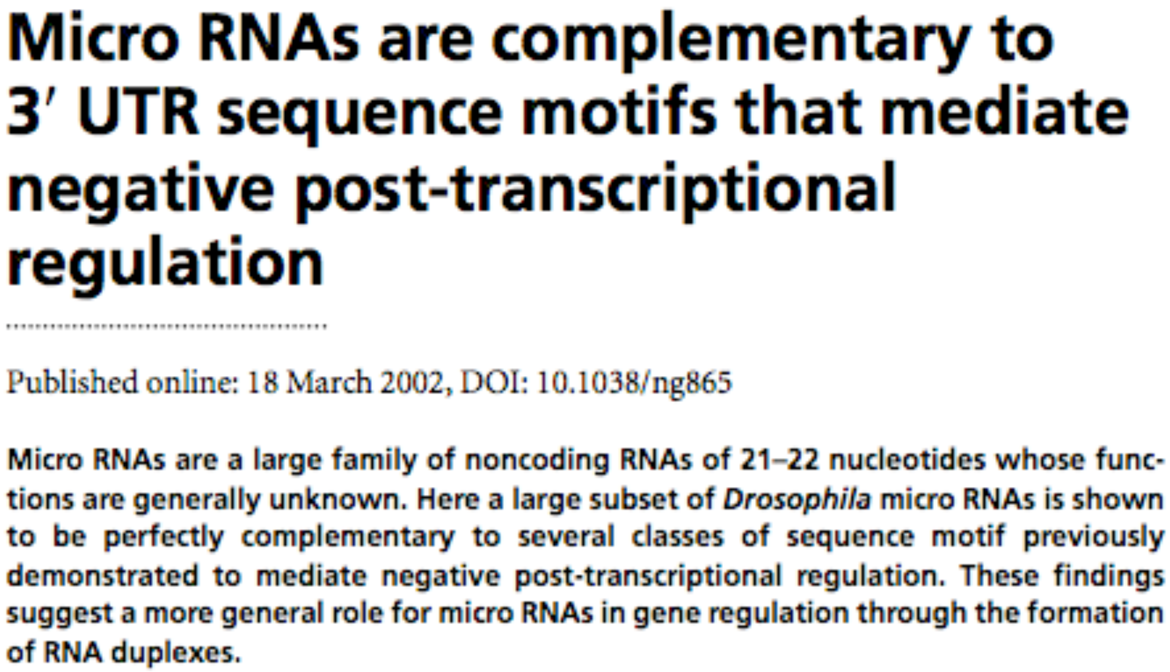
Nature Reviews | Genetics

**Ambros, Ruvkun, Lai and Slack**

# Approaches to target identification

## Contribution of Genetics

Drosophila:



### Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation

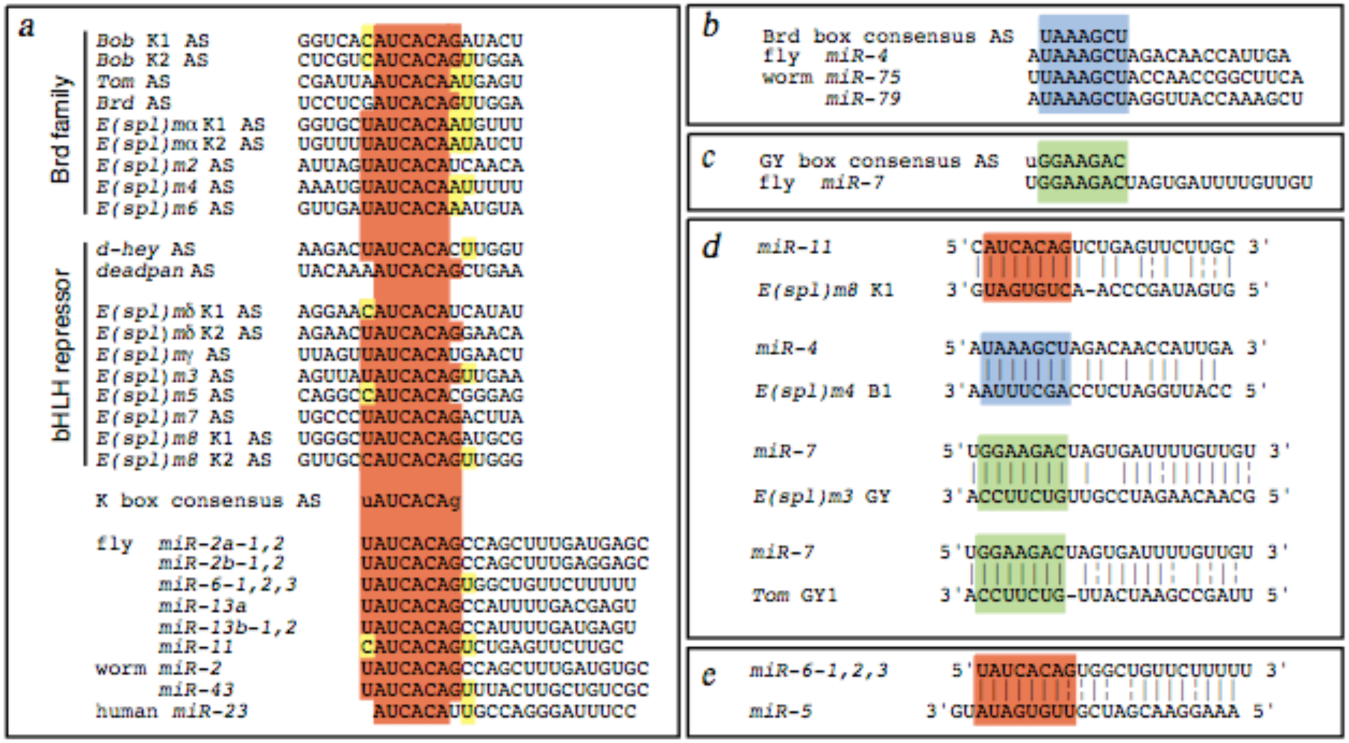
Micro RNAs are a large family of noncoding RNAs of 21–22 nucleotides whose functions are generally unknown. Here a large subset of *Drosophila* micro RNAs is shown to be perfectly complementary to several classes of sequence motif previously demonstrated to mediate negative post-transcriptional regulation. These findings suggest a more general role for micro RNAs in gene regulation through the formation of RNA duplexes.

**Lai**

# Approaches to target identification

## Contribution of Genetics

Drosophila: motifs in some mRNAs that caused post-transcriptional repression of these mRNAs
These motifs were complementary to some miRNAs 5’ end



**Lai**

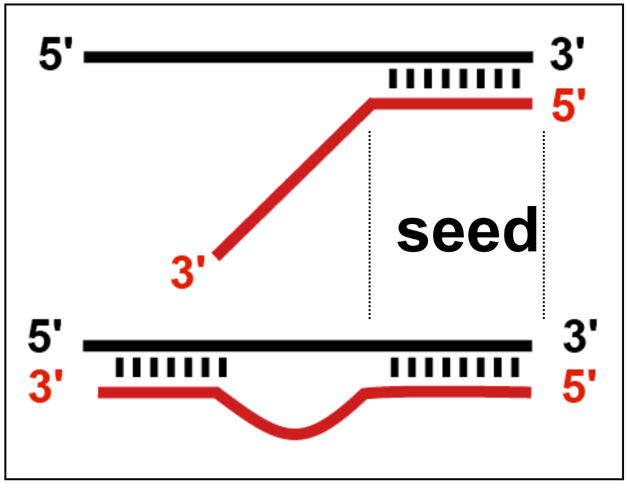
# Basic Rules for miRNA target recognition

Plants: “Perfect” sequence complementarity

Animals: 3’UTR (but also can happen in ORF or 5’UTR)

Seed is the major determinant

Target mRNA



microRNA

Mutagenesis defined the seed

# Basic Rules for miRNA target recognition

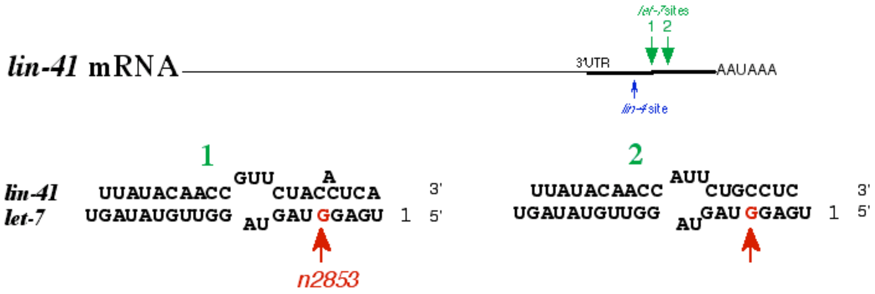
Plants: "Perfect" sequence complementarity

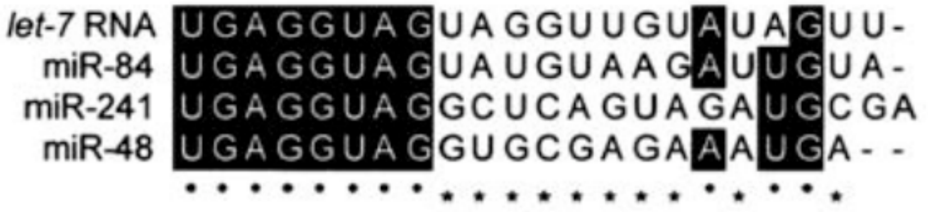
Animals: 3'UTR (but also can happen in ORF or 5'UTR)

Seed is the major determinant

Exceptions:

Let 7 regulation of lin-41





Slack 2004, Ambros 2005

# Approaches to target identification

Genetics
Computational
Experimental

# Approaches to target identification

## Computational approaches

Look at sequence complementary to miRNAs

Reinforced the idea of the seed

Too many sites: Use of conservation

Performance of a sliding window over the miRNA
Compared to a shuffled miRNA, keeping dinucleotide bias

This approach identified >100 target mRNAs per miRNA

**Bartel, Rajewsky, Cohen, Sander, Gerstein…**

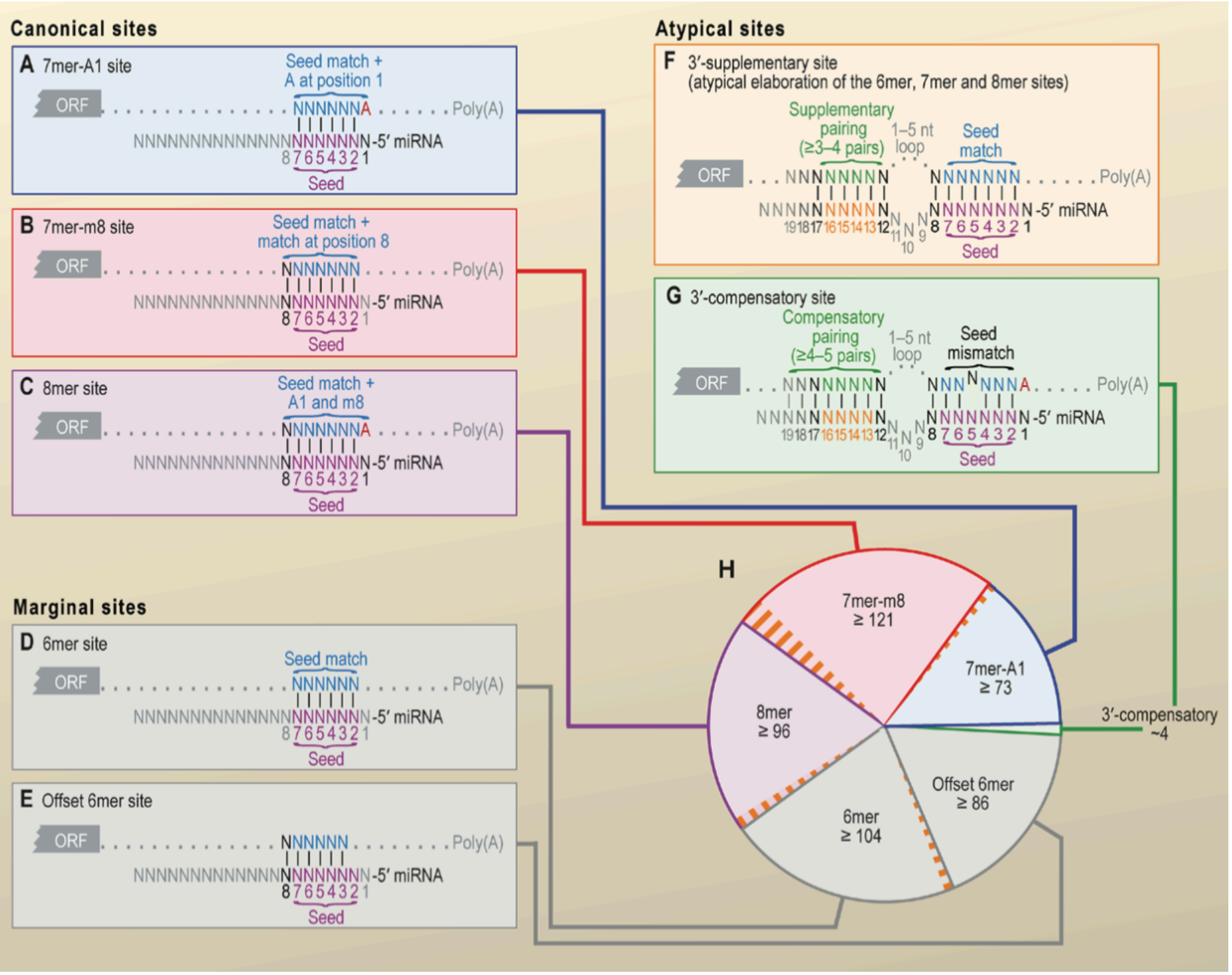
# Approaches to target identification

## Computational approaches

## Seed is a major determinant

Specificity

Sensitivity



Bartel 2009

# Approaches to target identification

## Experimental approaches

**Transfect miRNAs in HeLa Cells**

**Analyze the gene expression profile by microarrays**

**mRNAs downregulated in the presence of the miRNA**

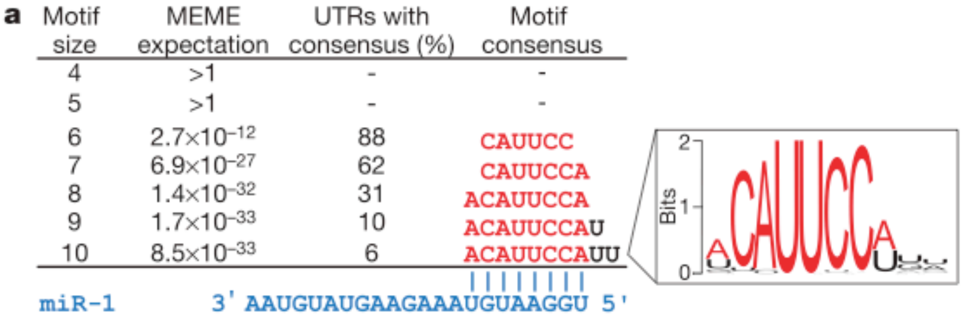
Lim 2005

# Approaches to target identification

## Experimental approaches

**Motif enriched in the genes downregulated was complementary to the miRNA seed.**

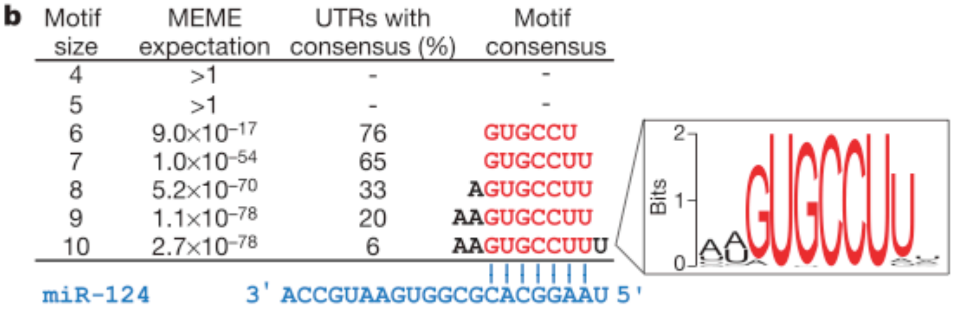
**But these targets were not preferentially conserved**



**a**

| Motif size | MEME expectation | UTRs with consensus (%) | Motif consensus |
|---|---|---|---|
| 4 | >1 | - | - |
| 5 | >1 | - | - |
| 6 | $2.7\times10^{-12}$ | 88 | CAUUCC |
| 7 | $6.9\times10^{-27}$ | 62 | CAUUCCA |
| 8 | $1.4\times10^{-32}$ | 31 | ACAUUCCA |
| 9 | $1.7\times10^{-33}$ | 10 | ACAUUCCAU |
| 10 | $8.5\times10^{-33}$ | 6 | ACAUUCCAUU |

miR-1 3' AAUGUAUGAAGAAAUGUAAGGU 5'



**b**

| Motif size | MEME expectation | UTRs with consensus (%) | Motif consensus |
|---|---|---|---|
| 4 | >1 | - | - |
| 5 | >1 | - | - |
| 6 | $9.0\times10^{-17}$ | 76 | GUGCCU |
| 7 | $1.0\times10^{-54}$ | 65 | GUGCCUU |
| 8 | $5.2\times10^{-70}$ | 33 | AGUGCCUU |
| 9 | $1.1\times10^{-78}$ | 20 | AAGUGCCUU |
| 10 | $2.7\times10^{-78}$ | 6 | AAGUGCCUUU |

miR-124 3' ACCGUAAGUGGCGCACGGAAU 5'

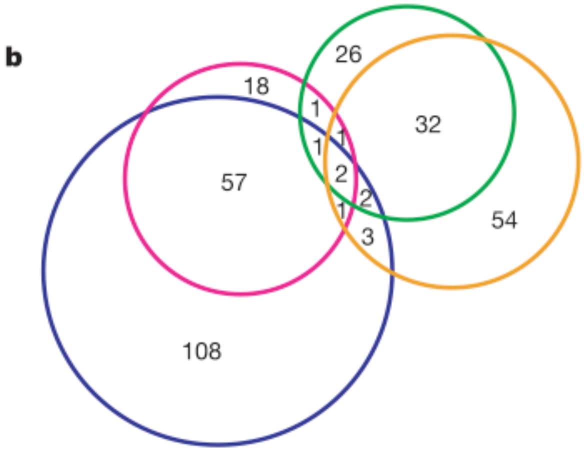
Lim 2005

# Approaches to target identification

## Experimental approaches

Generated chimeric miRNAs

5’ end is most important for target regulation



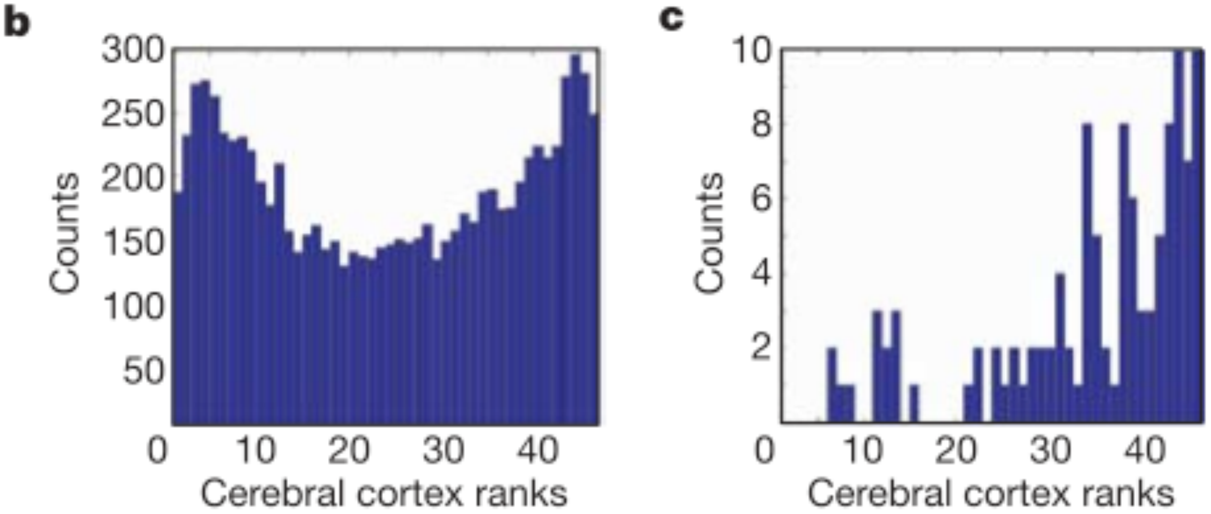
| | |
|---|---|
| miR-1 | UGGAAUGUAAAGAAGUAUGUAA |
| miR-124 | UAAGGCACGCGGUGAAUGCCA |
| chimiR-124/1 | UAAGGCACGCAGAAGUAUGUAA |
| chimiR-1/124 | UGGAAUGUAAGGUGAAUGCCA |

Lim 2005

# Approaches to target identification

## Experimental approaches

## What is the correlation with gene expression in vivo?



**Targets of a brain miRNA tend to be expressed lower in The brain compared to other tissues.**

**Caveat: Remember experiment done in HeLa cells**

Lim 2005

# Approaches to target identification

## Experimental/computational approaches

## What about the genes that are highly expressed in one tissue?

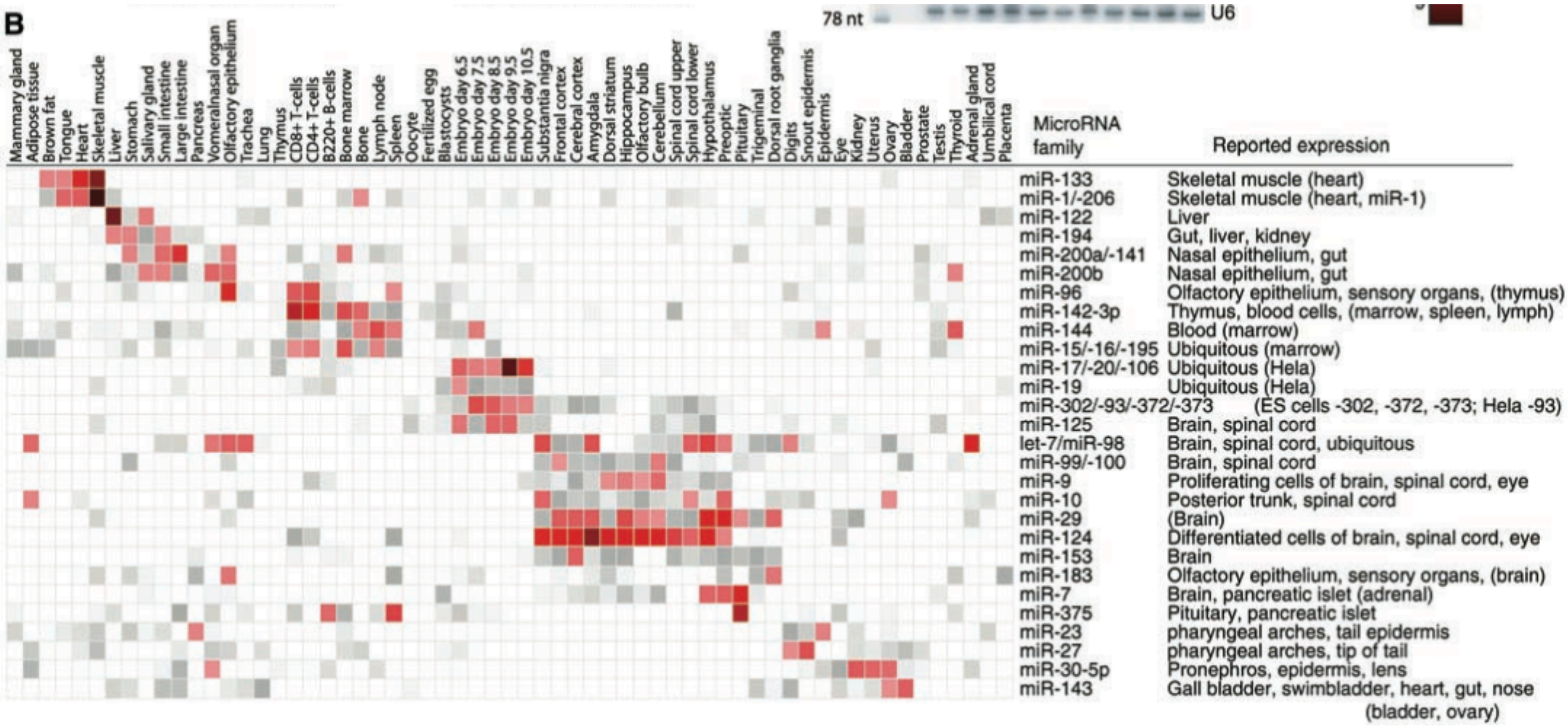
Concept of target avoidance: mRNAs highly expressed in one tissue tend to lack miRNA target sites for miRNAs expressed in that tissue

Lim 2005

# Approaches to target identification

## Experimental/computational approaches

## What about the genes that are highly expressed in one tissue?



| MicroRNA family | Reported expression |
|---|---|
| miR-133 | Skeletal muscle (heart) |
| miR-1/-206 | Skeletal muscle (heart, miR-1) |
| miR-122 | Liver |
| miR-194 | Gut, liver, kidney |
| miR-200a/-141 | Nasal epithelium, gut |
| miR-200b | Nasal epithelium, gut |
| miR-96 | Olfactory epithelium, sensory organs, (thymus) |
| miR-142-3p | Thymus, blood cells, (marrow, spleen, lymph) |
| miR-144 | Blood (marrow) |
| miR-15/-16/-195 | Ubiquitous (marrow) |
| miR-17/-20/-106 | Ubiquitous (Hela) |
| miR-19 | Ubiquitous (Hela) |
| miR-302/-93/-372/-373 | (ES cells -302, -372, -373; Hela -93) |
| miR-125 | Brain, spinal cord |
| let-7/miR-98 | Brain, spinal cord, ubiquitous |
| miR-99/-100 | Brain, spinal cord |
| miR-9 | Proliferating cells of brain, spinal cord, eye |
| miR-10 | Posterior trunk, spinal cord |
| miR-29 | (Brain) |
| miR-124 | Differentiated cells of brain, spinal cord, eye |
| miR-153 | Brain |
| miR-183 | Olfactory epithelium, sensory organs, (brain) |
| miR-7 | Brain, pancreatic islet (adrenal) |
| miR-375 | Pituitary, pancreatic islet |
| miR-23 | pharyngeal arches, tail epidermis |
| miR-27 | pharyngeal arches, tip of tail |
| miR-30-5p | Pronephros, epidermis, lens |
| miR-143 | Gall bladder, swimbladder, heart, gut, nose (bladder, ovary) |

Farh 2005

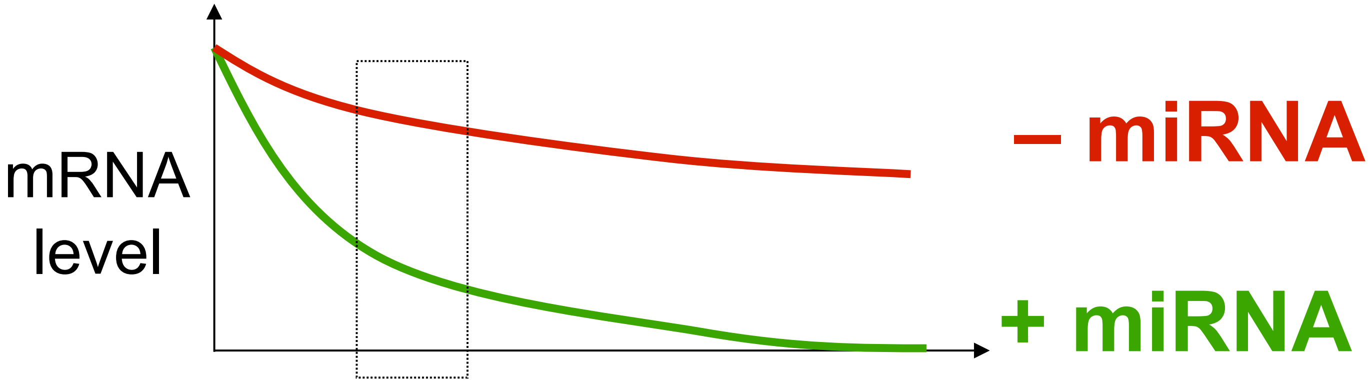
# microRNAs accelerate target mRNA degradation

## How can we identify physiological miRNA targets?

**miRNA**

**REPRESSION**

**mRNA DECAY**



# microRNAs accelerate target mRNA degradation

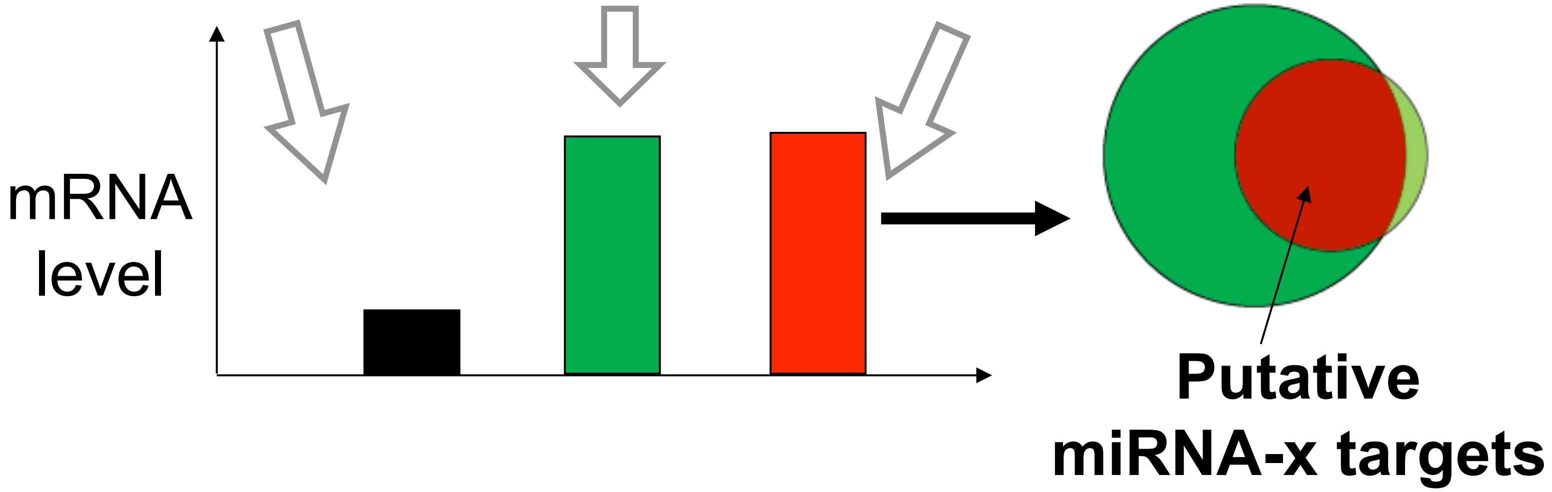
## What about the genes that are highly expressed in one tissue?

**miRNA**

**REPRESSION** **mRNA DECAY**

**Wild type** **Mutant** **Rescued**

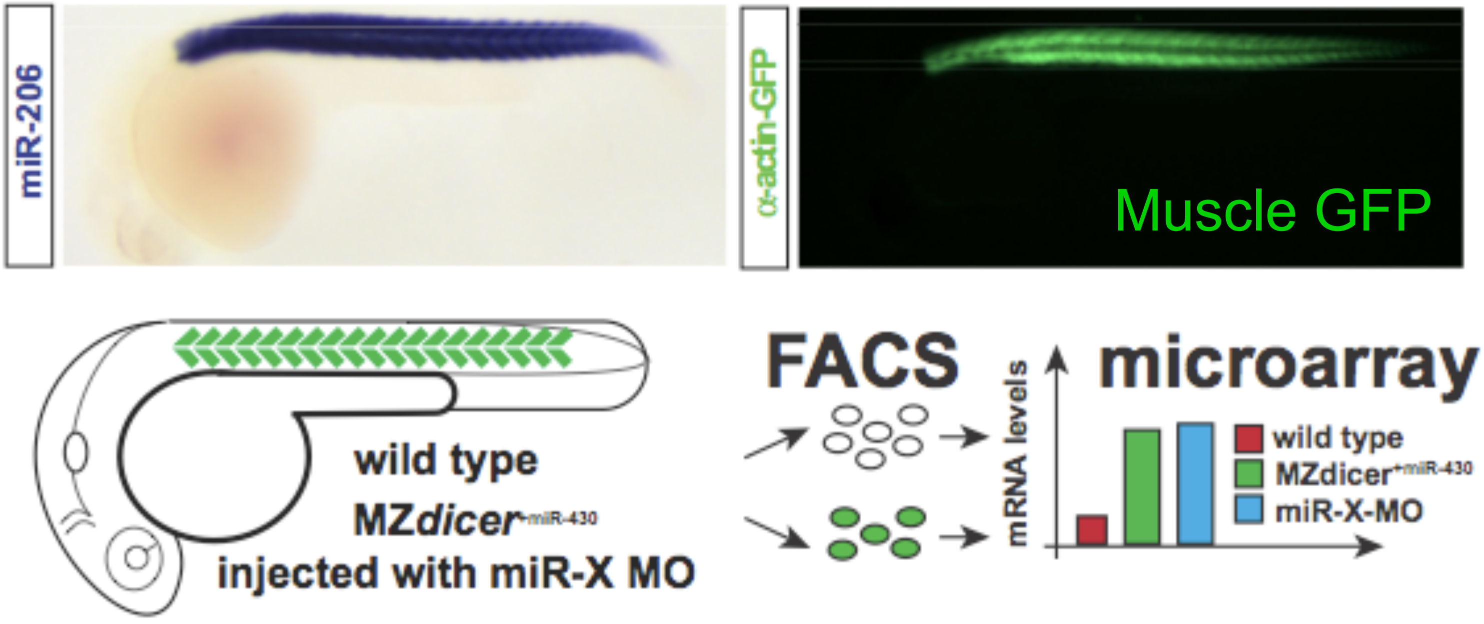
**wt miRNA** **– all miRNAs** **- miRNA x**



# What are the targets of miRNAs in muscle?

**miRNA**

**REPRESSION** **mRNA DECAY**

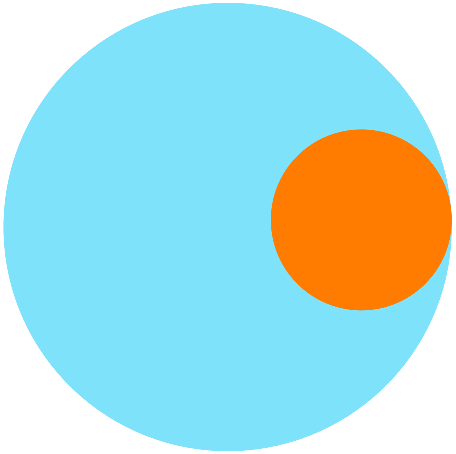


**Genes upregulated in the absence of the microRNAs**

Mishima et al G&D 2009

# Compare wild type vs. MZ*dicer*$^{+miR\text{-}430}$ muscle

Total : 6825 mRNAs (with UTR information)



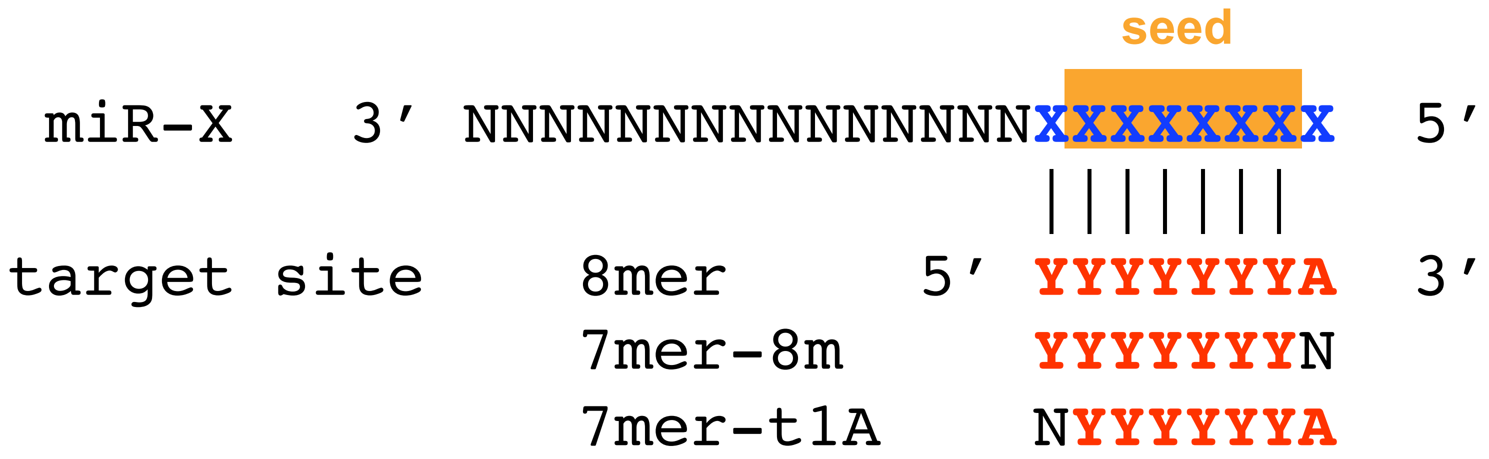
Upregulated in MZdicer:
666 mRNAs

( > 1.3 fold upregulated)

**Genes upregulated in muscle depleted of microRNAs**

# miRNA target site analysis

miRNA target site (TS) = complementary to the seed (7mer - 8mer)

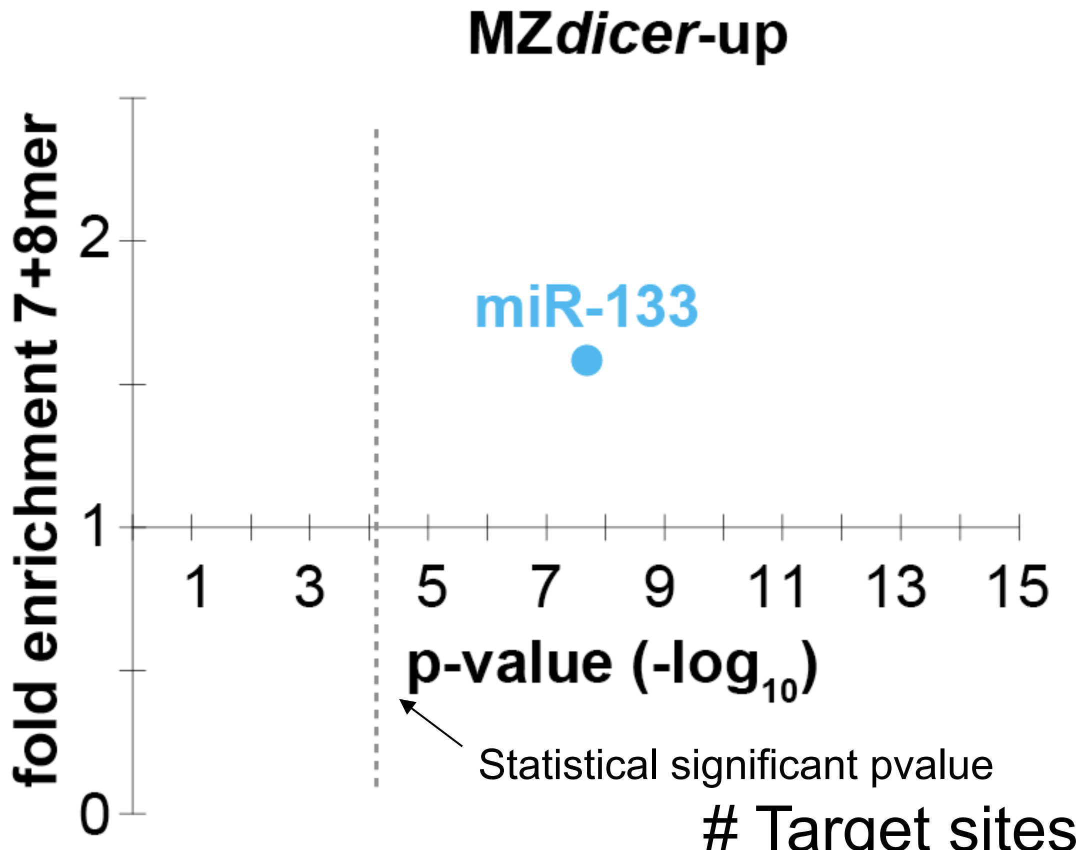


```
 miR-X    3' NNNNNNNNNNNNNNNXXXXXXXX   5'
                           |||||||
target site    8mer     5' YYYYYYYA   3'
               7mer-8m     YYYYYYYN
               7mer-t1A    NYYYYYYA
```

**1) TS preferentially enriched = miR-X is active**

**2) target genes for miR-X = upregulated genes with TS**

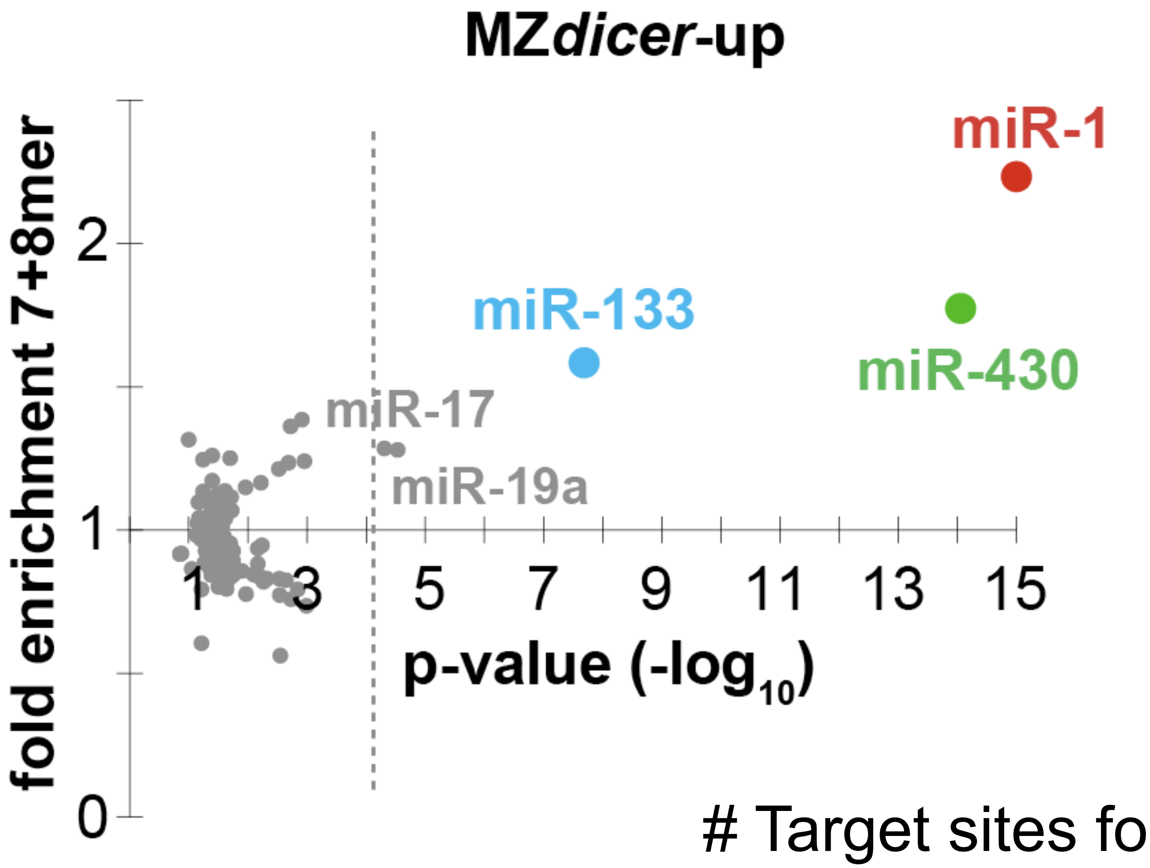
# miR-1/206, miR-133 and miR-430 TS are enriched in genes upregulated in MZdicer muscle


**MZ*dicer*-up**

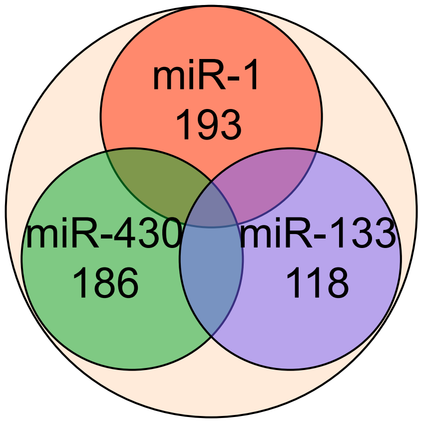
Statistical significant pvalue

$$\text{Fold enrichment} = \frac{\dfrac{\#\ \text{Target sites for mR-X}}{\#\ \text{Target sites for all miRNAs}} \quad \text{experiment}}{\dfrac{\#\ \text{Target sites for mR-X}}{\#\ \text{Target sites for all miRNAs}} \quad \text{control}}$$

# miR-1/206, miR-133 and miR-430 TS are enriched in genes upregulated in MZdicer muscle

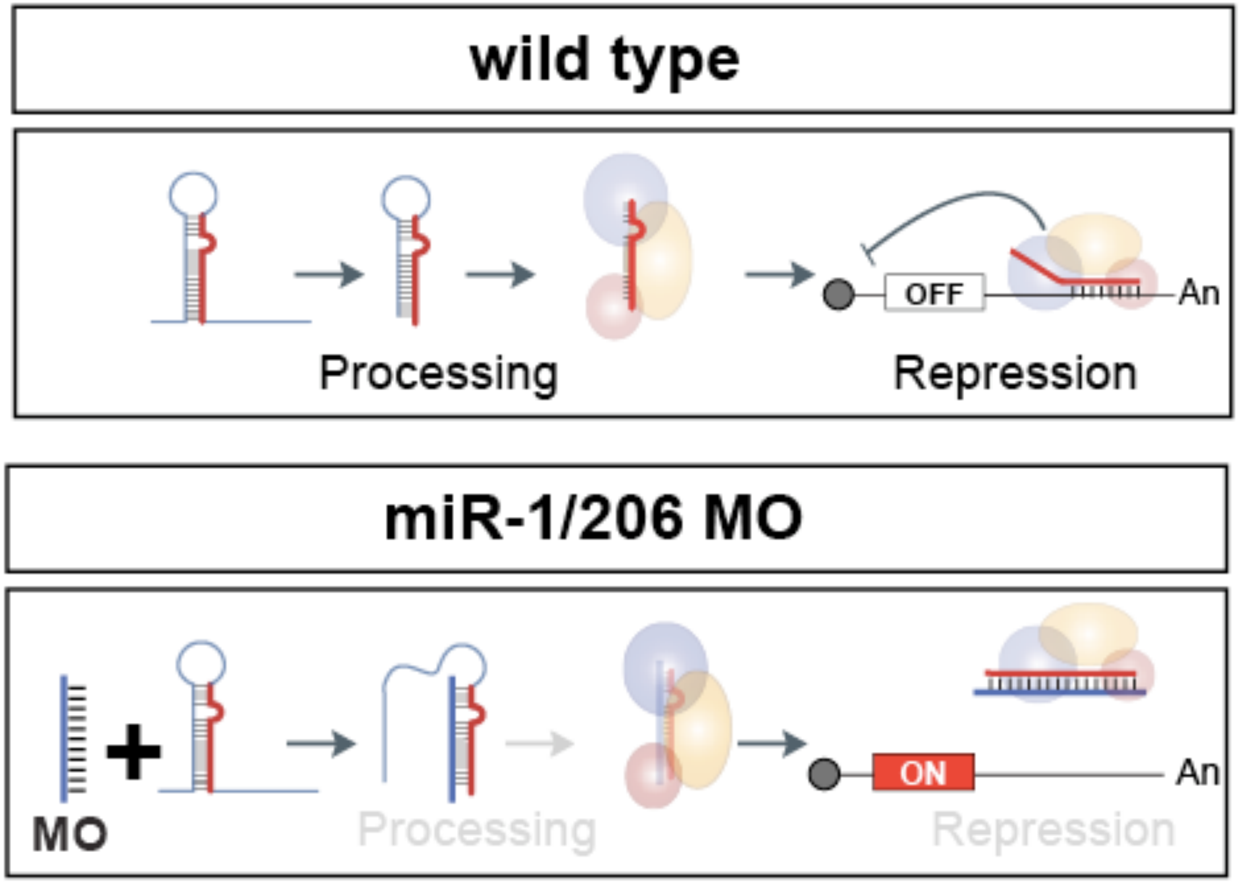


Up in MZdicer 666 mRNAs
with TS 274 mRNA



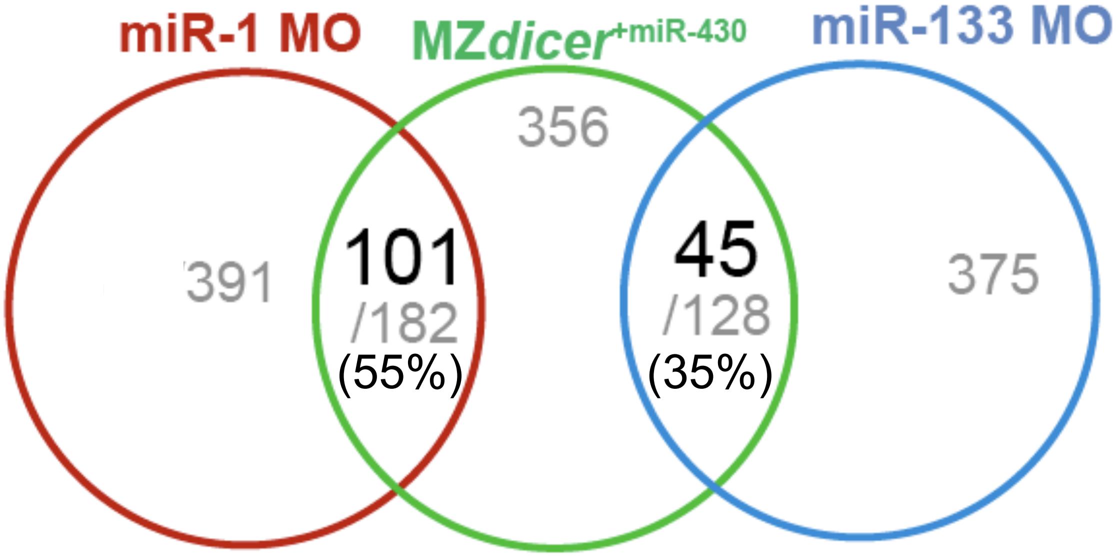
$$\text{Fold enrichment} = \frac{\dfrac{\#\text{ Target sites for mR-X}}{\#\text{ Target sites for all miRNAs}}\quad\text{experiment}}{\dfrac{\#\text{ Target sites for mR-X}}{\#\text{ Target sites for all miRNAs}}\quad\text{control}}$$

# Inhibition of individual miRNAs using antisense MO

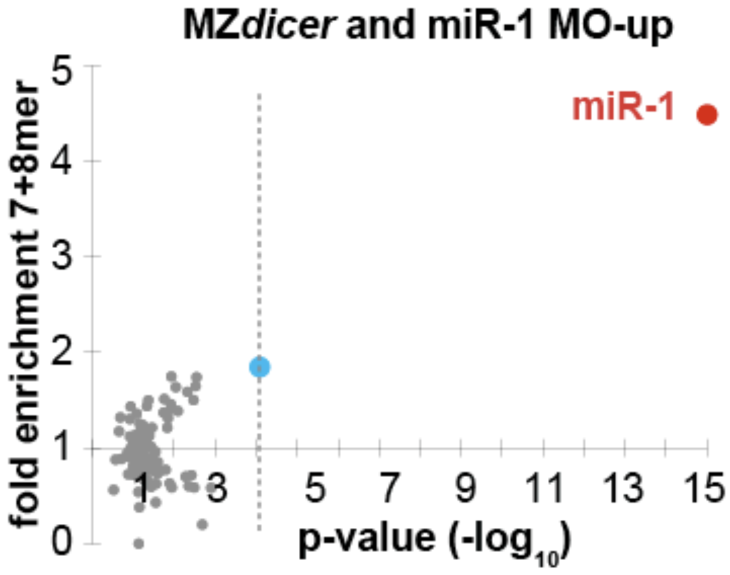


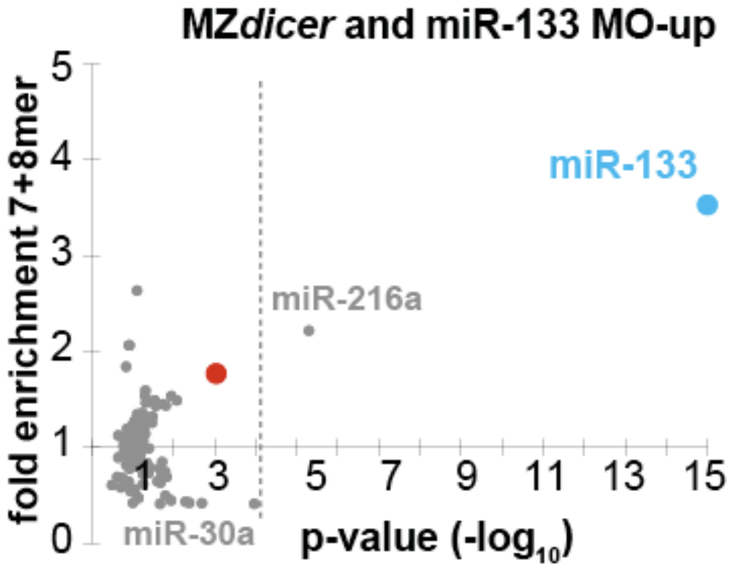
# Single miRNA TS are enriched in genes upregulated in MO injected embryos and MZdicer

Control
892/6825
(13%)

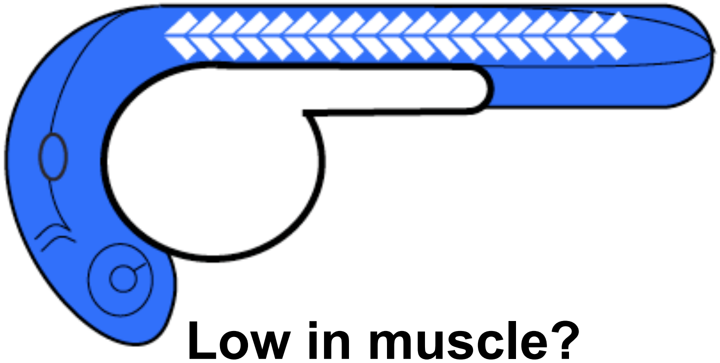


Control
722/6825
(11%)


**MZ*dicer* and miR-1 MO-up**

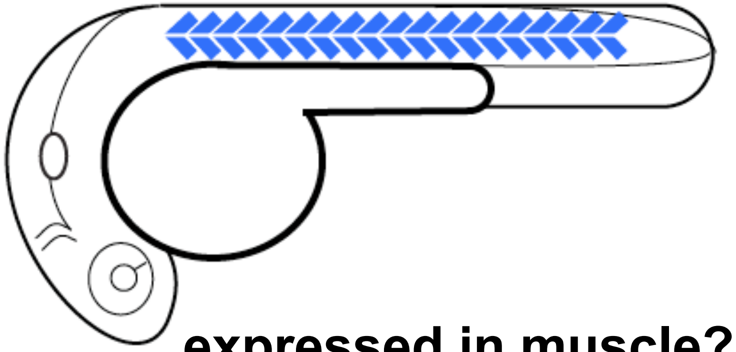

**MZ*dicer* and miR-133 MO-up**

# What is the expression pattern of muscle miRNA targets?



Low in muscle?

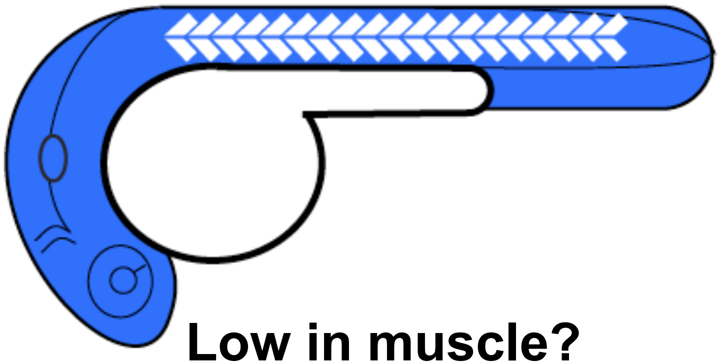
GFP VS. **non-GFP**

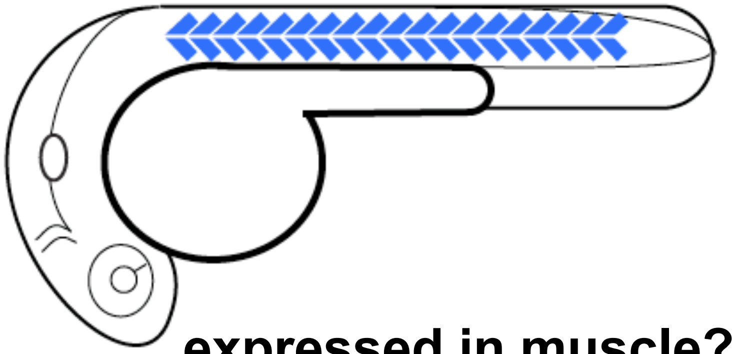


expressed in muscle?

**GFP** VS. non-GFP

# Muscle miRNA targets tend to be expressed at lower levels in muscle



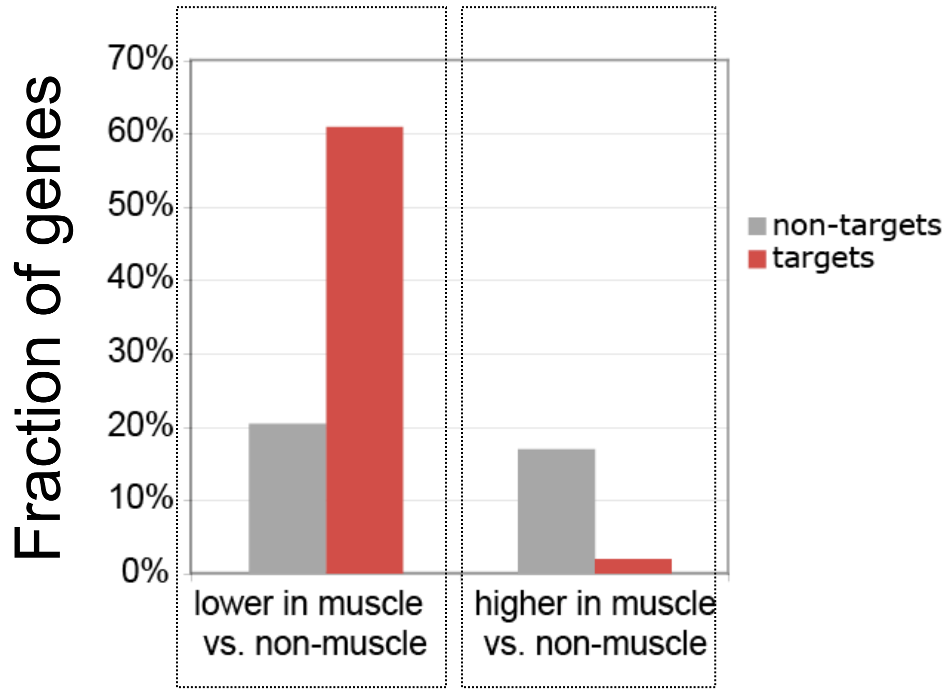


**Low in muscle?**

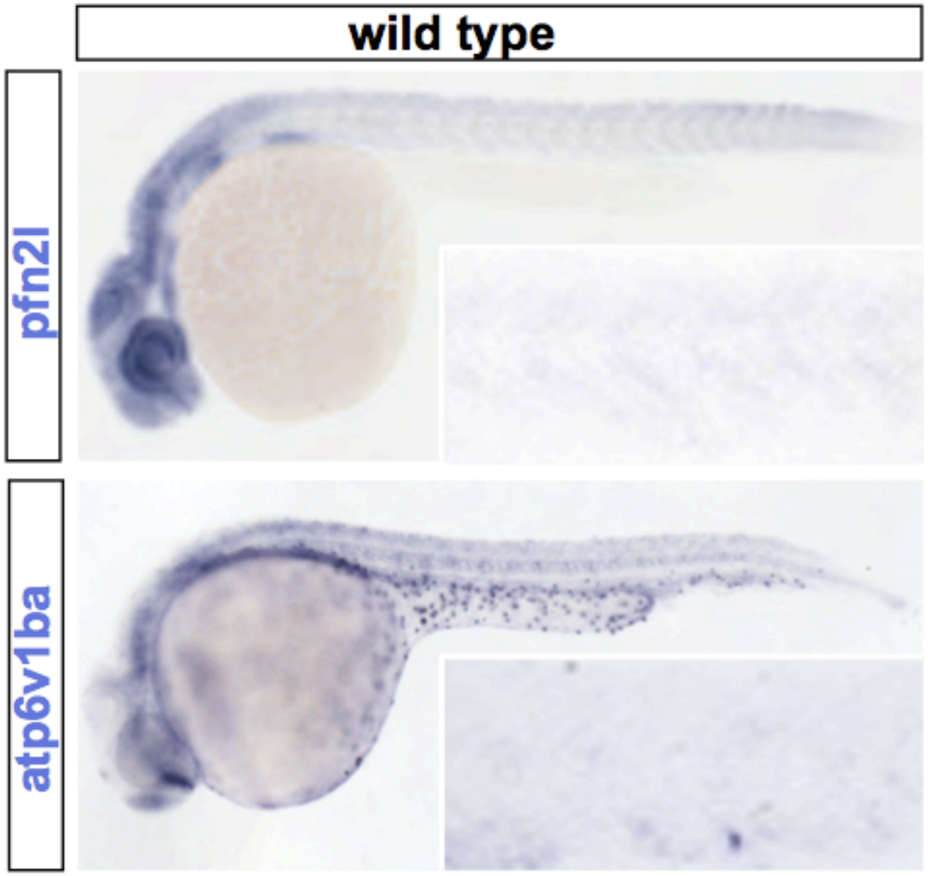
**expressed in muscle?**

GFP vs. **non-GFP**

**GFP** vs. non-GFP



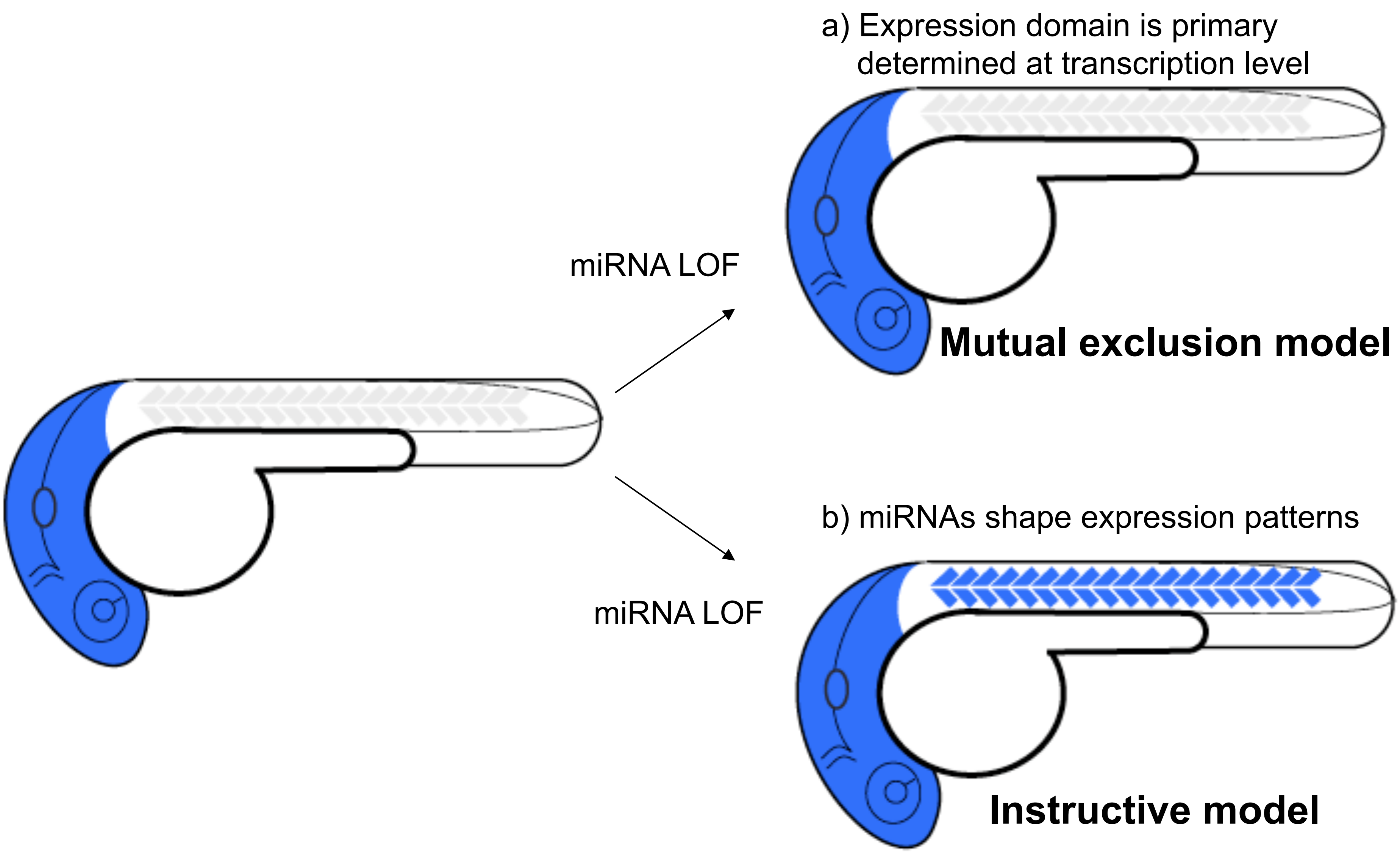
# miRNA targets are low in muscle



miRNA and targets are expressed in mutually exclusive manner

(Stark et al 2005, Farh et al 2005)

# Do miRNAs influence the expression pattern of the targets?



a) Expression domain is primary
determined at transcription level

miRNA LOF

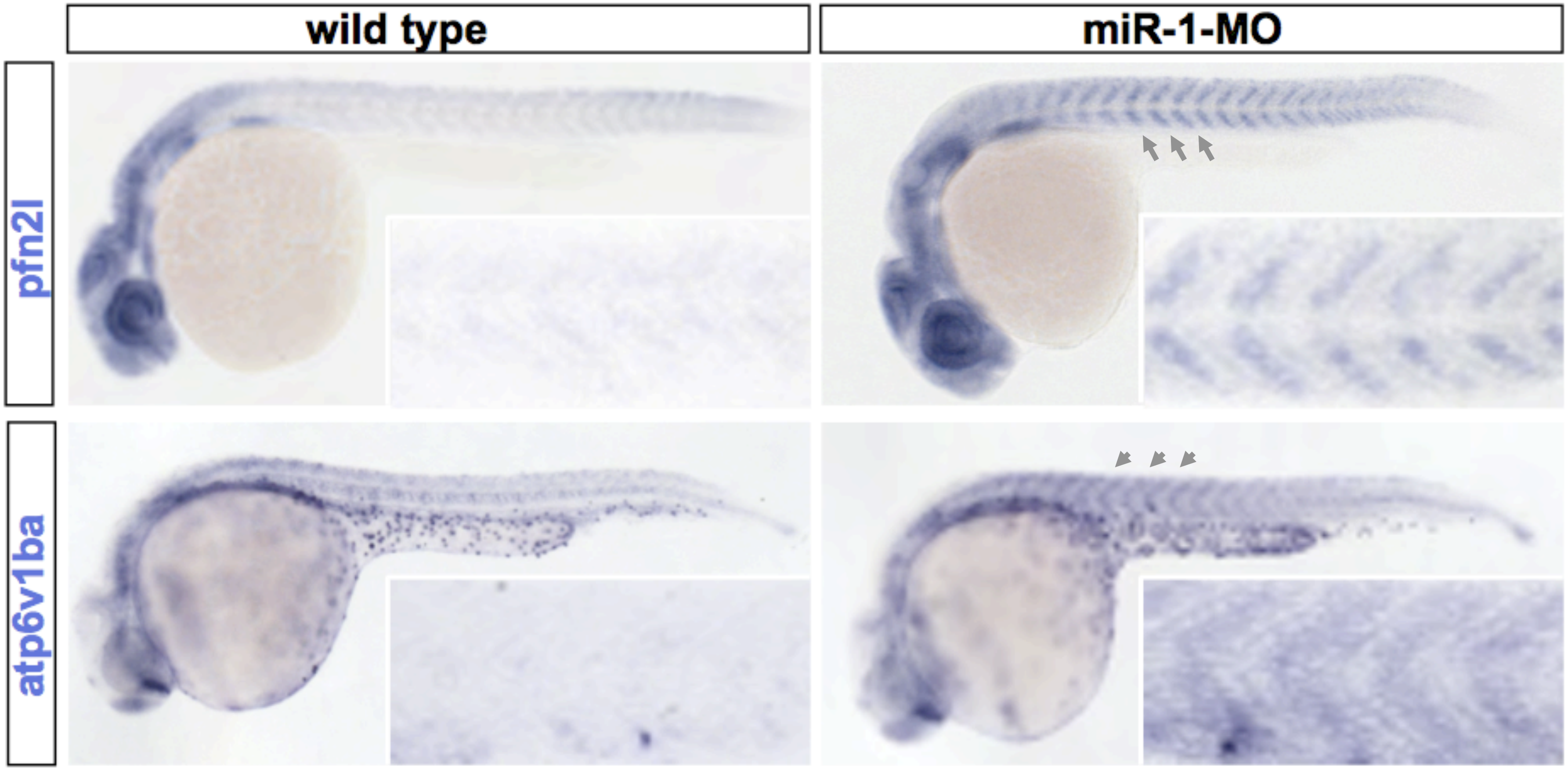
**Mutual exclusion model**

miRNA LOF

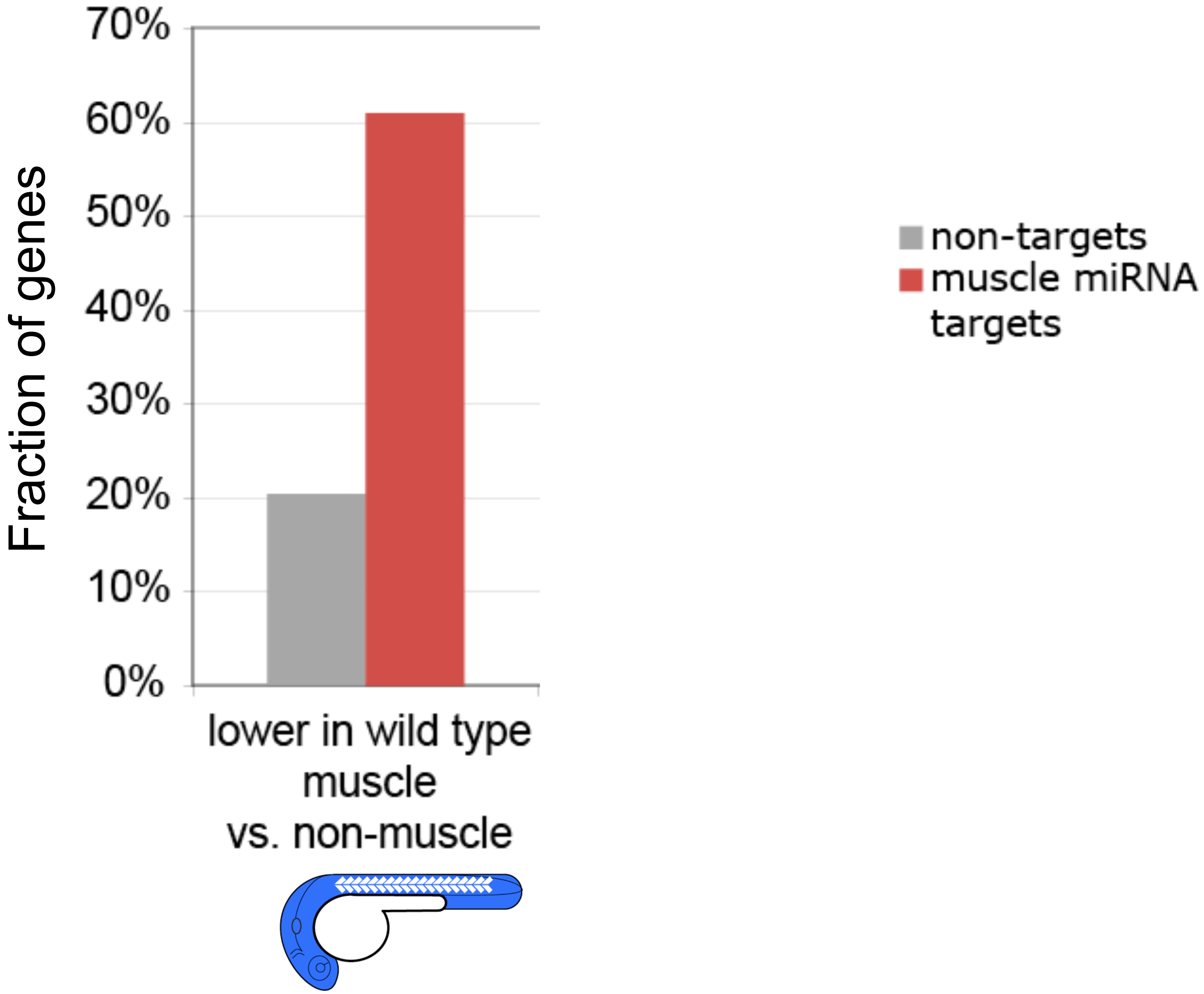
b) miRNAs shape expression patterns

**Instructive model**

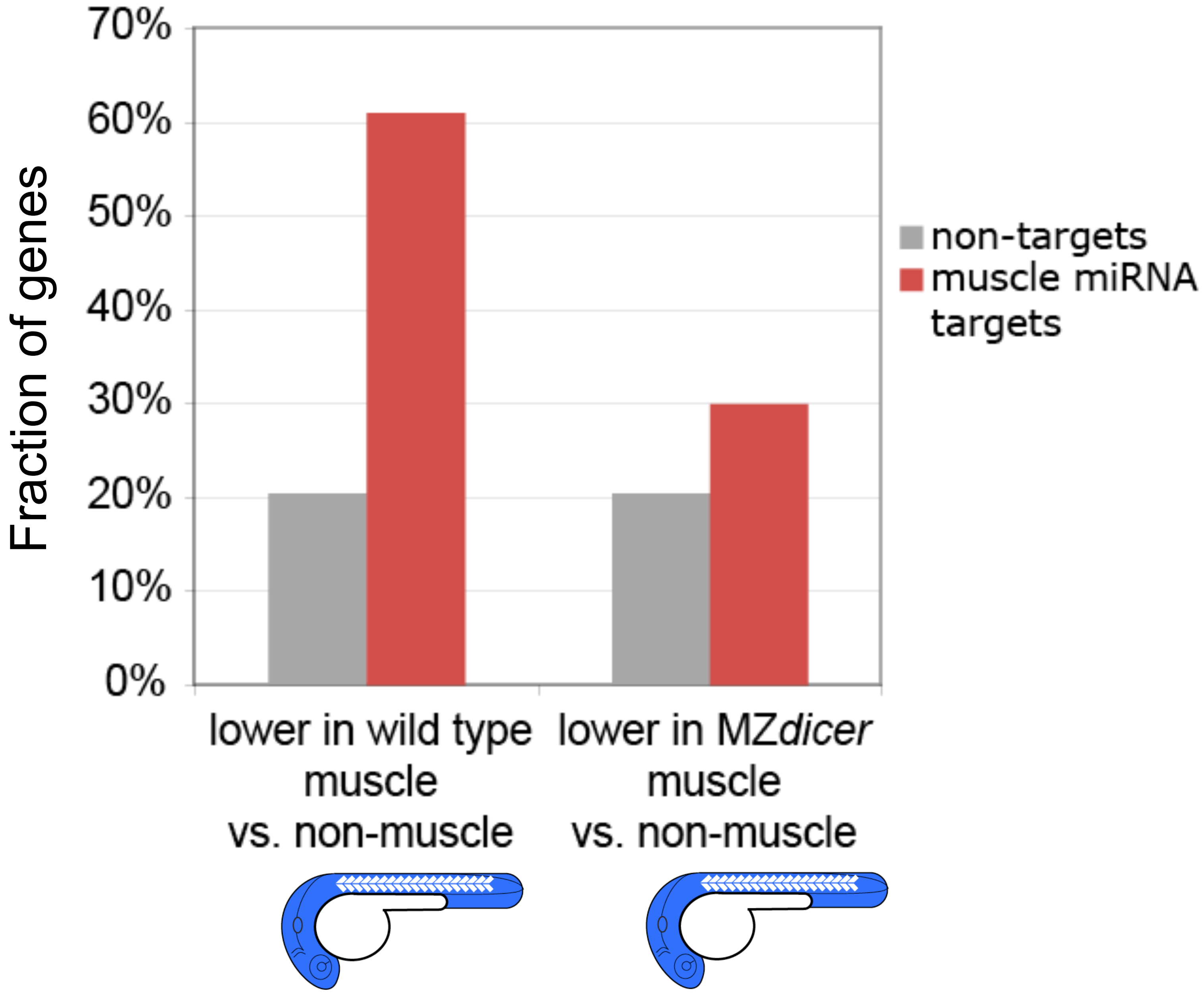
# miRNAs regulate embryonic gene expression patterns post-transcriptionally



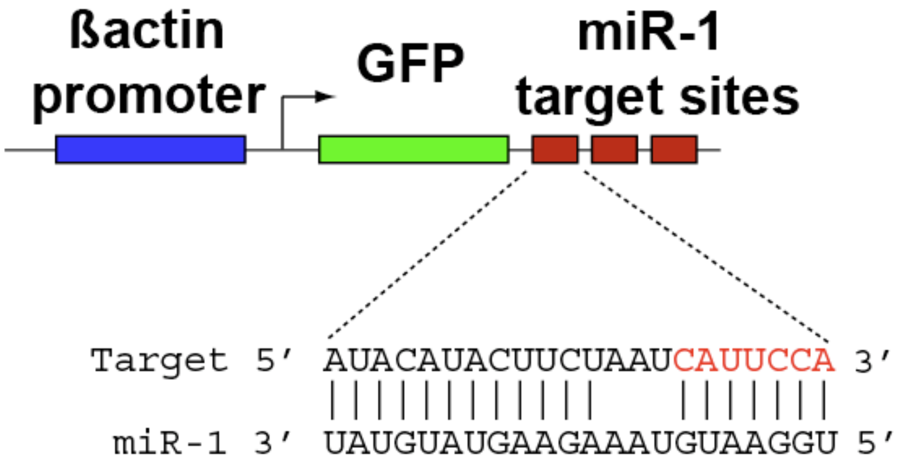
# Muscle miRNA targets tend to be lower in muscle vs non-muscle



# Muscle miRNA targets tend to be lower in muscle vs non-muscle
# This bias on gene expression depends largely on miRNAs



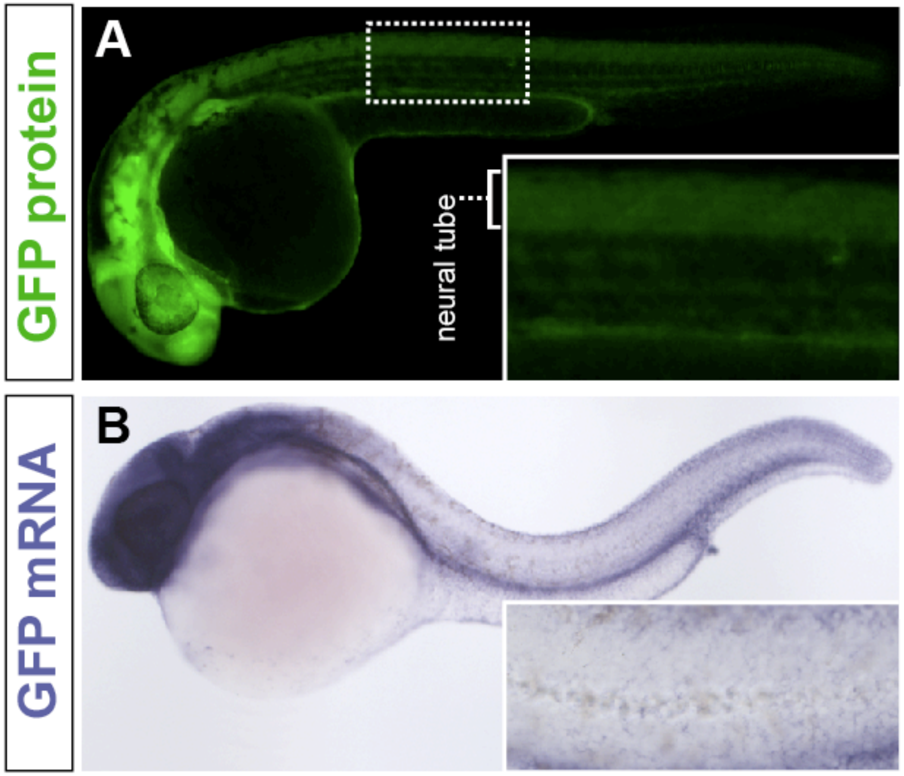
# Can miRNAs shape the expression of a ubiquitously transcribed gene?

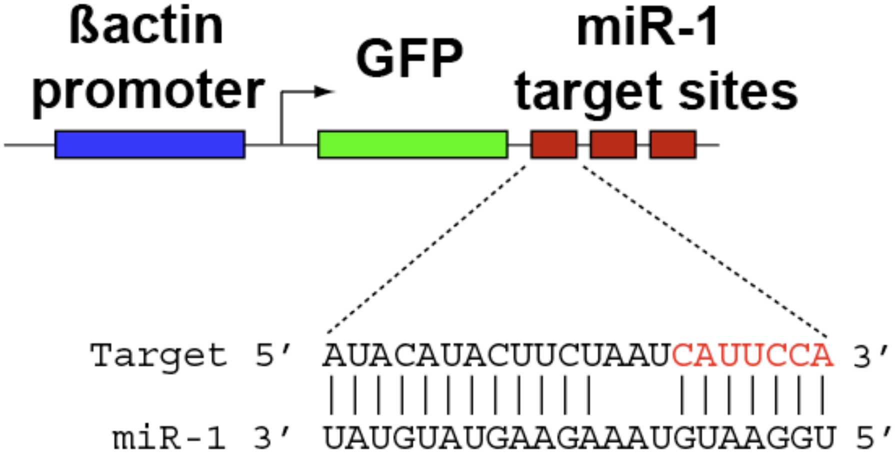


# miRNAs regulate embryonic gene expression patterns post-transcriptionally

**Tg: [β-actin-promoter]-GFP-3xIPT-miR-1**

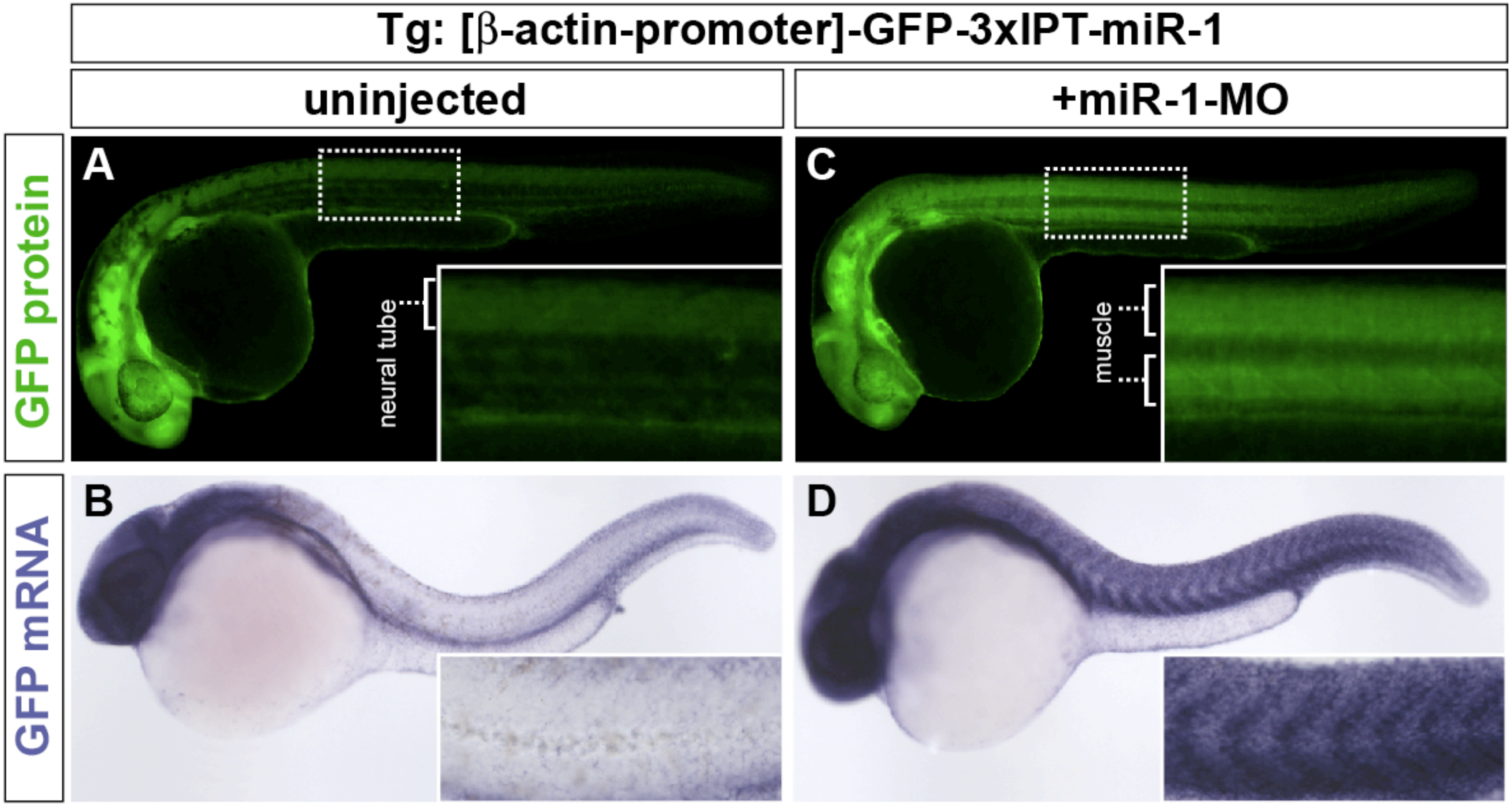
**uninjected**





```
Target 5′ AUACAUACUUCUAAUCAUUCCA 3′
          ||||||||||||    |||||||
 miR-1 3′ UAUGUAUGAAGAAAUGUAAGGU 5′
```

# miRNAs regulate embryonic gene expression patterns post-transcriptionally



# Basic Rules for miRNA target recognition

**Plants: “Perfect” sequence complementarity**

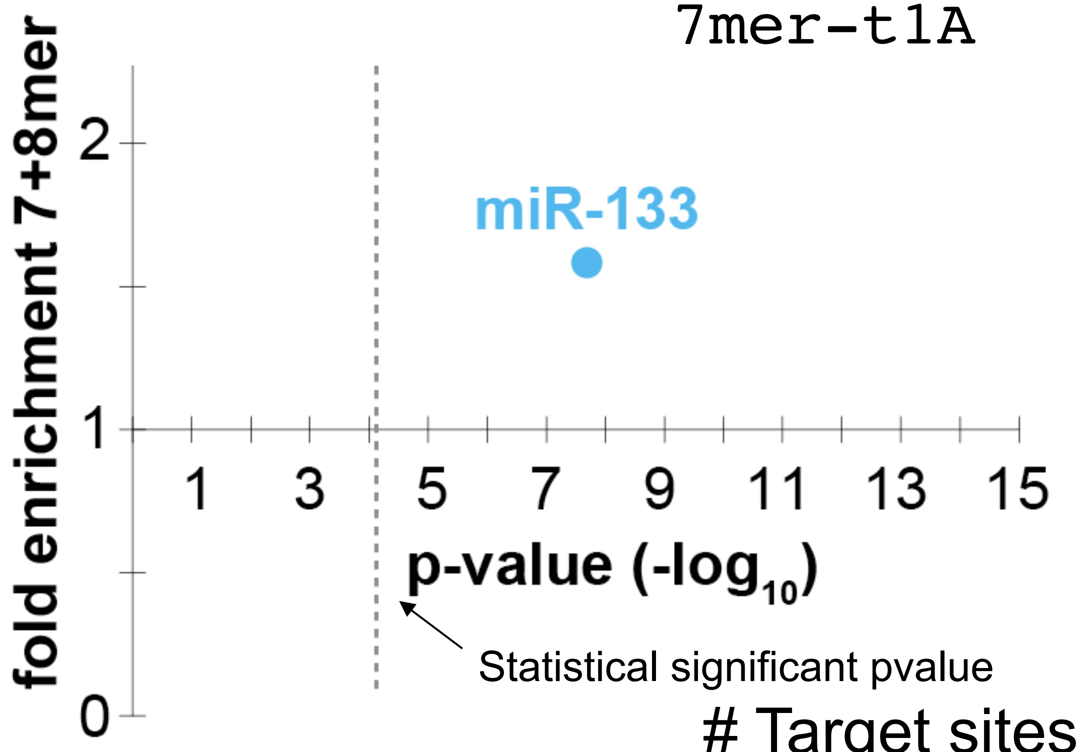
**Animals: 3’UTR (but also can happen in ORF or 5’UTR)**

**Seed is the major determinant**
**Exceptions:**
**3’UTR Context:**
**15nt away from stop codon**
**Away from the center in long 3’UTRs**
**AU rich sequences flanking site**
**Proximity of sites to = or ≠ miRNAs that are co-expressed**

**miRNAs are likely to influence 3’UTR evolution and gene expression patterns in the tissues.**

Bartel Cell 2009

```
miR-X    3′ NNNNNNNNNNNNNNNXXXXXXXX   5′
                           |||||||
target site     8mer      5′ YYYYYYYA   3′
                7mer-8m      YYYYYYYN
                7mer-t1A     NYYYYYYA
```



$$\text{Fold enrichment} = \frac{\dfrac{\#\ \text{Target sites for mR-X}}{\#\ \text{Target sites for all miRNAs}}\ \text{experiment}}{\dfrac{\#\ \text{Target sites for mR-X}}{\#\ \text{Target sites for all miRNAs}}\ \text{control}}$$