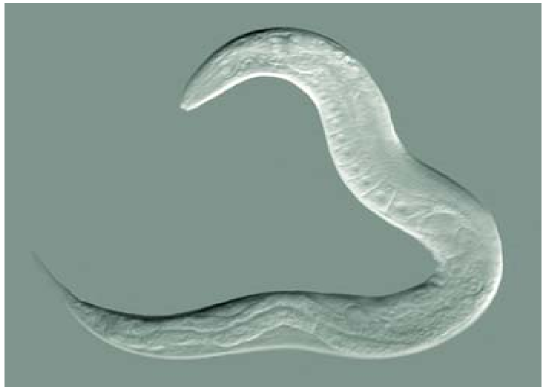
Identifying piRNA targeting rules in *C. elegans*



Wei-Sheng Wu
Professor
Department of Electrical Engineering
National Cheng Kung University
2018.10.20

Outline

- Introduction of piRNA

- **Identifying piRNA targeting rules** in *C. elegans* (Science 2018)

  (**Heng-Chi Lee**, Department of Molecular Genetics and Cell Biology, University of Chicago)

- **pirScan:** a webserver to **predict piRNA targeting sites** and to **avoid transgene silencing** in *C. elegans* (NAR web server 2018)

- **piRTarBase:** a database of **piRNA targeting sites** and their roles in gene regulation (NAR database 2019)

What is piRNA?

- **Piwi-interacting RNA** (**piRNA**) is the largest class of small non-coding RNA molecules expressed in animal cells.

- Over 17000 distinct piRNAs are encoded by the *C. elegans* genome.

- piRNAs in *C. elegans* are also called **21U-RNAs**, for their length of 21 nucleotides with strong 5’ U preference.

What is piRNA?

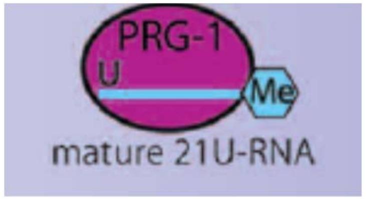
- **Piwi-interacting RNA** (**piRNA**) is the largest class of small non-coding RNA molecules expressed in animal cells.
- Over 17000 distinct piRNAs are encoded by the *C. elegans* genome.
- piRNAs in *C. elegans* are also called **21U-RNAs**, for their length of 21 nucleotides with strong 5’ U preference.

21ur-5557	UUGGCGUGGUCUUUUUAAUUU
21ur-11888	UUUGGUCGAUCGCUAAAAUCG
21ur-15651	CAACUGCAUAUUGAUGAUUUG
21ur-13181	UGGGAGAAAUCUACGGAAAUU
21ur-11965	UUGUUUGAAGUAGGAUUUUGA
21ur-7861	UAGUAAUUUAAUGAUAUGACA
21ur-10491	GUAAGUUCGAUGGUUUGAAUU
21ur-12527	UAACGUCUGAAAUAUGUGGAU
21ur-15658	AUUCCUGUUAGCAAUGAUUAU

U: 96.76%
A: 1.16%
G: 0.86%
C: 1.22%

Function of piRNA

- piRNAs interact with **Piwi-related Argonaute proteins** (**PRG-1** in *C. elegans*) to **defend the genome by silencing foreign nucleic acids** (e.g. **transposons** or **transgenes**).

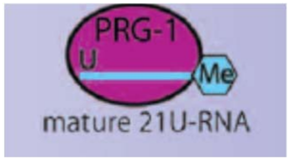


- piRNA function extends beyond genome defense.
- piRNAs can also regulate the expression of endogenous genes in *C. elgans*.

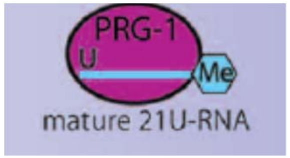
- Kim et al. (2018). A Neuronal piRNA Pathway Inhibits Axon Regeneration in *C. elegans*. Neuron 97, 511–519.e6.
- Tang et al. (2018). A sex chromosome piRNA promotes robust dosage compensation and sex determination in *C. elegans*. Dev. Cell 44, 762–770.

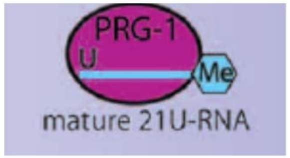
>17000 distinct piRNAs form a foreign gene (Non-self) surveillance system

- **piRNAs can silence most foreign genes.**

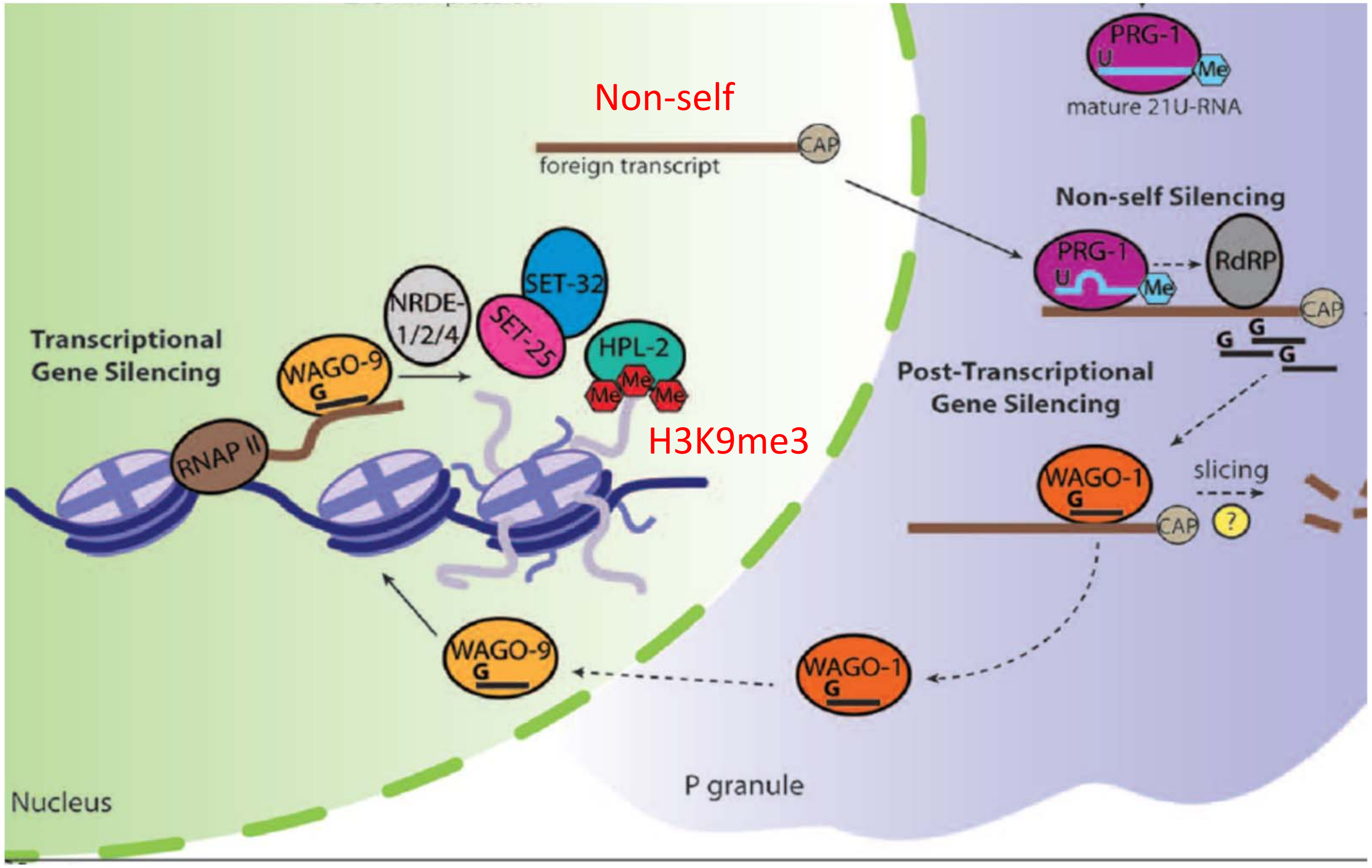


Like Google search engine





RNA-dependent RNA Polymerases (RdRPs)



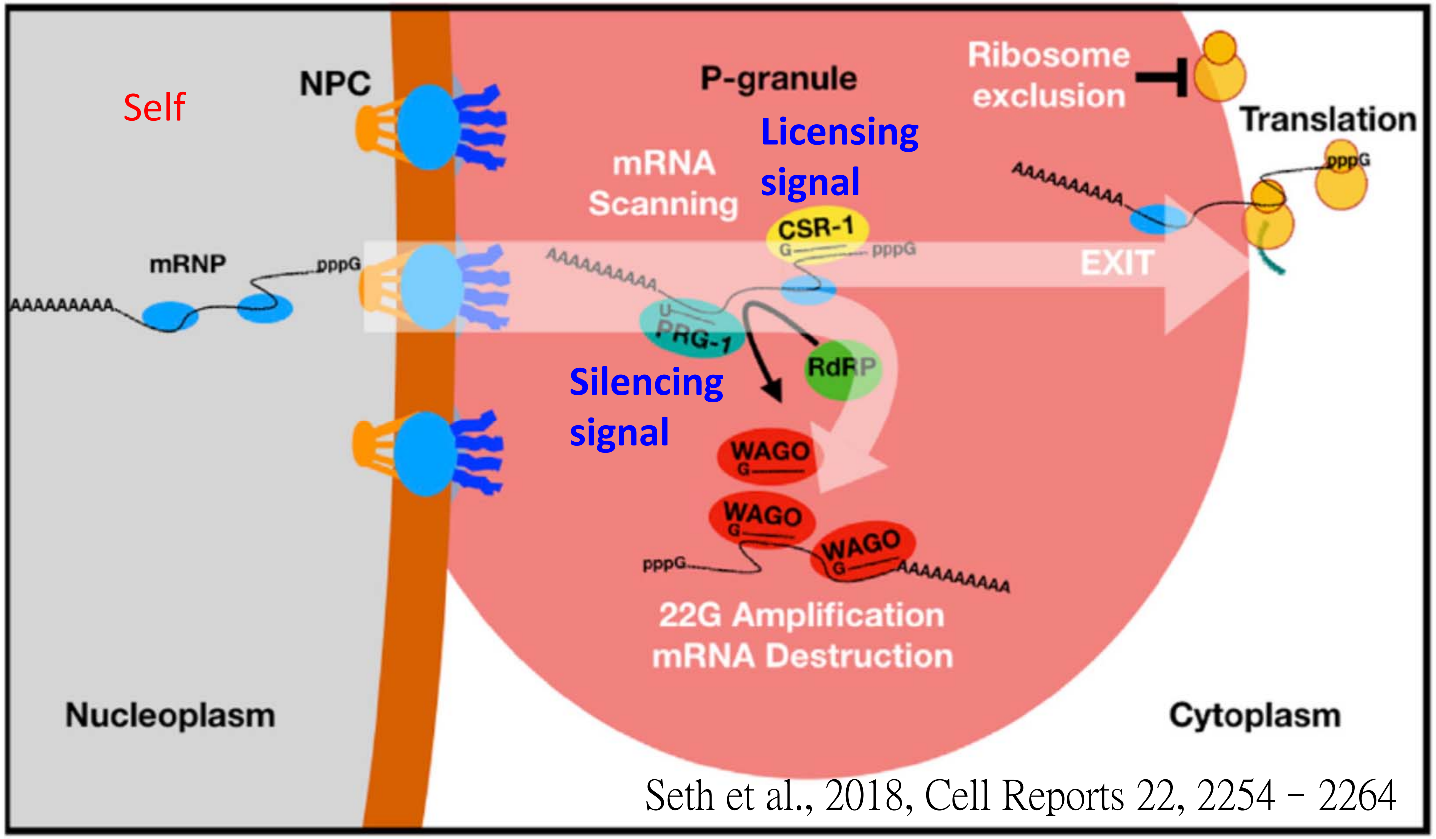
Worm-specific ArGOnautes (WAGOs)

Epigenetics 9:1, 62 – 74; January 2014

Only some endogenous genes are silenced by piRNAs but most are not.

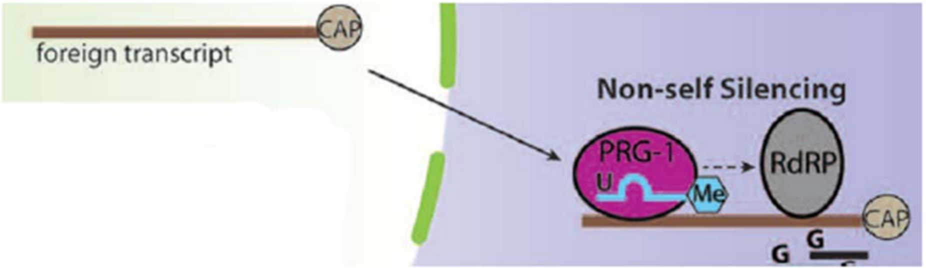
- piRNAs can target essentially all germline mRNAs.
- Endogenous mRNAs (**Self**) are **protected** by CSR-1 Argonaute Pathway.
- The **balance** of licensing (CSR-1) and silencing (PRG-1) signals along an mRNA determines the fraction of mRNA molecules that escape destruction and gain access to the translation machinery.

P granules: a class of perinuclear RNA **granules** specific to the germline in *C. elegans*.



Seth et al., 2018, Cell Reports 22, 2254 – 2264

How do piRNAs recognize their targets?



siRNA Argonaute

miRNA Argonaute

PIWI Argonaute

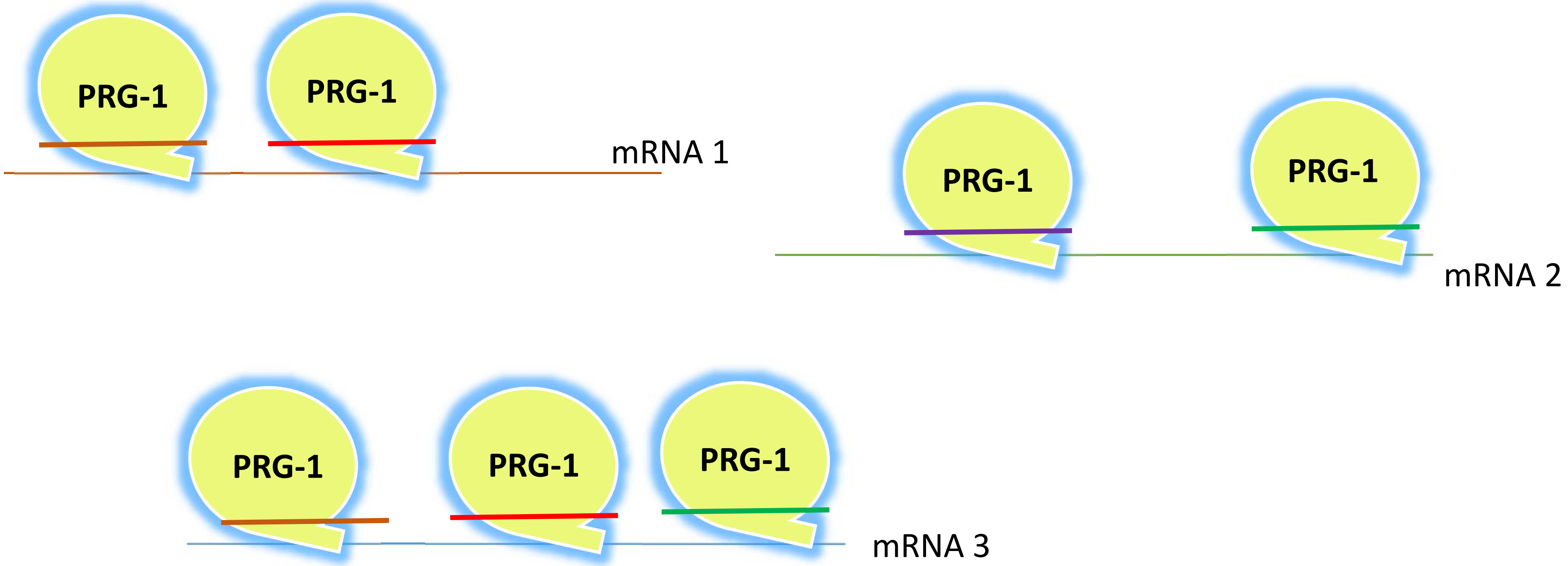
near-perfect matches

base pairing at
“Seed region”

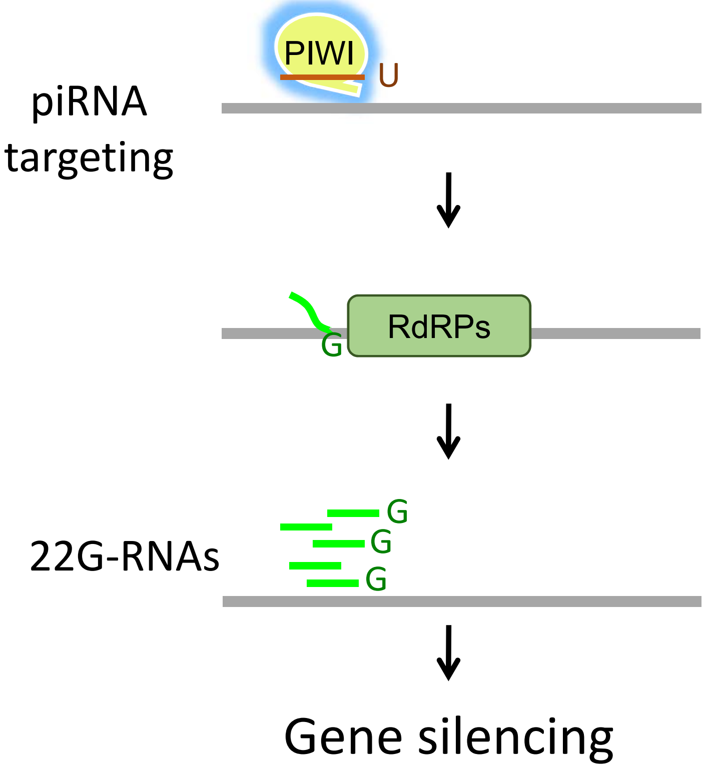
? Answered by our Science paper

Identification of piRNA targeting rule is difficult

- C. elegans has >17000 piRNAs and >33000 mRNAs



22G-RNAs can serve as a “mark” for piRNA targeting sites *in vivo*



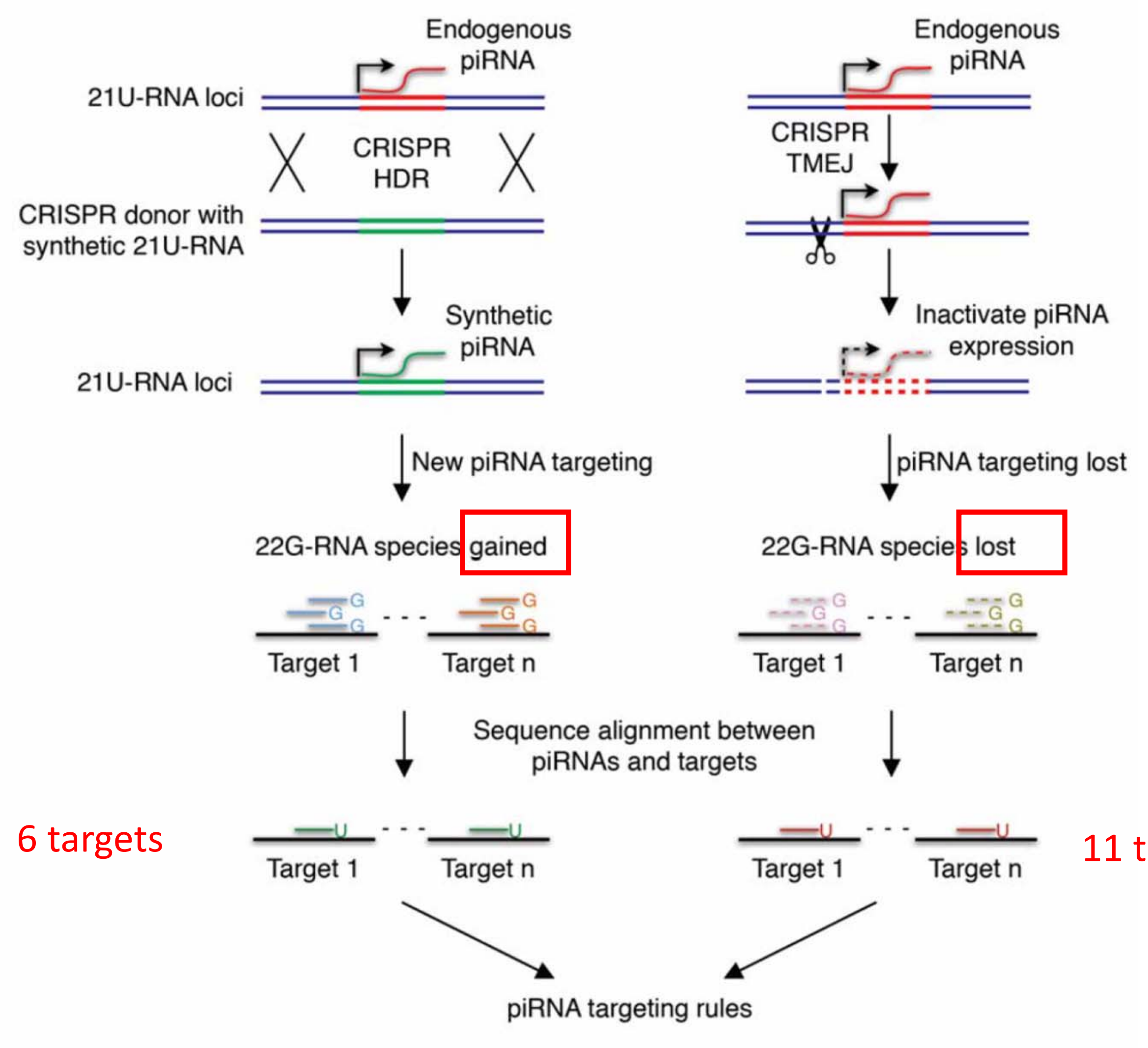
piRNA targeting

22G-RNAs

Gene silencing

Bagijn MP et al., 2012
Lee HC et al., 2012

Identify the mRNA targets of a single piRNA



6 targets

11 targets

Synthetic/Deleted piRNAs induce/lose 22G-RNAs and reveal piRNA targeting signature

```
                                X  seed    Non-seed
piRNA        Synthetic piRNA 5′-U GUUUCA UAUGAUCUGGGUAU
                                  |||||| ||::| |||||:||
Target mRNA  T10B11.2        3′-C CAAAGU AUGUUUGACCCGUA
```

Non-GU:1
GU:3

```
             Synthetic piRNA 5′-U GUUUCA UAUGAUCUGGGUAU
                                : |||||| | ||  |:|||||
             C05F2.2         3′-G CAAAGU UUUCUCUAUCCAUA
```

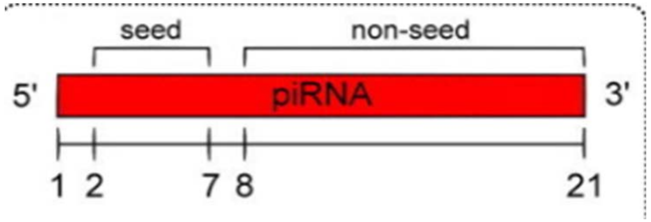
Non-GU:4
GU:1

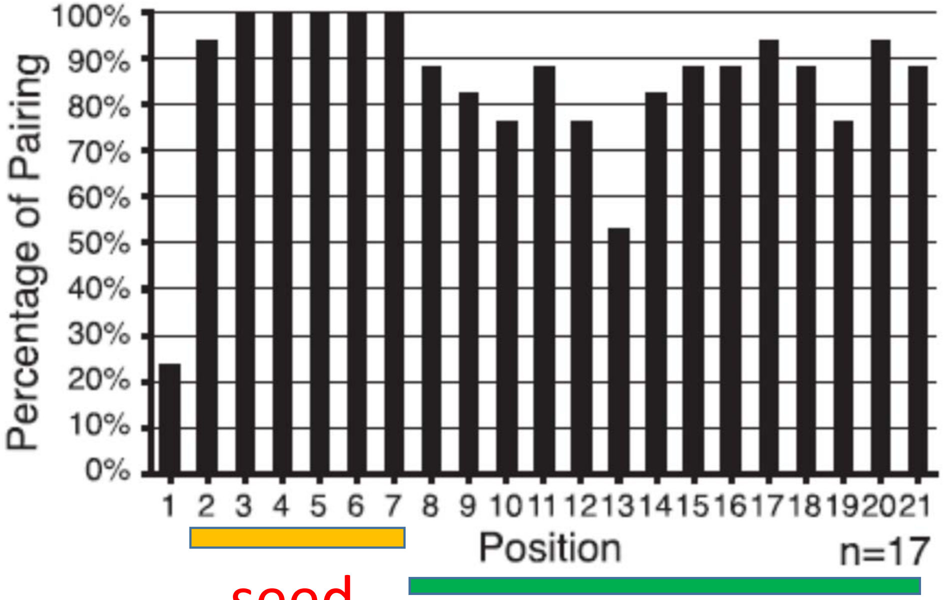
```
             Synthetic piRNA 5′-U GUUUCA UAUGAUCUGGGUAU
                                  ||||||  ||| |||:|||||
             T26A5.2         3′-U CAAAGU UUACAAGAUCCAUA
```

Non-GU:2
GU:1

```
             21U-RNA-X1      5′-U GUGUUU UUUGUUGAGGUAUC
                                  |||||: : | ||||||||||
             F59A3.9         3′-U CACAAG GCAAAACUCCAUAG
```

Non-GU:2
GU:2



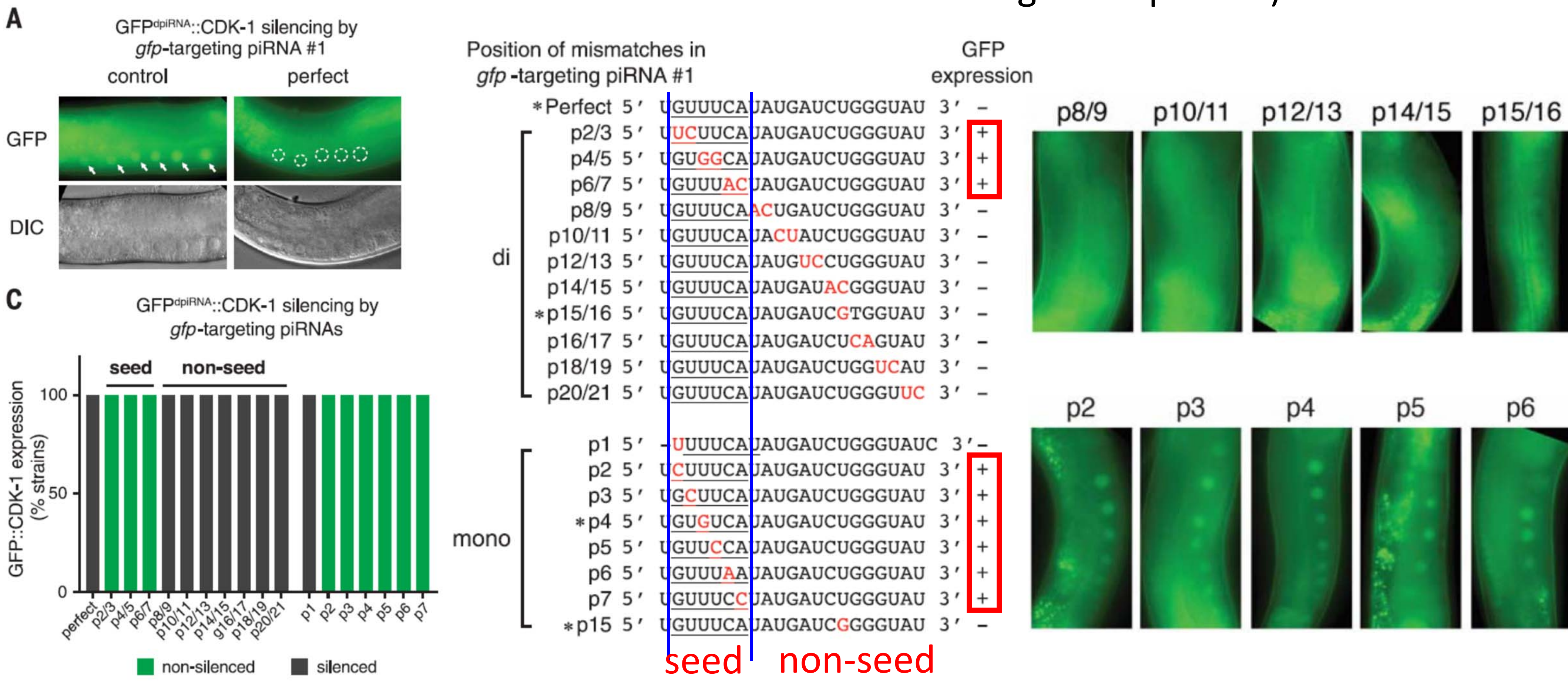


seed

non-seed

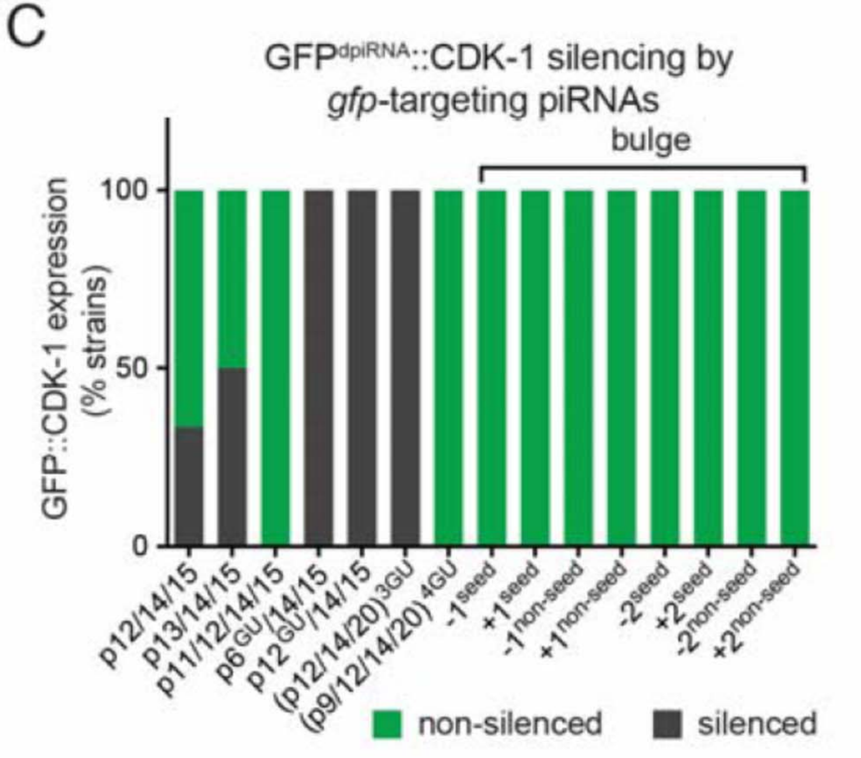
A piRNA reporter assay to investigate the piRNA targeting rules

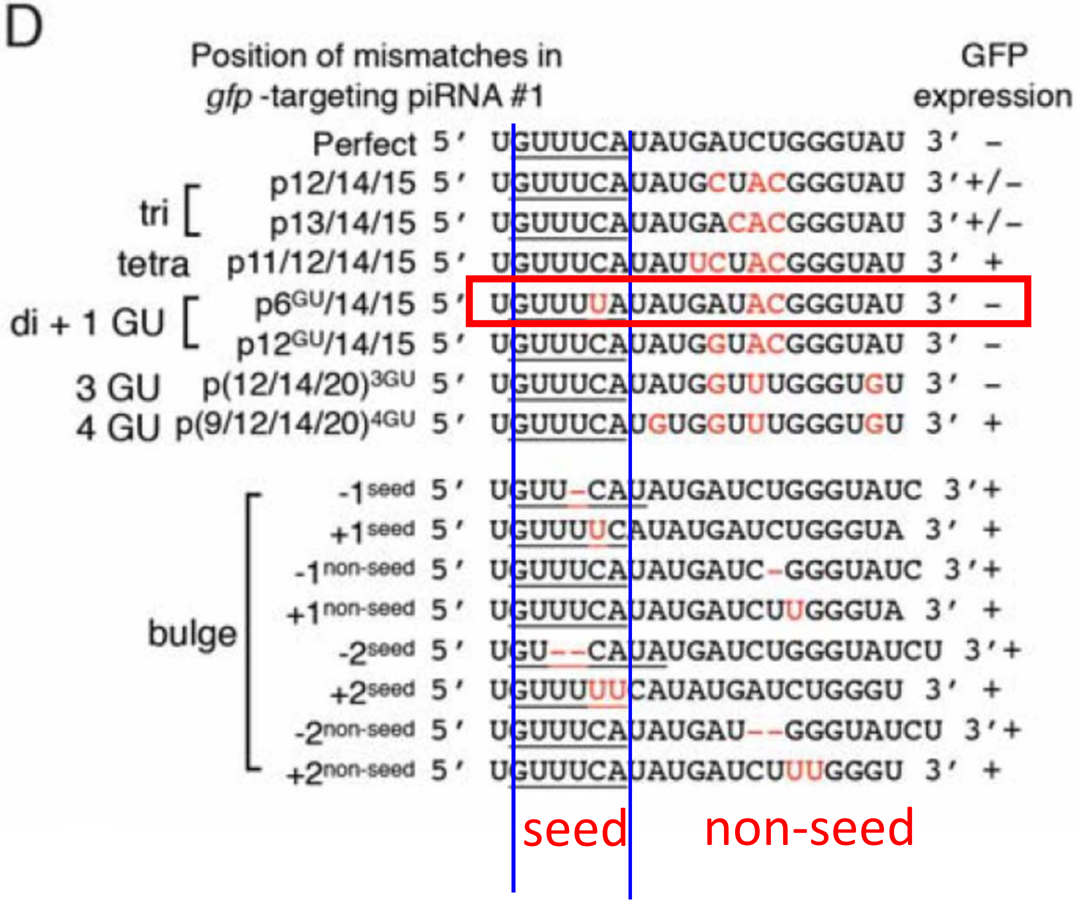
GFP^{dpiRNA}:CDK-1transgene (depletion of piRNA targeting sites, in which the GFP sequence has been recoded to avoid silencing by endogenous piRNAs)



	Position of mismatches in *gfp* -targeting piRNA #1		GFP expression
	*Perfect	5′ UGUUUCAUAUGAUCUGGGUAU 3′	–
di	p2/3	5′ UUCUUCAUAUGAUCUGGGUAU 3′	+
di	p4/5	5′ UGUGGCAUAUGAUCUGGGUAU 3′	+
di	p6/7	5′ UGUUUACUAUGAUCUGGGUAU 3′	+
di	p8/9	5′ UGUUUCAACUGAUCUGGGUAU 3′	–
di	p10/11	5′ UGUUUCAUACUAUCUGGGUAU 3′	–
di	p12/13	5′ UGUUUCAUAUGUCCUGGGUAU 3′	–
di	p14/15	5′ UGUUUCAUAUGAUACGGGUAU 3′	–
di	*p15/16	5′ UGUUUCAUAUGAUCGTGGUAU 3′	–
di	p16/17	5′ UGUUUCAUAUGAUCUCAGUAU 3′	–
di	p18/19	5′ UGUUUCAUAUGAUCUGGUCAU 3′	–
di	p20/21	5′ UGUUUCAUAUGAUCUGGGUUC 3′	–
	p1	5′ -UUUUCAUAUGAUCUGGGUAUC 3′	–
mono	p2	5′ UCUUUCAUAUGAUCUGGGUAU 3′	+
mono	p3	5′ UGCUUCAUAUGAUCUGGGUAU 3′	+
mono	*p4	5′ UGUGUCAUAUGAUCUGGGUAU 3′	+
mono	p5	5′ UGUUCCAUAUGAUCUGGGUAU 3′	+
mono	p6	5′ UGUUUAAUAUGAUCUGGGUAU 3′	+
mono	p7	5′ UGUUUCCUAUGAUCUGGGUAU 3′	+
mono	*p15	5′ UGUUUCAUAUGAUCGGGGUAU 3′	–

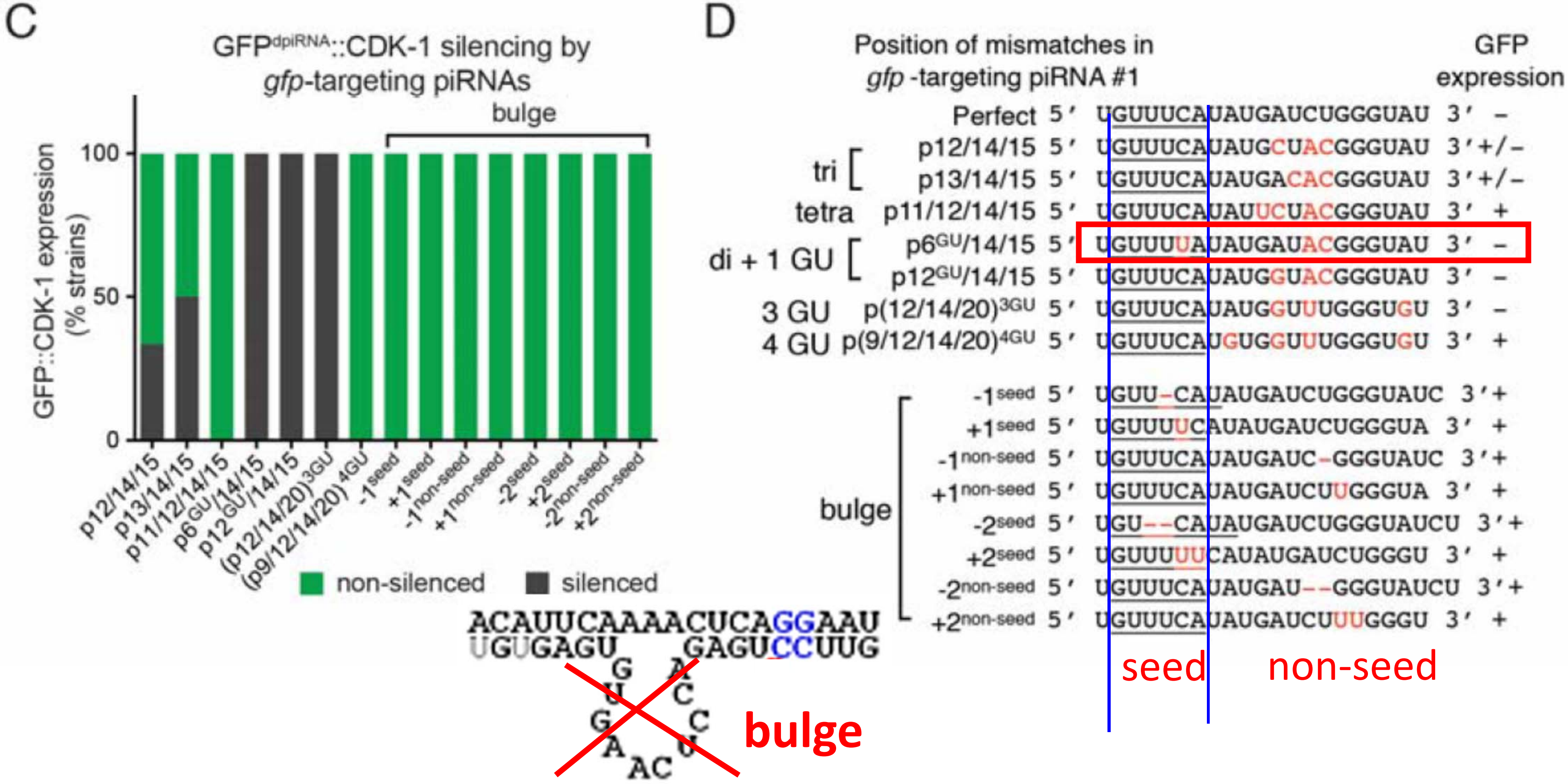
A piRNA reporter assay to investigate the piRNA targeting rules





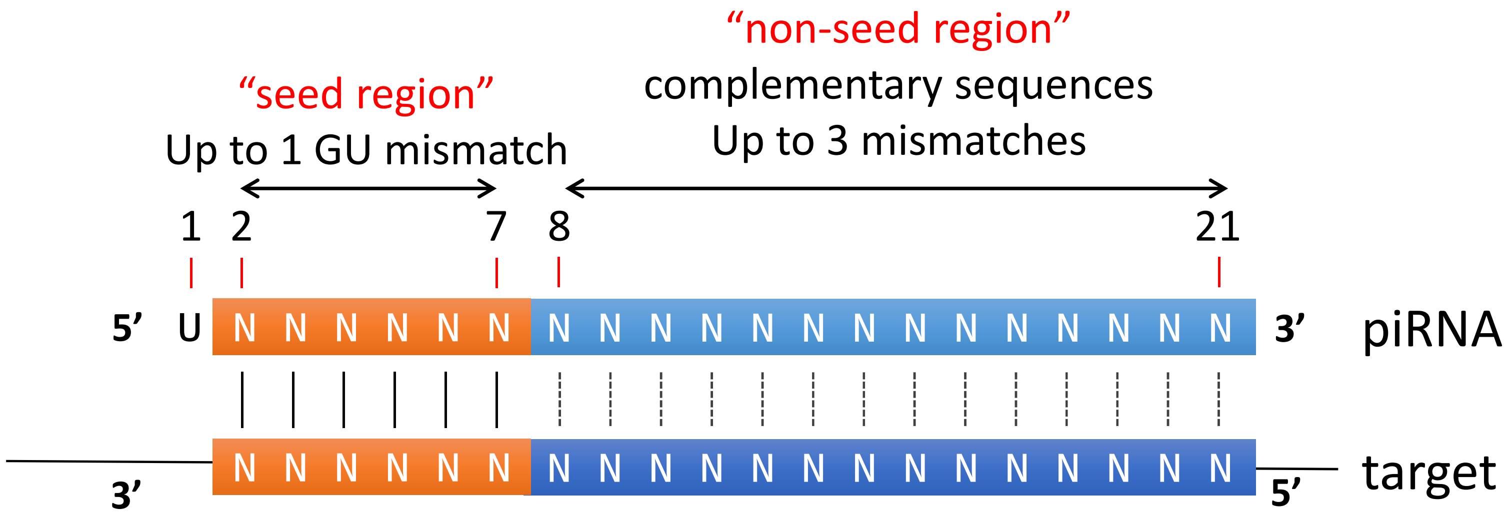
	Position of mismatches in *gfp*-targeting piRNA #1		GFP expression
	Perfect	5′ UGUUUCAUAUGAUCUGGGUAU 3′	–
tri	p12/14/15	5′ UGUUUCAUAUGCUACGGGUAU 3′	+/–
tri	p13/14/15	5′ UGUUUCAUAUGACACGGGUAU 3′	+/–
tetra	p11/12/14/15	5′ UGUUUCAUAUUCUACGGGUAU 3′	+
di + 1 GU	p6^{GU}/14/15	5′ UGUUUUAUAUGAUACGGGUAU 3′	–
di + 1 GU	p12^{GU}/14/15	5′ UGUUUCAUAUGGUACGGGUAU 3′	–
3 GU	p(12/14/20)^{3GU}	5′ UGUUUCAUAUGGUUUGGGUGU 3′	–
4 GU	p(9/12/14/20)^{4GU}	5′ UGUUUCAUGUGGUUUGGGUGU 3′	+
bulge	-1^{seed}	5′ UGUU–CAUAUGAUCUGGGUAUC 3′	+
bulge	+1^{seed}	5′ UGUUUUCAUAUGAUCUGGGUA 3′	+
bulge	-1$^{non-seed}$	5′ UGUUUCAUAUGAUC–GGGUAUC 3′	+
bulge	+1$^{non-seed}$	5′ UGUUUCAUAUGAUCUUGGGUA 3′	+
bulge	-2^{seed}	5′ UGU––CAUAUGAUCUGGGUAUCU 3′	+
bulge	+2^{seed}	5′ UGUUUUUCAUAUGAUCUGGGU 3′	+
bulge	-2$^{non-seed}$	5′ UGUUUCAUAUGAU––GGGUAUCU 3′	+
bulge	+2$^{non-seed}$	5′ UGUUUCAUAUGAUCUUUGGGU 3′	+

A piRNA reporter assay to investigate the piRNA targeting rules



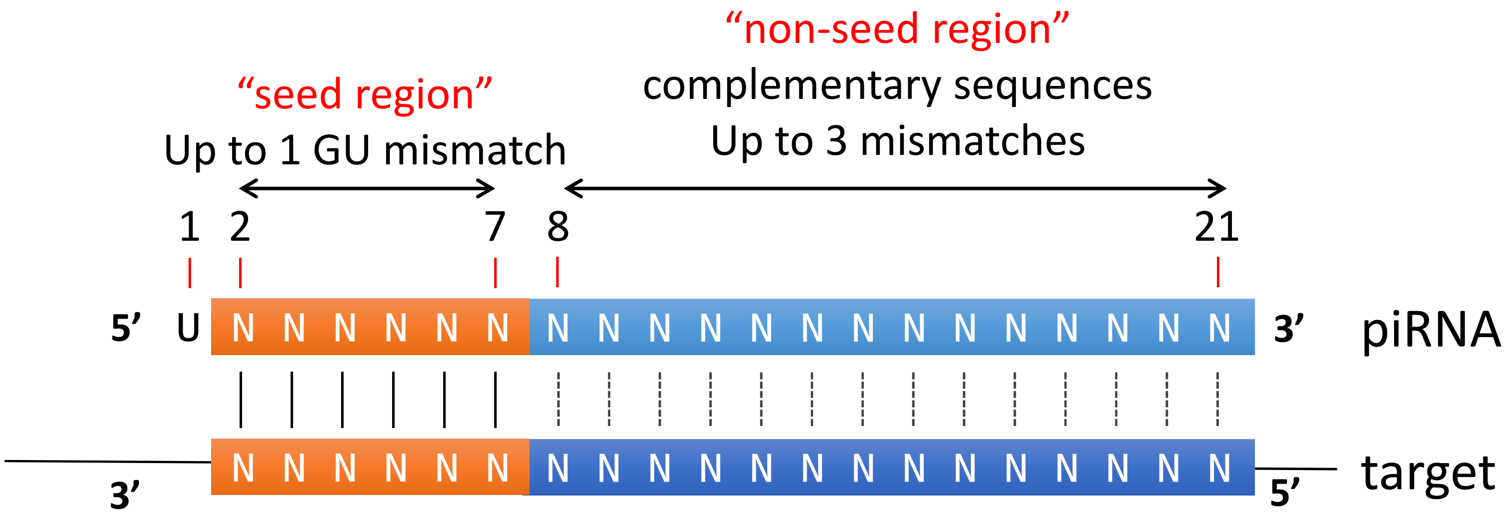
		Position of mismatches in *gfp* -targeting piRNA #1	GFP expression
	Perfect	5′ UGUUUCAUAUGAUCUGGGUAU 3′	–
tri	p12/14/15	5′ UGUUUCAUAUGCUACGGGUAU 3′	+/–
	p13/14/15	5′ UGUUUCAUAUGACACGGGUAU 3′	+/–
tetra	p11/12/14/15	5′ UGUUUCAUAUUCUACGGGUAU 3′	+
di + 1 GU	p6^{GU}/14/15	5′ UGUUUUAUAUGAUACGGGUAU 3′	–
	p12^{GU}/14/15	5′ UGUUUCAUAUGGUACGGGUAU 3′	–
3 GU	p(12/14/20)^{3GU}	5′ UGUUUCAUAUGGUUUGGGUGU 3′	–
4 GU	p(9/12/14/20)^{4GU}	5′ UGUUUCAUGUGGUUUGGGUGU 3′	+
bulge	-1^{seed}	5′ UGUU–CAUAUGAUCUGGGUAUC 3′	+
	+1^{seed}	5′ UGUUUUCAUAUGAUCUGGGUA 3′	+
	-1$^{non-seed}$	5′ UGUUUCAUAUGAUC–GGGUAUC 3′	+
	+1$^{non-seed}$	5′ UGUUUCAUAUGAUCUUGGGUA 3′	+
	-2^{seed}	5′ UGU––CAUAUGAUCUGGGUAUCU 3′	+
	+2^{seed}	5′ UGUUUUUCAUAUGAUCUGGGU 3′	+
	-2$^{non-seed}$	5′ UGUUUCAUAUGAU––GGGUAUCU 3′	+
	+2$^{non-seed}$	5′ UGUUUCAUAUGAUCUUUGGGU 3′	+

Identified piRNA targeting rules



- piRNA targeting in *C. elegans* prefers **near-perfect pairing at the piRNA seed region** (**one GU wobble pair is tolerated**).
- Supplementary pairing outside of the seed region also contributes to piRNA targeting, but few mismatches are tolerated (**piRNAs tolerate up to three non-seed mismatches but not RNA bulges**).
- **GU pairs are moderately more tolerated** than non-GU mismatches in the non-seed region.

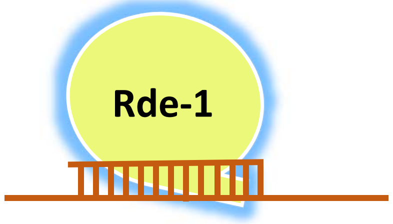
Identified piRNA targeting rules



"non-seed region"
complementary sequences
Up to 3 mismatches

"seed region"
Up to 1 GU mismatch

siRNA Argonaute



near-perfect matches

miRNA Argonaute

base pairing at
"Seed region"

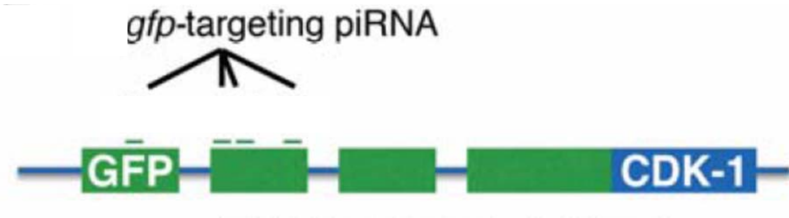
Transgenes carrying foreign nucleic acids are frequently silenced in the germline of *C. elegans*

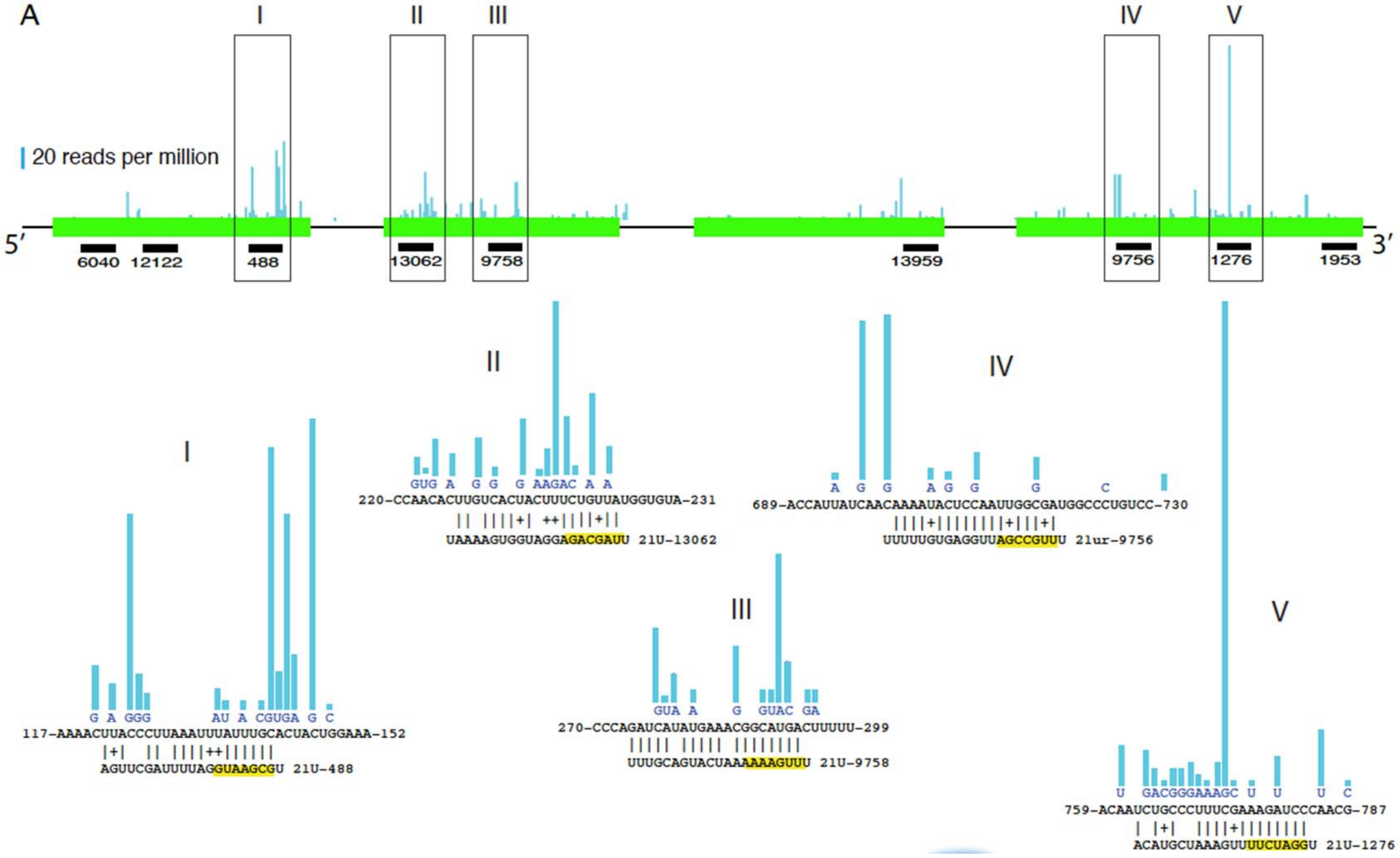
Transgene	% expression
Cas9	*0*
I-Sce I	*0*
Mos transposase	*0*
gfp::*csr-1*	6
gfp::*rde-3*	17
gfp::*cdk-1*	0
flag::*csr-1*	83
flag::*rde-3*	100

Shirayama M et al., 2012

Do piRNA recognize foreign nucleic acids and silence them?

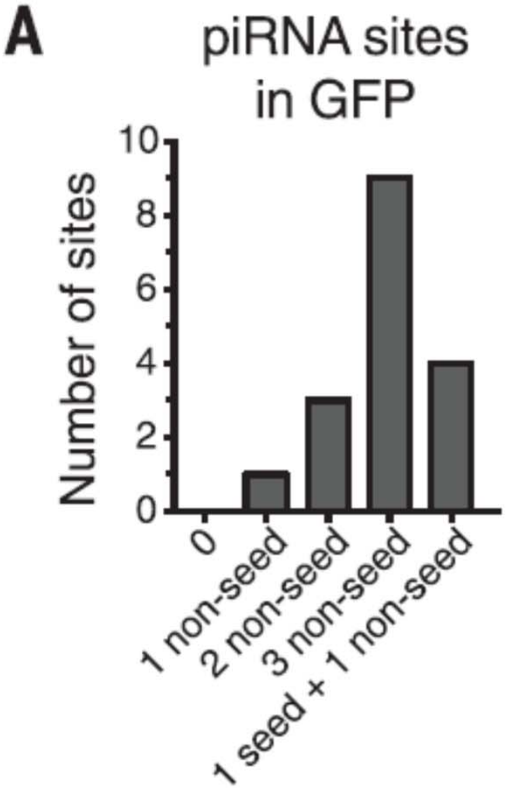
Can we avoid piRNA recognition?

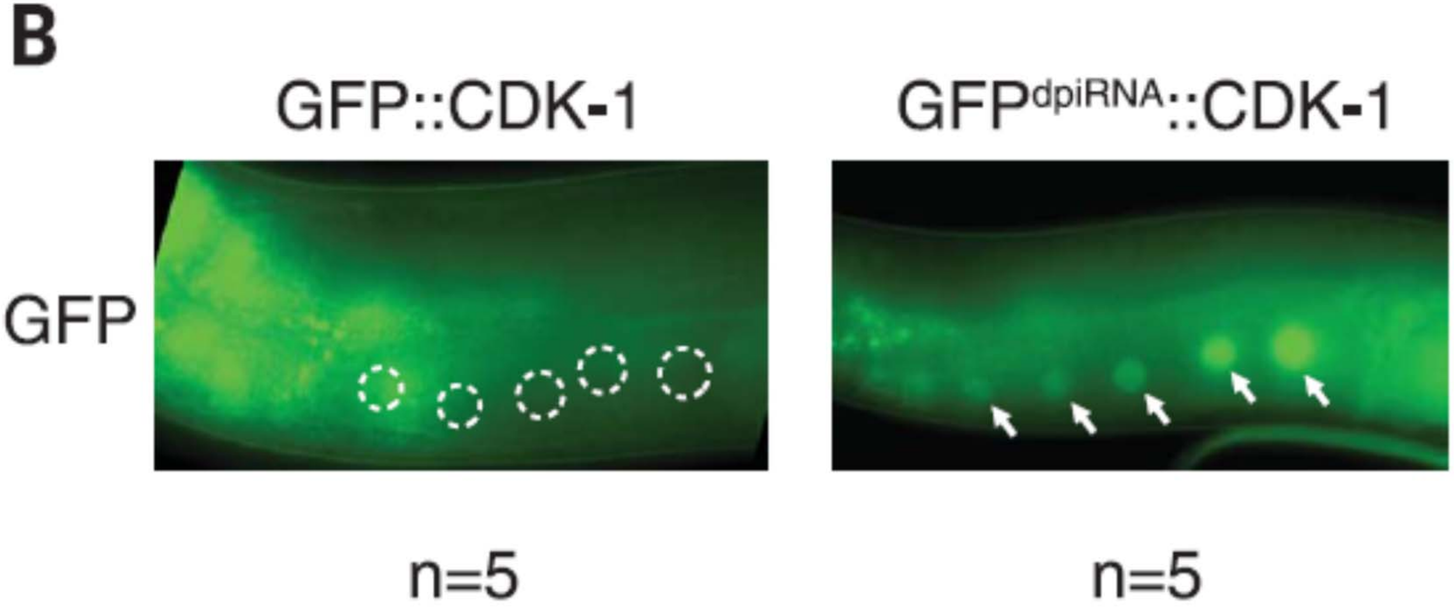




Transgene	% expression
Cas9	*0*
I-Sce I	*0*
Mos transposase	*0*
gfp::csr-1	6
gfp::rde-3	17
gfp::cdk-1	0

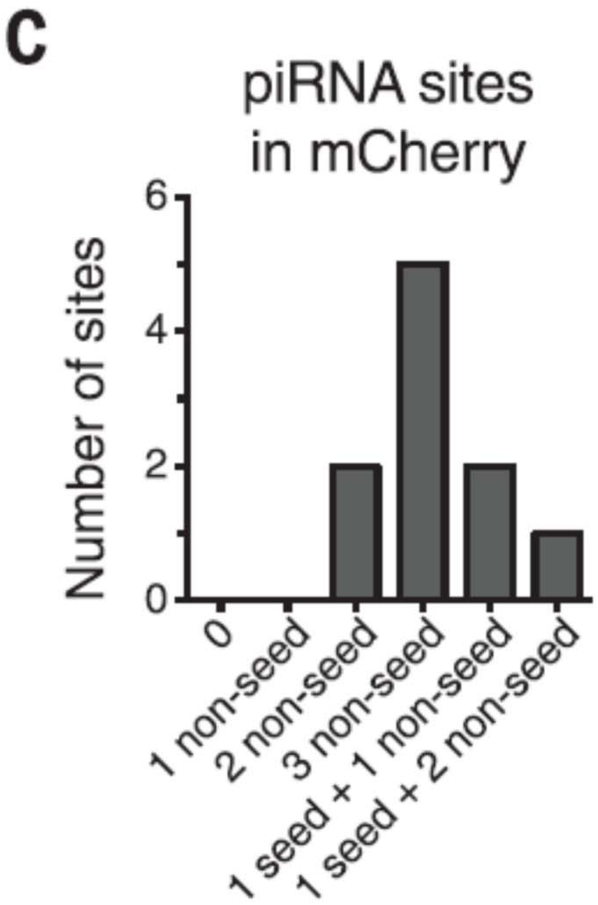
Silencing-prone transgenes can be expressed in the germline by avoiding piRNA targeting

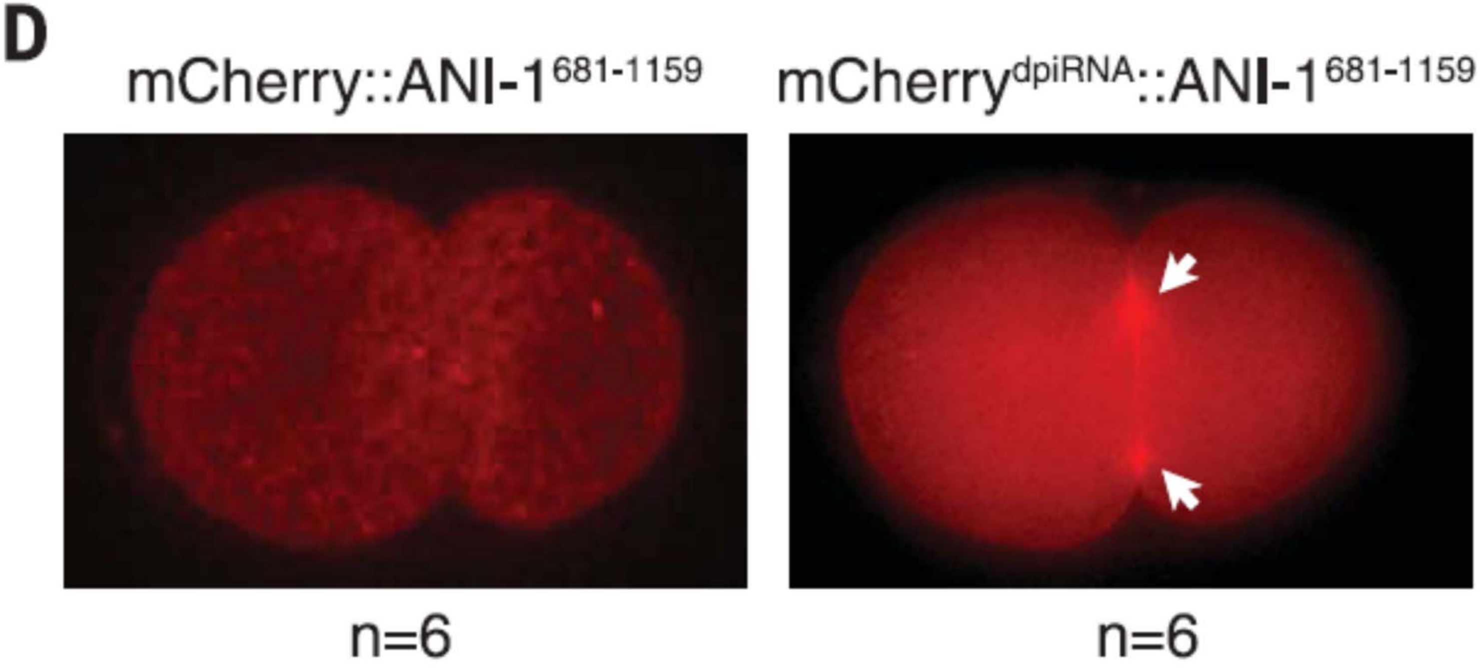




Introduce synonymous mutations on the GFP DNA sequence in piRNA targeting sites without changing its amino acid sequence

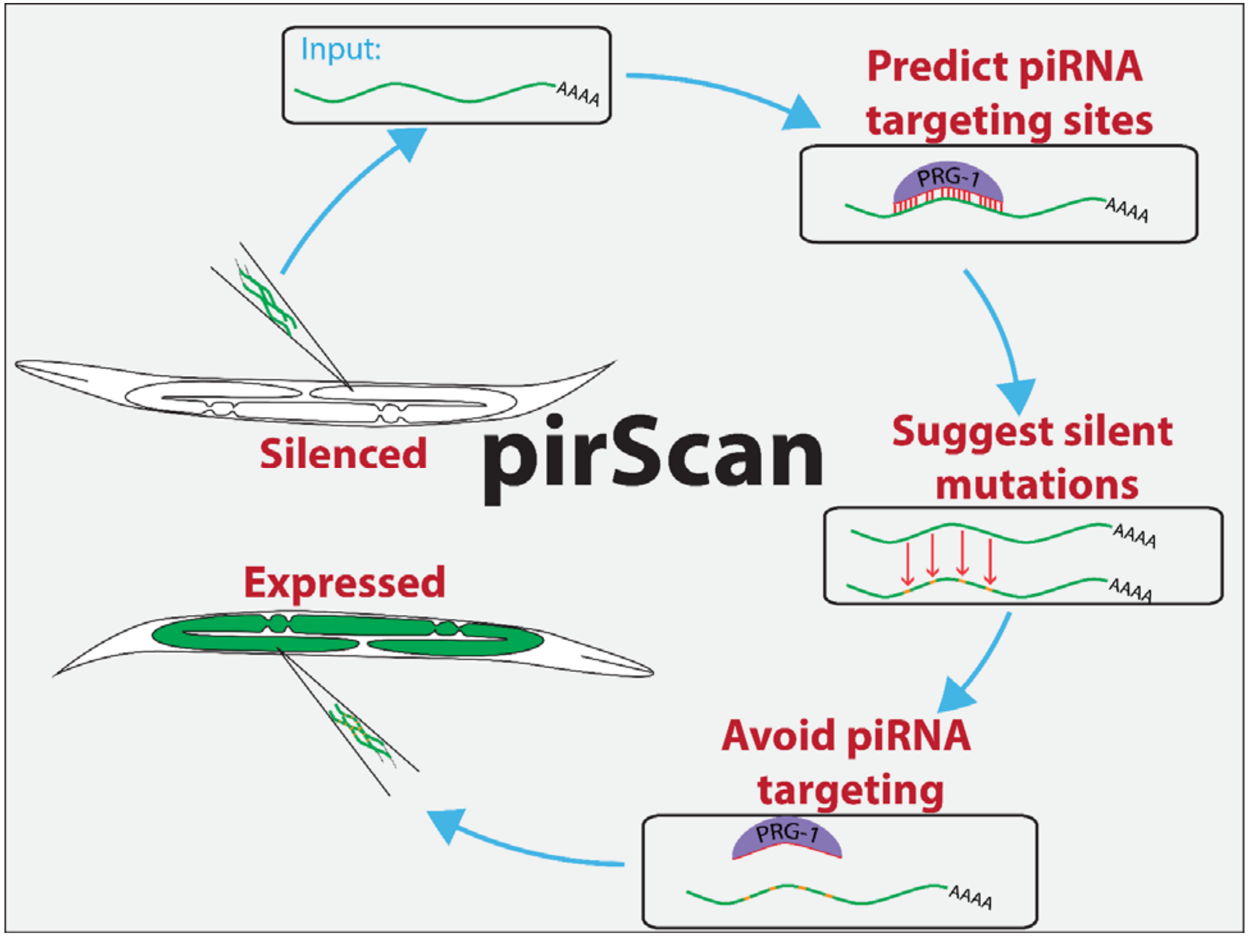
Silencing-prone transgenes can be expressed in the germline by avoiding piRNA targeting



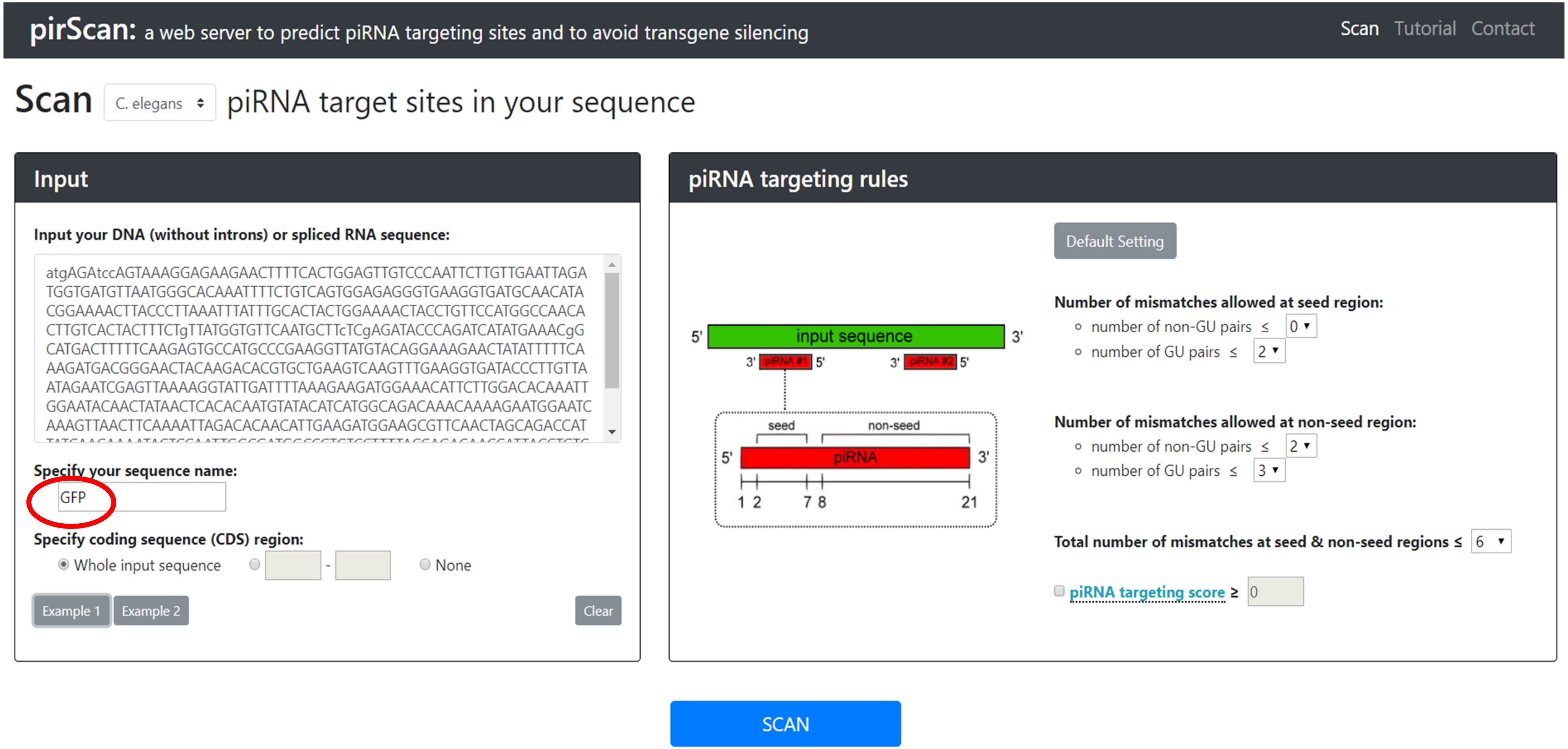


Introduce synonymous mutations on the GFP DNA sequence **in piRNA targeting sites** without changing its amino acid sequence

pirScan: a web server to predict piRNA targeting sites and to avoid transgene silencing



Input page of pirScan



pirScan: a web server to predict piRNA targeting sites and to avoid transgene silencing

Scan C. elegans piRNA target sites in your sequence

Input

Input your DNA (without introns) or spliced RNA sequence:

```
atgAGAtccAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGA
TGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATA
CGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACA
CTTGTCACTACTTTCTgTTATGGTGTTCAATGCTTcTCgAGATACCCAGATCATATGAAACgG
CATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCA
AAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTA
ATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATT
GGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATC
AAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCAT
```

Specify your sequence name: GFP

Specify coding sequence (CDS) region:

● Whole input sequence ○ - ○ None

piRNA targeting rules

Number of mismatches allowed at seed region:

- number of non-GU pairs ≤ 0
- number of GU pairs ≤ 2

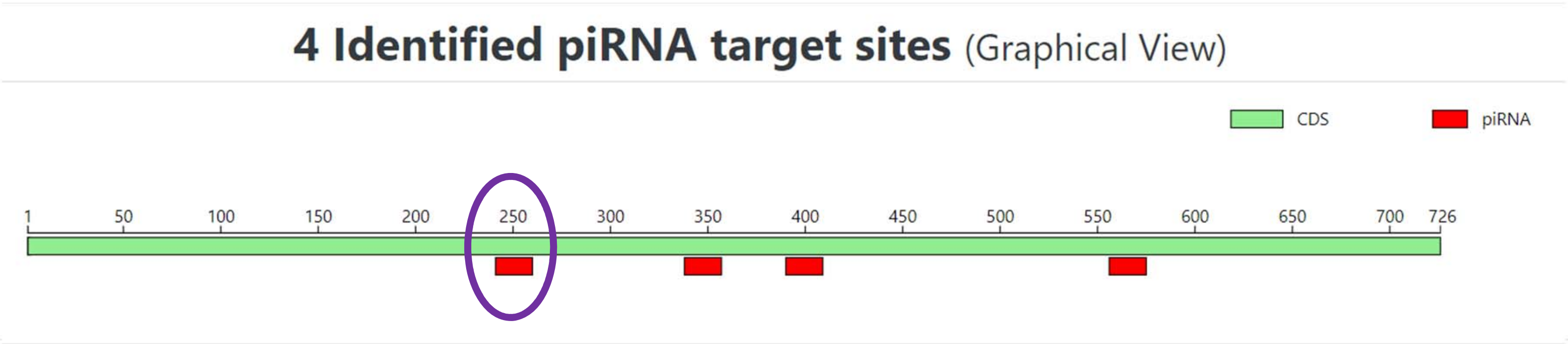
Number of mismatches allowed at non-seed region:

- number of non-GU pairs ≤ 2
- number of GU pairs ≤ 3

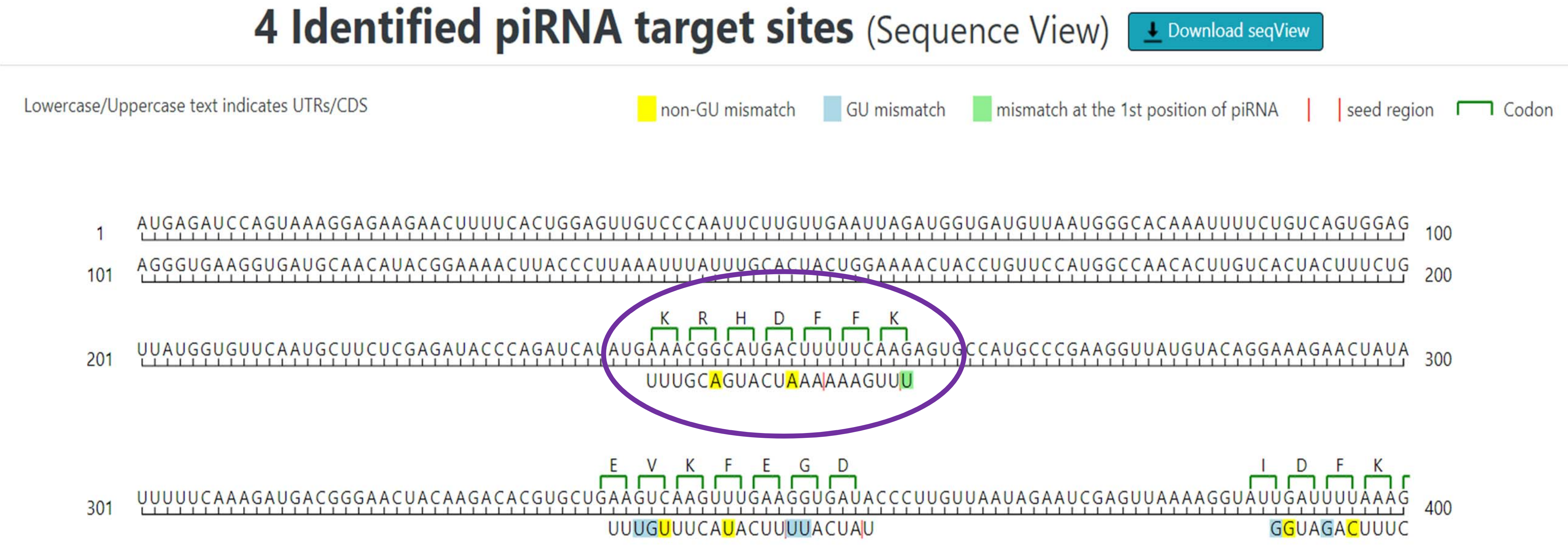
Total number of mismatches at seed & non-seed regions ≤ 6

☐ piRNA targeting score ≥ 0

Scan results



4 Identified piRNA target sites (Graphical View)



4 Identified piRNA target sites (Sequence View)

Lowercase/Uppercase text indicates UTRs/CDS

🟨 non-GU mismatch　🟦 GU mismatch　🟩 mismatch at the 1st position of piRNA　| | seed region　⊓ Codon

1 AUGAGAUCCAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGUGAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAG 100

101 AGGGUGAAGGUGAUGCAACAUACGGAAAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUGUCACUACUUUCUG 200

K R H D F F K

201 UUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGCAUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUA 300

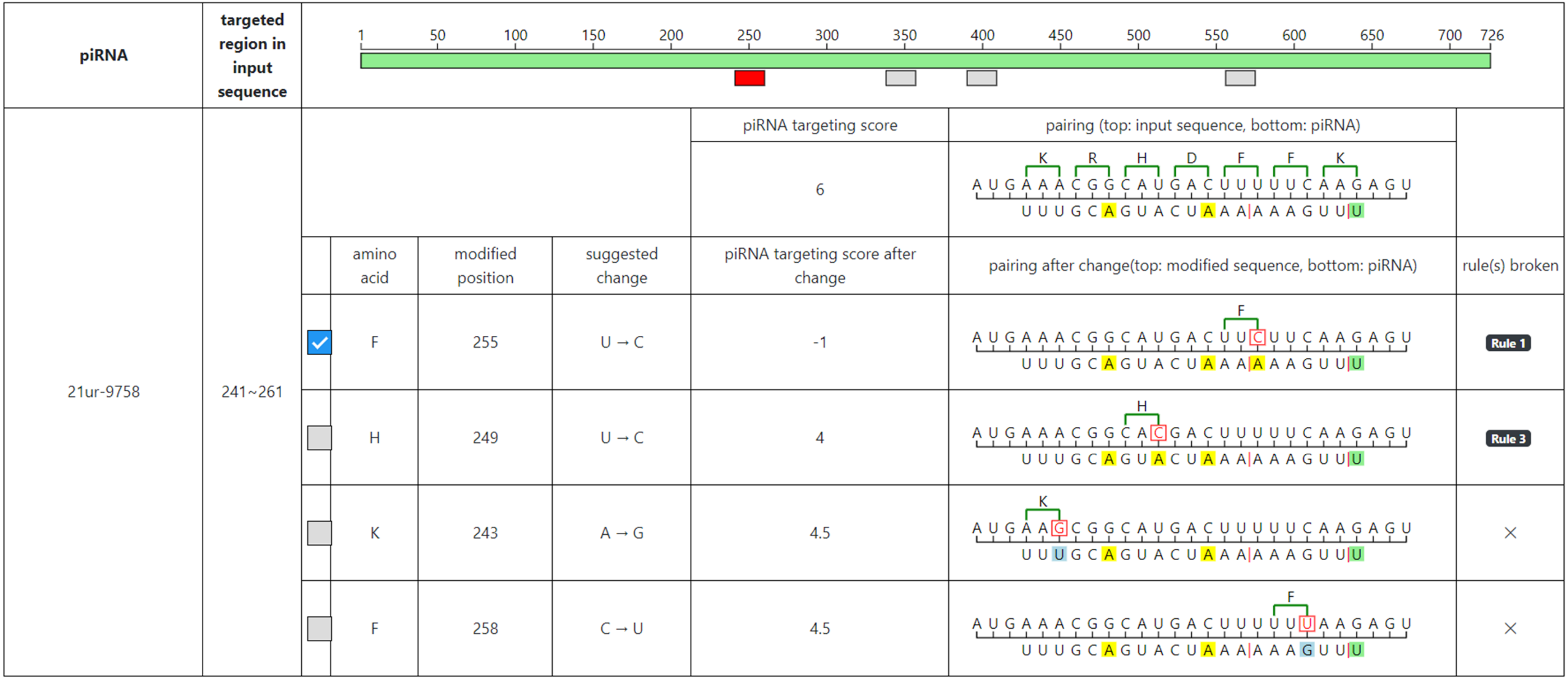
UUUGCAGUACUAAAAAGUUU

E V K F E G D　I D F K

301 UUUUUCAAAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUAAUAGAAUCGAGUUAAAAGGUAUUGAUUUUAAAG 400

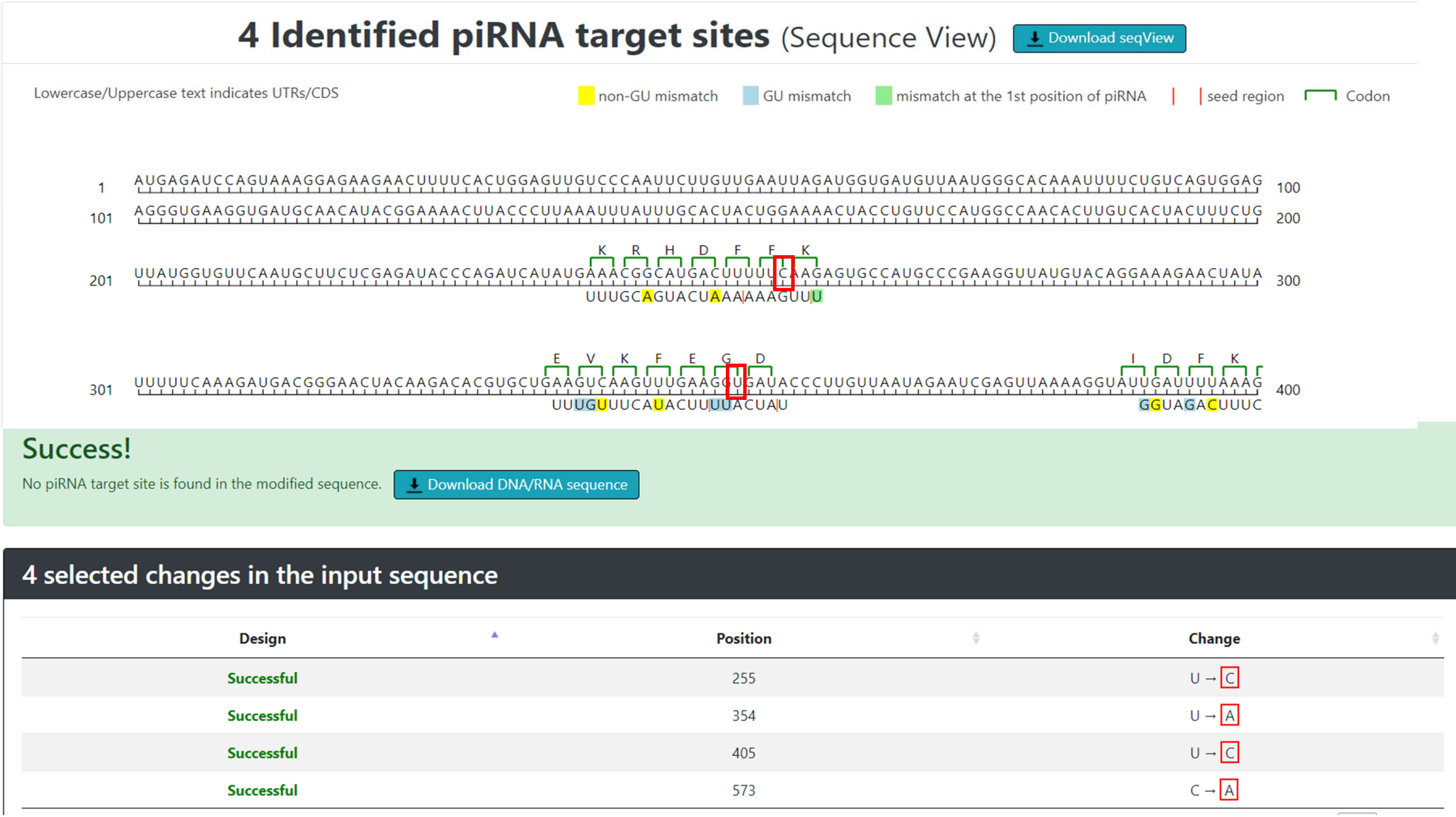
UUUGUUUCAUACUUUUACUAU　GGUAGACUUUC

Suggest synonymous mutations



piRNA	targeted region in input sequence	piRNA targeting score	pairing (top: input sequence, bottom: piRNA)
21ur-9758	241~261	6	K R H D F F K A U G A A A C G G C A U G A C U U U U U C A A G A G U U U U G C A G U A C U A A A|A A A G U U|U

	amino acid	modified position	suggested change	piRNA targeting score after change	pairing after change(top: modified sequence, bottom: piRNA)	rule(s) broken
☑	F	255	U → C	-1	F A U G A A A C G G C A U G A C U U C U U C A A G A G U U U U G C A G U A C U A A A|A A A G U U|U	Rule 1
☐	H	249	U → C	4	H A U G A A A C G G C A C G A C U U U U U C A A G A G U U U U G C A G U A C U A A A|A A A G U U|U	Rule 3
☐	K	243	A → G	4.5	K A U G A A G C G G C A U G A C U U U U U C A A G A G U U U U G C A G U A C U A A A|A A A G U U|U	×
☐	F	258	C → U	4.5	F A U G A A A C G G C A U G A C U U U U U U A A G A G U U U U G C A G U A C U A A A|A A A G U U|U	×



4 Identified piRNA target sites (Sequence View)

Lowercase/Uppercase text indicates UTRs/CDS

```
  1 AUGAGAUCCAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGUGAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAG 100
101 AGGGUGAAGGUGAUGCAACAUACGGAAAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUGUCACUACUUUCUG 200
                                                  K  R  H  D  F  F  K
201 UUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGCAUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUA 300
                                                UUUGCAGUACUAAAAAAGUUU
                                             E  V  K  F  E  G  D                                     I  D  F  K
301 UUUUUCAAAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUAAUAGAAUCGAGUUAAAAGGUAUUGAUUUUAAAG 400
                                             UUUGUUUCAUACUUUUACUAU                                  GGUAGACUUUC
```

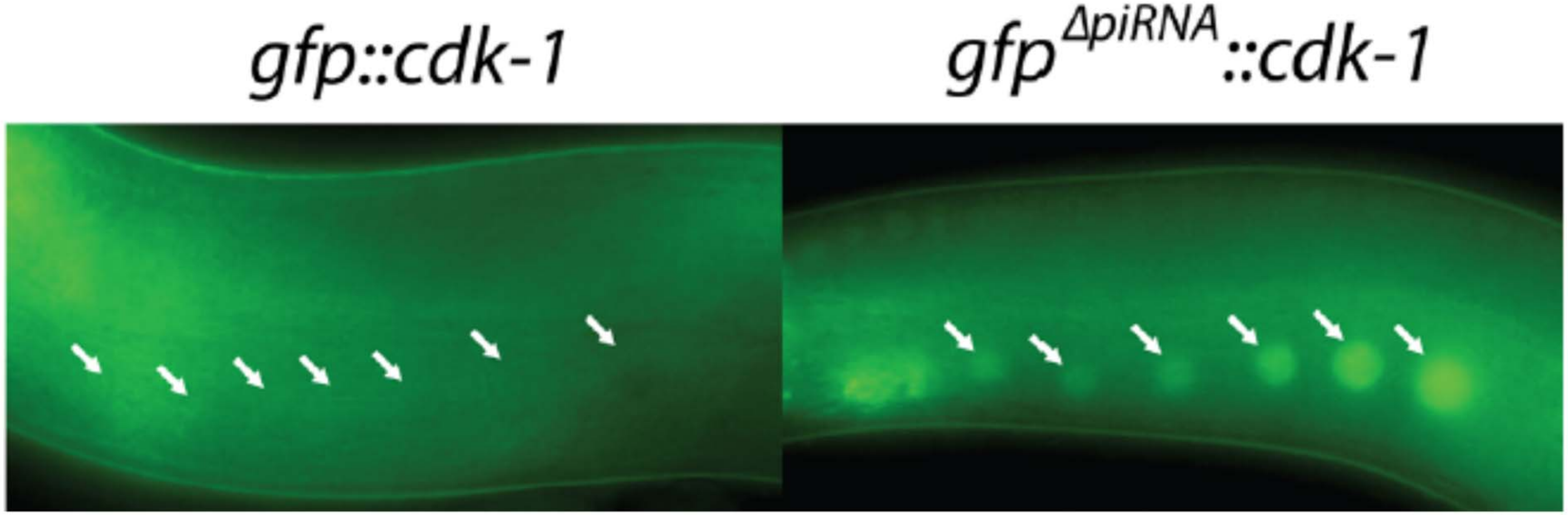
Success!

No piRNA target site is found in the modified sequence.

4 selected changes in the input sequence

Design	Position	Change
Successful	255	U → C
Successful	354	U → A
Successful	405	U → C
Successful	573	C → A

Experimental validation of our modified GFP sequence



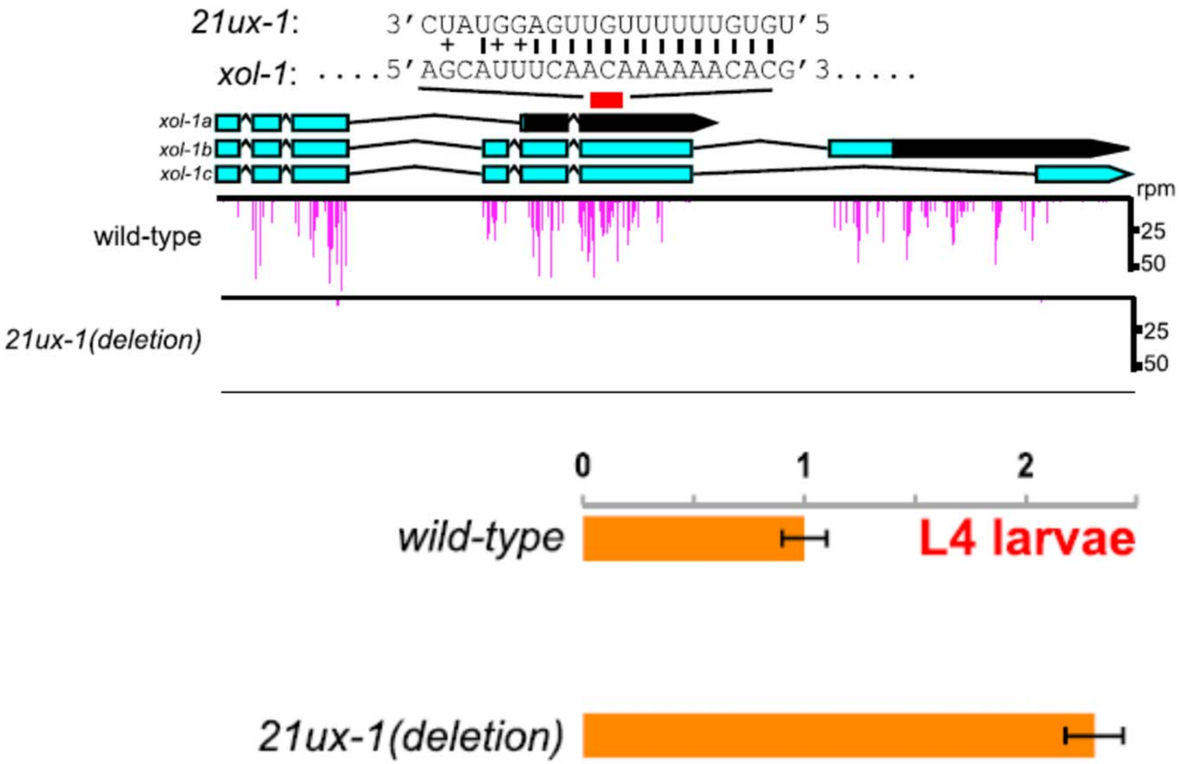
Transgene	% expression
Cas9	*0*
I-Sce I	*0*
Mos transposase	*0*
gfp::csr-1	6
gfp::rde-3	17
gfp::cdk-1	0

piRNA can regulate endogenous genes

- The X chromosome-derived piRNA 21ux-1 downregulates XOL-1 (XO Lethal), a master regulator of X chromosome dosage compensation and sex determination in C. elegans.

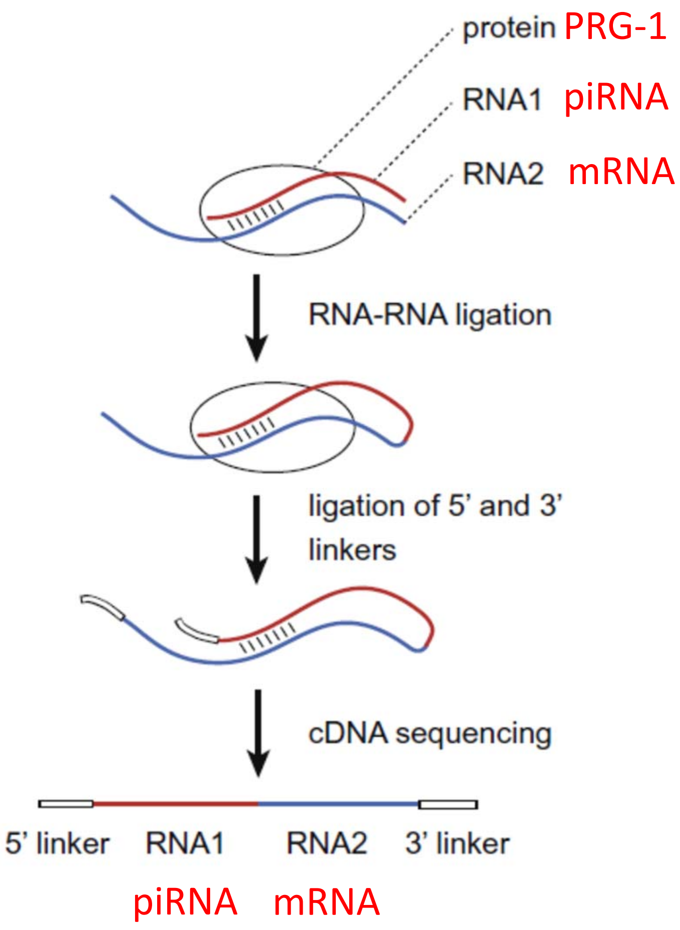
- Mutations in 21ux-1 sensitize hermaphrodites to dosage compensation and sex determination defects.

- 21ux-1 targets the xol-1 transcript, induces 22G-RNA production, and represses xol-1 expression.



CLASH: Crosslinking, Ligation And Sequencing of Hybrids

- The CLASH technique captures RNA–RNA interactions by physically joining two RNA molecules associated with a protein complex into a single chimeric RNA molecule.
- Shen et al. (2018) used CLASH technique to identify piRNA-mRNA interactions transcriptome-wide in *C. elegans*.
- CLASH technique contains false positives.

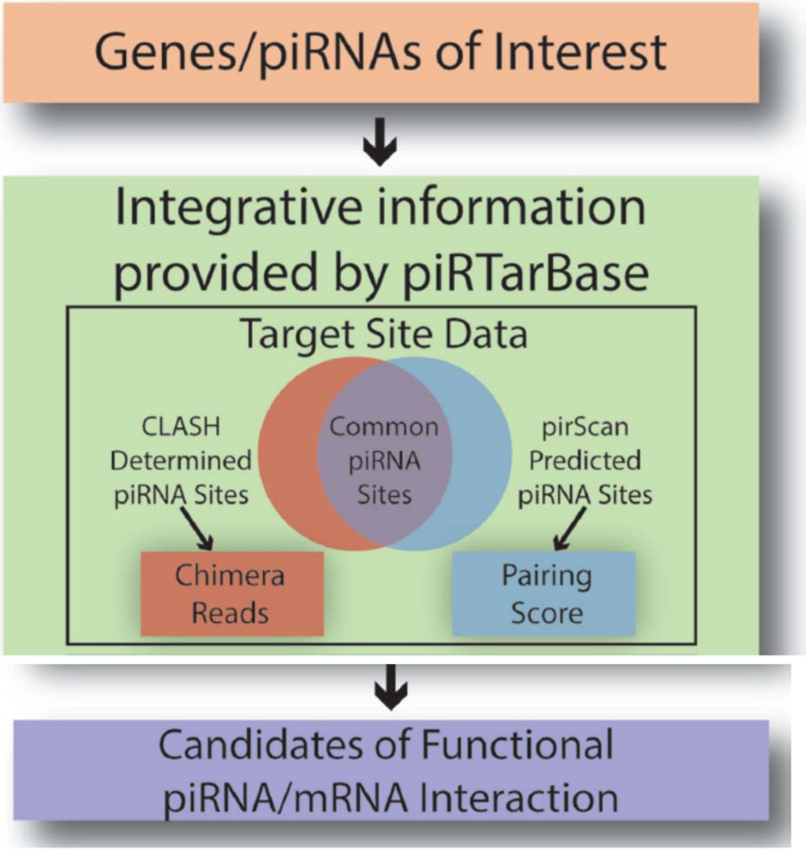


piRTarBase: a database of piRNA targeting sites and their roles in gene regulation

- For each mRNA/piRNA, we provide

1. **predicted piRNA target sites**
   (from our pirScan webtool)

2. **experimental determined piRNA target sites** (from CLASH data)



Genes/piRNAs of Interest

Integrative information provided by piRTarBase

Candidates of Functional piRNA/mRNA Interaction

Input page of piRTarBase

piRTarBase: a database of piRNA targeting sites and their roles in gene regulation

Search predicted piRNA-target pairs in worms

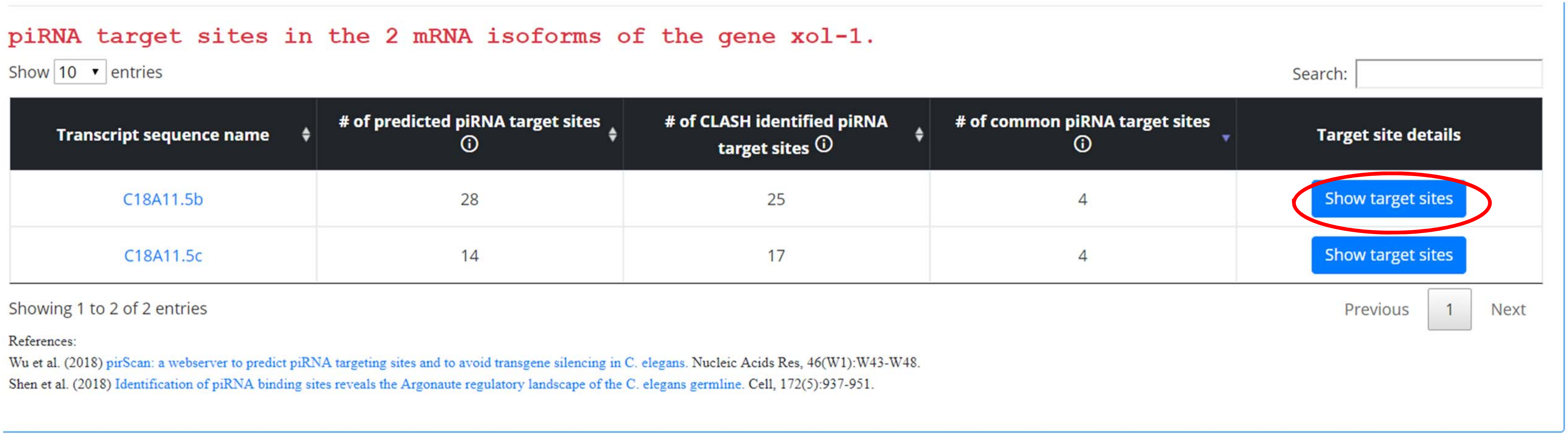
Select a Species : ● C. elegans ○ C. briggsae

Choose piRNA targeting rules : ● stringent ⓘ ○ relaxed ⓘ

Search by : ● gene/transcript ○ piRNA

Input a gene or a transcript (e.g. Y40B10A.2a, xol-1, WBGene00006962**)** xol-1

Search Result



piRNA target sites in the 2 mRNA isoforms of the gene xol-1.

Show 10 ▾ entries

Search:

Transcript sequence name	# of predicted piRNA target sites	# of CLASH identified piRNA target sites	# of common piRNA target sites	Target site details
C18A11.5b	28	25	4	Show target sites
C18A11.5c	14	17	4	Show target sites

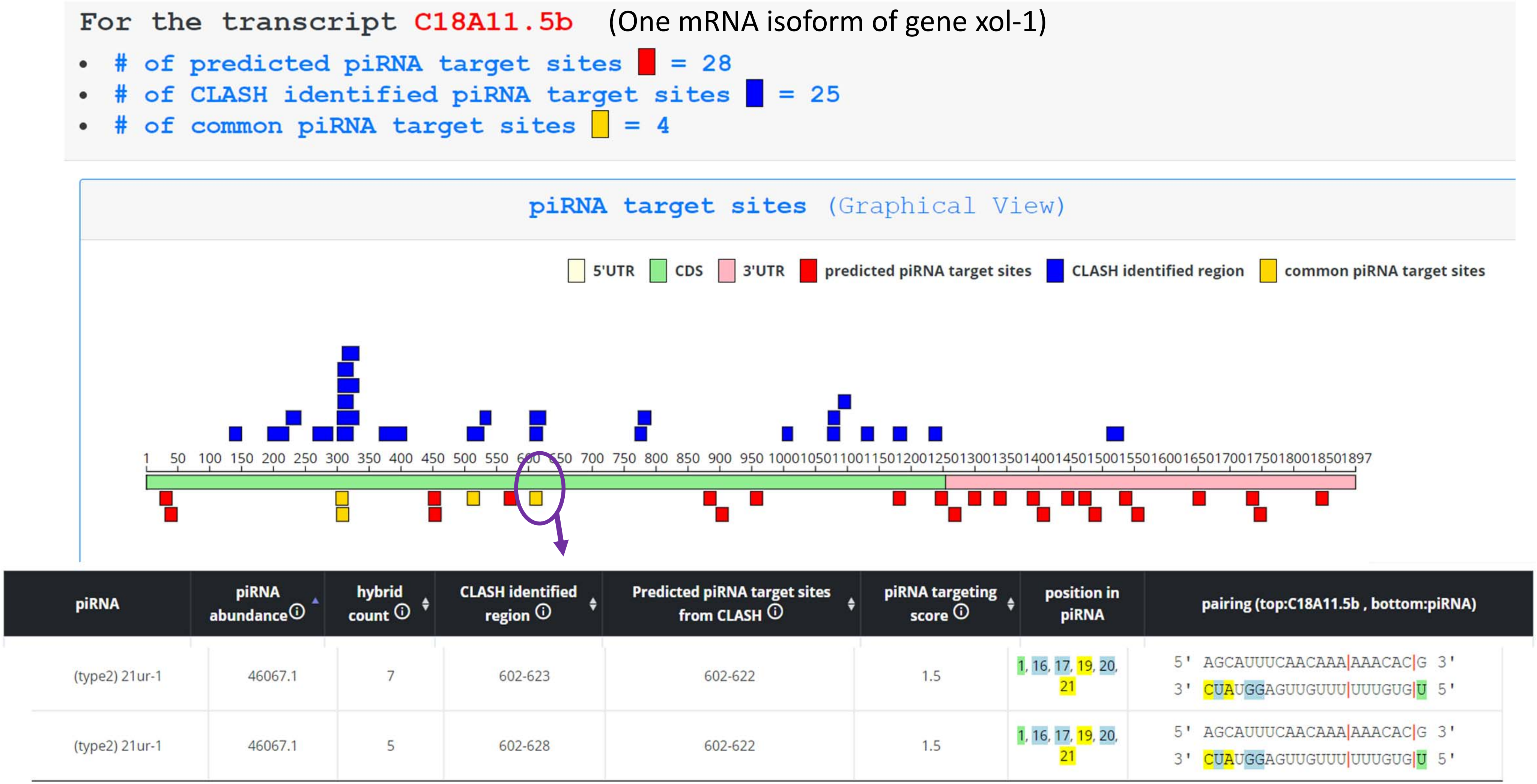
Showing 1 to 2 of 2 entries

Previous 1 Next

References:

Wu et al. (2018) pirScan: a webserver to predict piRNA targeting sites and to avoid transgene silencing in C. elegans. Nucleic Acids Res, 46(W1):W43-W48.

Shen et al. (2018) Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the C. elegans germline. Cell, 172(5):937-951.



For the transcript **C18A11.5b** (One mRNA isoform of gene xol-1)

- # of predicted piRNA target sites 🟥 = 28
- # of CLASH identified piRNA target sites 🟦 = 25
- # of common piRNA target sites 🟨 = 4

piRNA target sites (Graphical View)

piRNA	piRNA abundance ⓘ	hybrid count ⓘ	CLASH identified region ⓘ	Predicted piRNA target sites from CLASH ⓘ	piRNA targeting score ⓘ	position in piRNA	pairing (top:C18A11.5b , bottom:piRNA)
(type2) 21ur-1	46067.1	7	602-623	602-622	1.5	1, 16, 17, 19, 20, 21	5' AGCAUUUCAACAAA|AAACAC|G 3' 3' CUAUGGAGUUGUUU|UUUGUG|U 5'
(type2) 21ur-1	46067.1	5	602-628	602-622	1.5	1, 16, 17, 19, 20, 21	5' AGCAUUUCAACAAA|AAACAC|G 3' 3' CUAUGGAGUUGUUU|UUUGUG|U 5'

Tang et al. (2018). 21ux-1 targets the xol-1 transcript and downregulate its expression.

21ux-1 = (type 2) 21ur-1

Summary of my talk

- piRNA targeting in *C. elegans* prefers near-perfect pairing at the piRNA seed region. In addition, supplementary pairing in the non-seed region also contributes to piRNA targeting, but few mismatches are tolerated.

- piRNA-mediated gene silencing underlies the transgene silencing phenomenon in the germline of C. elegans and we provide a web tool (pirScan) to achieve transgene expression by avoiding piRNA recognition.

- We have constructed a database (pirTarBase) to provide predicted piRNA target sites (from our pirScan webtool) and experimental determined piRNA target sites (from CLASH data) for each mRNA in *C. elegans*.

Acknowledgment

Science Paper	pirScan (NAR paper)	pirTarBase (NAR paper)
• **Heng-Chi Lee's lab (University of Chicago)** Donglei Zhang Shikui Tu Michael Stubna Zhiping Weng Heng-Chi Lee	• **Heng-Chi Lee's lab (University of Chicago)** Jordan Brown Donglei Zhang Shikui Tu Heng-Chi Lee	**Heng-Chi Lee's lab (University of Chicago)** Jordan Brown Shikui Tu Heng-Chi Lee
• **Wei-Sheng Wu's lab (National Cheng Kung University)** Wei-Sheng Wu Wei-Che Huang	• **Wei-Sheng Wu's lab (National Cheng Kung University)** Wei-Sheng Wu Wei-Che Huang	• **Wei-Sheng Wu's lab (National Cheng Kung University)** Wei-Sheng Wu Tsung-Te Chen Yu-Han Chu Wei-Che Huang

Any Questions?

Outlines

- hggg

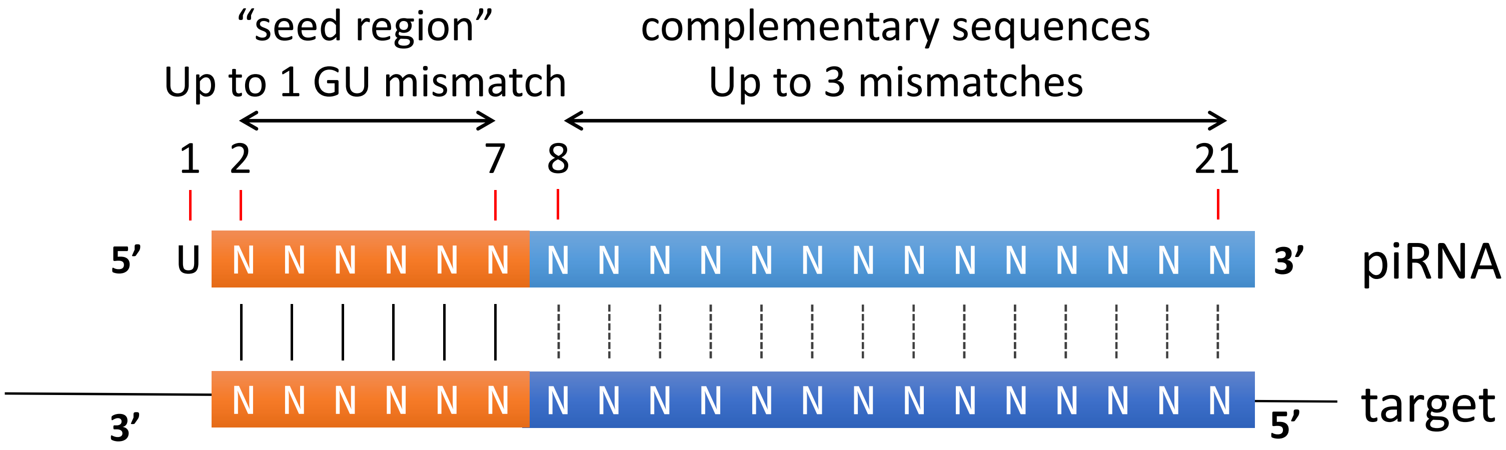
Outlines

- hggg

Outlines

- hggg

piRNA targeting rules in *C. elegans*



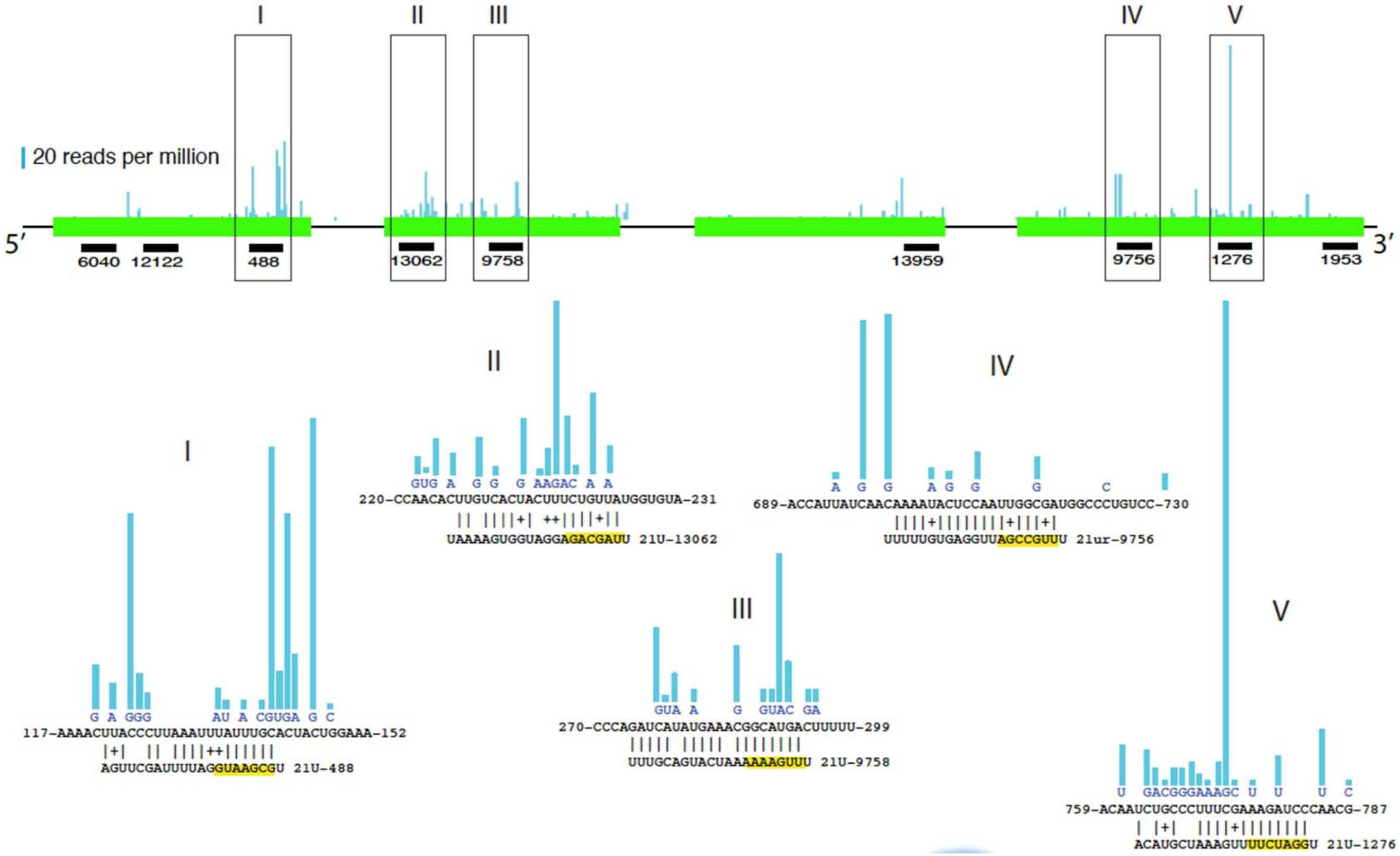
Transgenes carrying foreign nucleic acids are frequently silenced in the germline of *C. elegans*

Transgene	% expression
Cas9	*0*
I-Sce I	*0*
Mos transposase	*0*
gfp::*csr-1*	6
gfp::*rde-3*	17
gfp::*cdk-1*	0
flag::*csr-1*	83
flag::*rde-3*	100

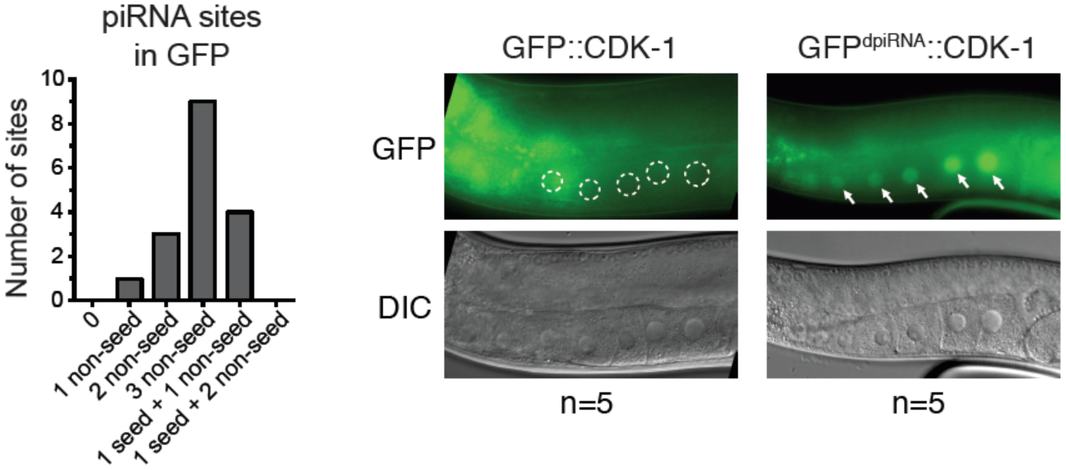
Shirayama M et al., 2012

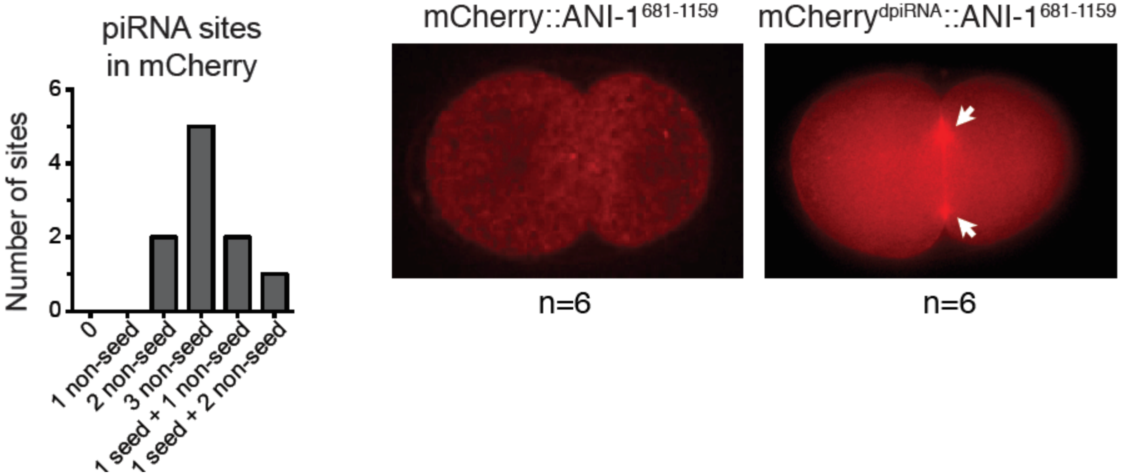
Do piRNA recognize foreign nucleic acids and silence them?

Can we avoid piRNA recognition?

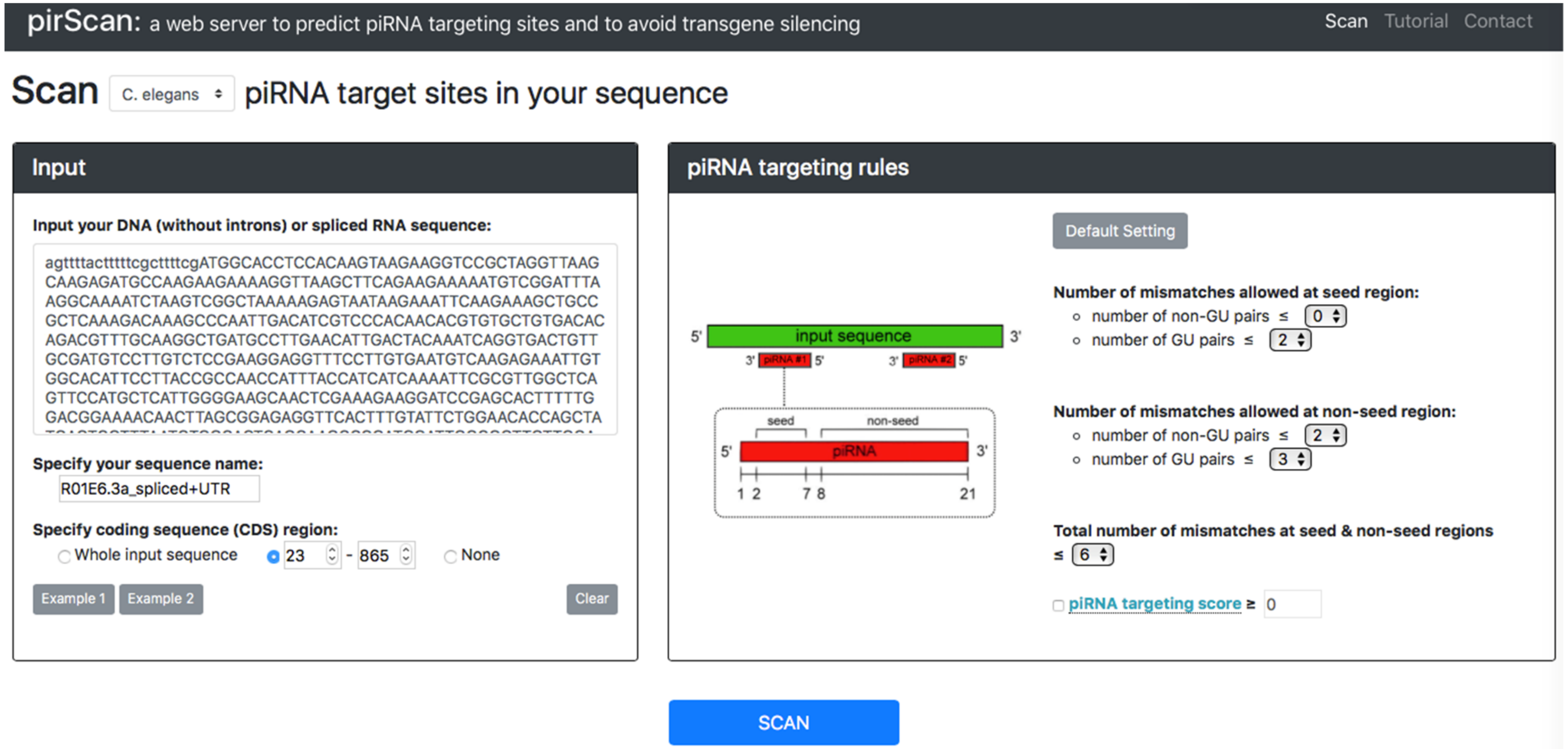


Silencing-prone transgenes can be expressed in the germline by avoiding piRNA targeting





pirScan: a web server to predict piRNA targeting sites and to avoid transgene silencing



Input your DNA (without introns) or spliced RNA sequence:

```
agttttactttttcgcttttcgATGGCACCTCCACAAGTAAGAAGGTCCGCTAGGTTAAG
CAAGAGATGCCAAGAAGAAAAGGTTAAGCTTCAGAAGAAAAATGTCGGATTTA
AGGCAAAATCTAAGTCGGCTAAAAAGAGTAATAAGAAATTCAAGAAAGCTGCC
GCTCAAAGACAAAGCCCAATTGACATCGTCCCACAACACGTGTGCTGTGACAC
AGACGTTTGCAAGGCTGATGCCTTGAACATTGACTACAAATCAGGTGACTGTT
GCGATGTCCTTGTCTCCGAAGGAGGTTTCCTTGTGAATGTCAAGAGAAATTGT
GGCACATTCCTTACCGCCAACCATTTACCATCATCAAAATTCGCGTTGGCTCA
GTTCCATGCTCATTGGGGAAGCAACTCGAAAGAAGGATCCGAGCACTTTTTG
GACGGAAAACAACTTAGCGGAGAGGTTCACTTTGTATTCTGGAACACCAGCTA
```

Specify your sequence name:

R01E6.3a_spliced+UTR

Specify coding sequence (CDS) region:

○ Whole input sequence ● 23 - 865 ○ None

Number of mismatches allowed at seed region:

- number of non-GU pairs ≤ 0
- number of GU pairs ≤ 2

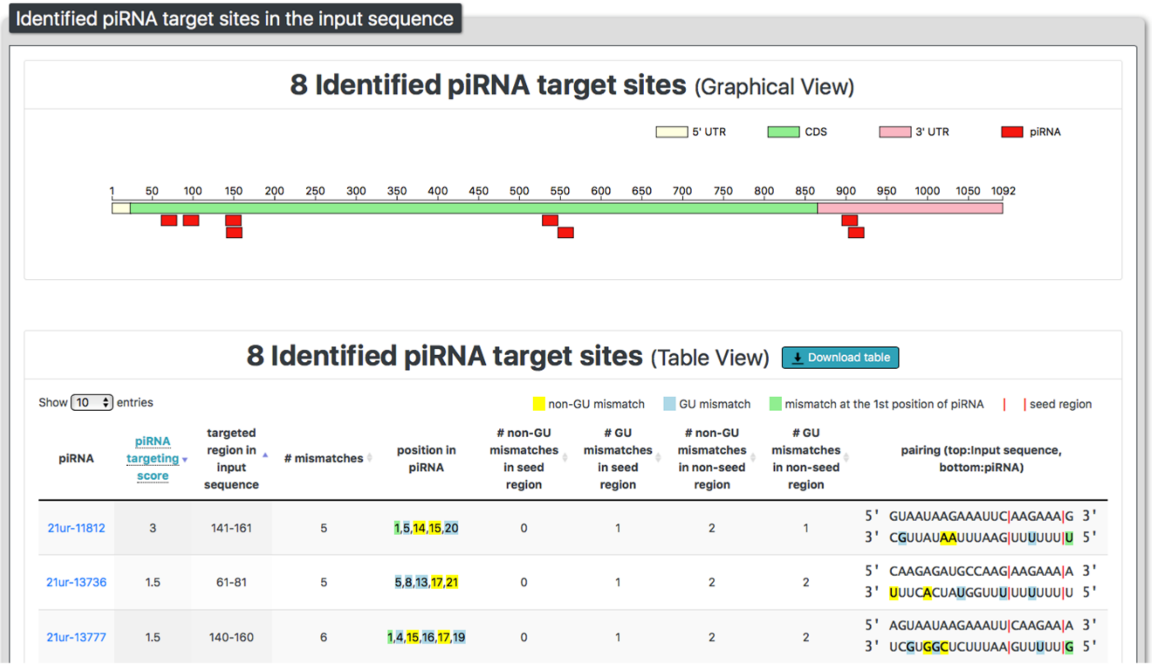
Number of mismatches allowed at non-seed region:

- number of non-GU pairs ≤ 2
- number of GU pairs ≤ 3

Total number of mismatches at seed & non-seed regions
≤ 6

☐ piRNA targeting score ≥ 0

pirScan predicts piRNA sites in the input seqequence and offer choices of silent mutations



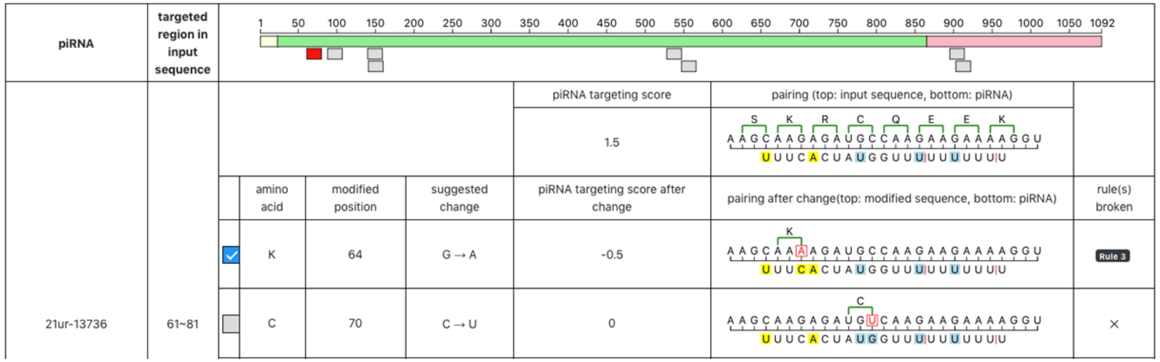
Identified piRNA target sites in the input sequence

8 Identified piRNA target sites (Graphical View)

8 Identified piRNA target sites (Table View)

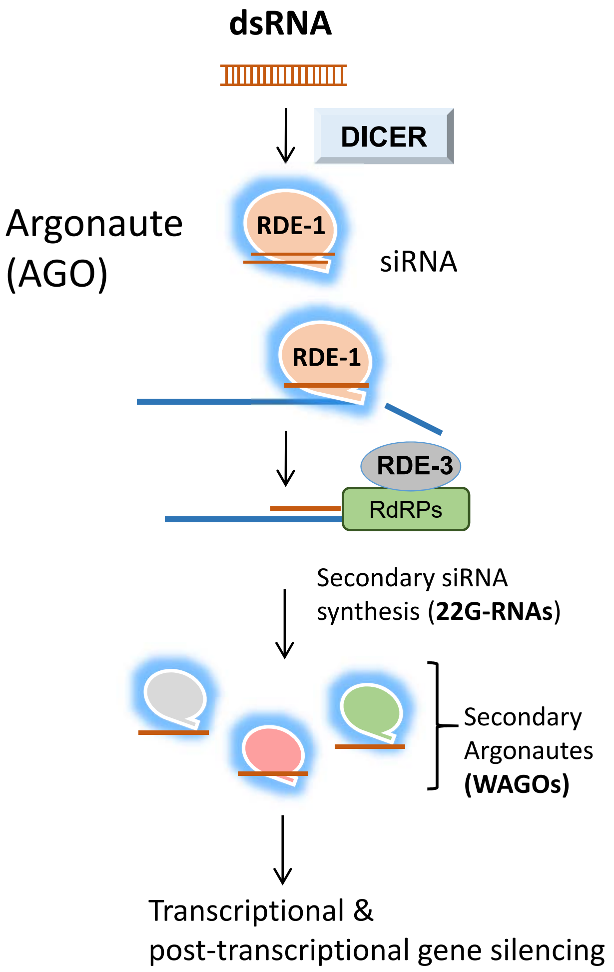
🟨 non-GU mismatch 🟦 GU mismatch 🟩 mismatch at the 1st position of piRNA | seed region

piRNA	piRNA targeting score	targeted region in input sequence	# mismatches	position in piRNA	# non-GU mismatches in seed region	# GU mismatches in seed region	# non-GU mismatches in non-seed region	# GU mismatches in non-seed region	pairing (top:Input sequence, bottom:piRNA)
21ur-11812	3	141-161	5	1,5,14,15,20	0	1	2	1	5' GUAAUAAGAAAUUC|AAGAAA|G 3' 3' CGUUAUAAUUUAAG|UUUUUU|U 5'
21ur-13736	1.5	61-81	5	5,8,13,17,21	0	1	2	2	5' CAAGAGAUGCCAAG|AAGAAA|A 3' 3' UUUCACUAUGGUUU|UUUUUU|U 5'
21ur-13777	1.5	140-160	6	1,4,15,16,17,19	0	1	2	2	5' AGUAAUAAGAAAUU|CAAGAA|A 3' 3' UCGUGGCUCUUUAA|GUUUUU|G 5'



piRNA	targeted region in input sequence		amino acid	modified position	suggested change	piRNA targeting score	pairing (top: input sequence, bottom: piRNA)	
						1.5	A A G C A A G A G A U G C C A A G A A G A A A A G G U U U U C A C U A U G G U U U|U U U U U U|U	
			amino acid	**modified position**	**suggested change**	**piRNA targeting score after change**	**pairing after change(top: modified sequence, bottom: piRNA)**	**rule(s) broken**
		☑	K	64	G → A	-0.5	A A G C A A A A G A U G C C A A G A A G A A A A G G U U U U C A C U A U G G U U U|U U U U U U|U	Rule 3
21ur-13736	61~81	☐	C	70	C → U	0	A A G C A A G A G A U G U C A A G A A G A A A A G G U U U U C A C U A U G G U U U|U U U U U U|U	×

RNAi pathway in C. *elegans*



dsRNA

DICER

Argonaute
(AGO)

RDE-1

siRNA

RDE-1

RDE-3

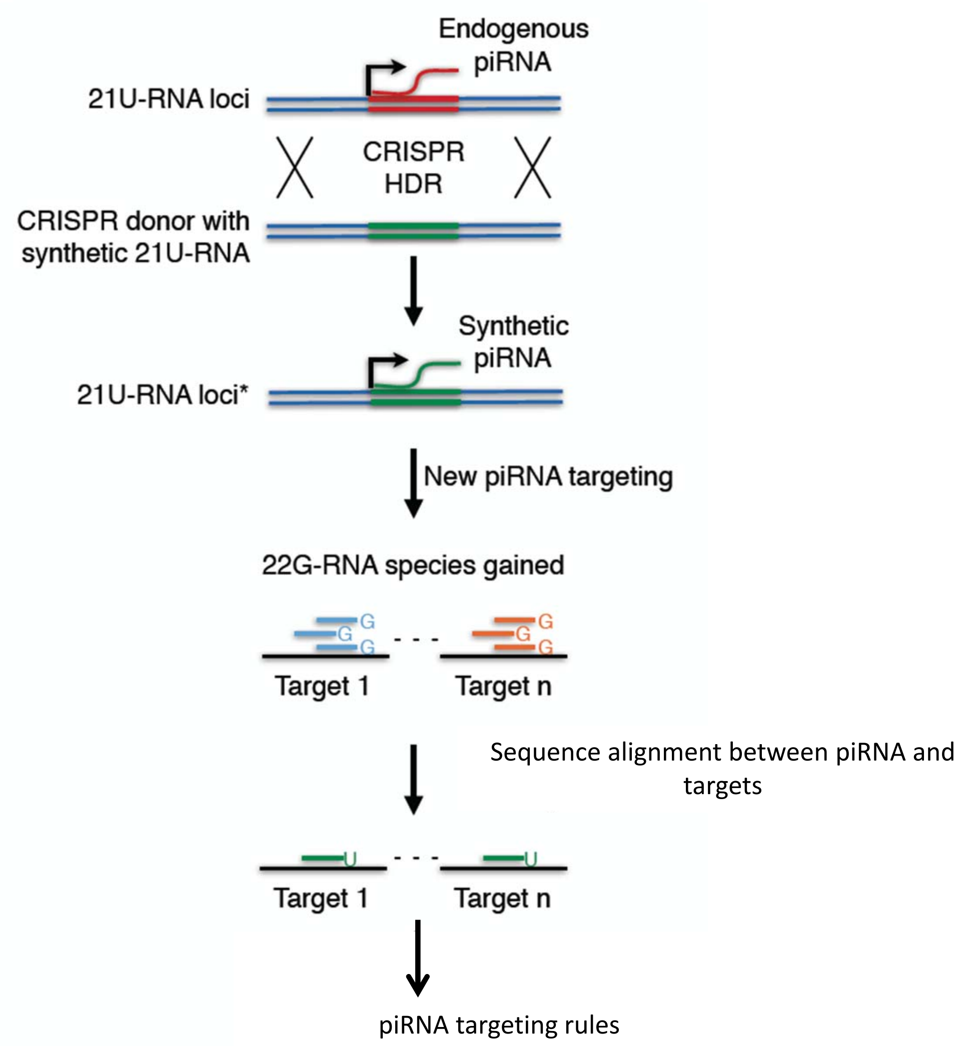
RdRPs

Secondary siRNA
synthesis (**22G-RNAs**)

Secondary
Argonautes
(WAGOs)

Transcriptional &
post-transcriptional gene silencing

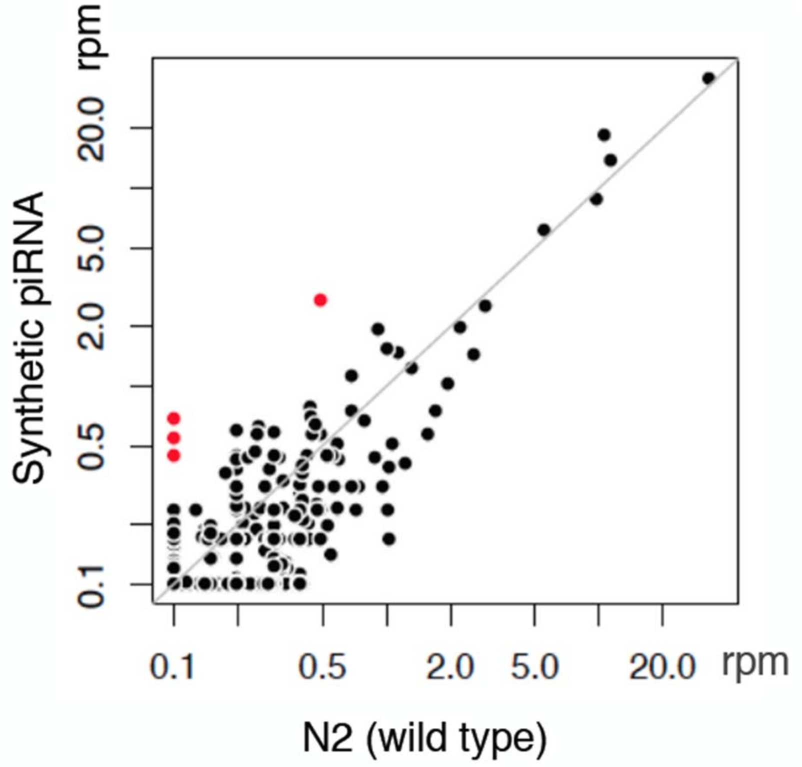
Identify the RNA targets of a single piRNA

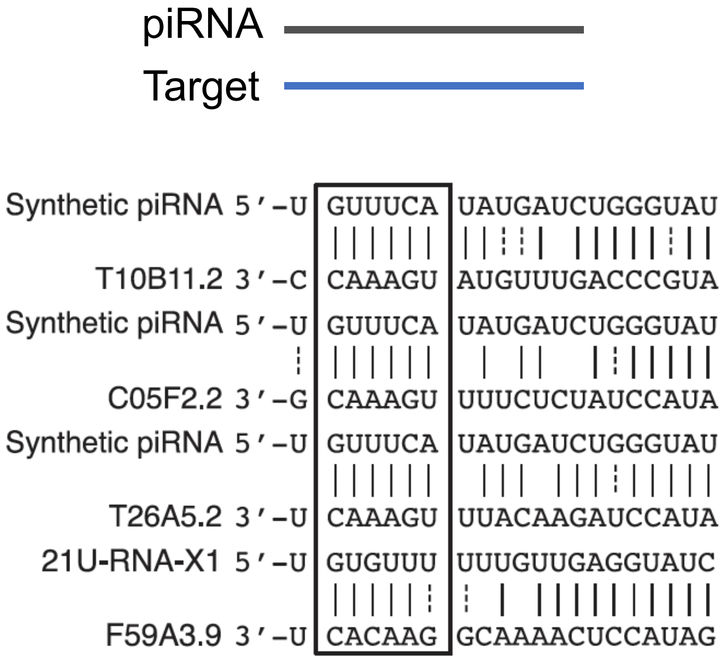


Sequence alignment between piRNA and targets

piRNA targeting rules

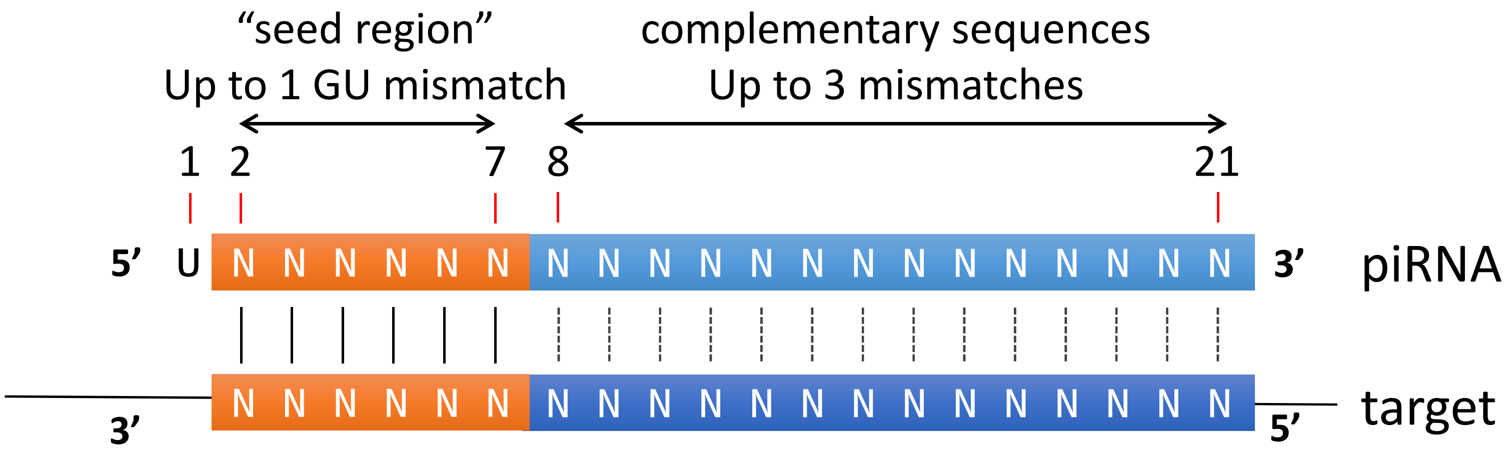
Synthetic piRNAs induce novel 22G-RNA sites and reveal piRNA targeting signature





```
Synthetic piRNA 5′-U GUUUCA UAUGAUCUGGGUAU
                     |||||| ||::| |||||:||
       T10B11.2 3′-C CAAAGU AUGUUUGACCCGUA
Synthetic piRNA 5′-U GUUUCA UAUGAUCUGGGUAU
                   : ||||||  | ||  |:|||||
        C05F2.2 3′-G CAAAGU UUUCUCUAUCCAUA
Synthetic piRNA 5′-U GUUUCA UAUGAUCUGGGUAU
                     ||||||  ||| |||:|||||
        T26A5.2 3′-U CAAAGU UUACAAGAUCCAUA
     21U-RNA-X1 5′-U GUGUUU UUUGUUGAGGUAUC
                     |||||: : | ||||||||||
        F59A3.9 3′-U CACAAG GCAAAACUCCAUAG
```

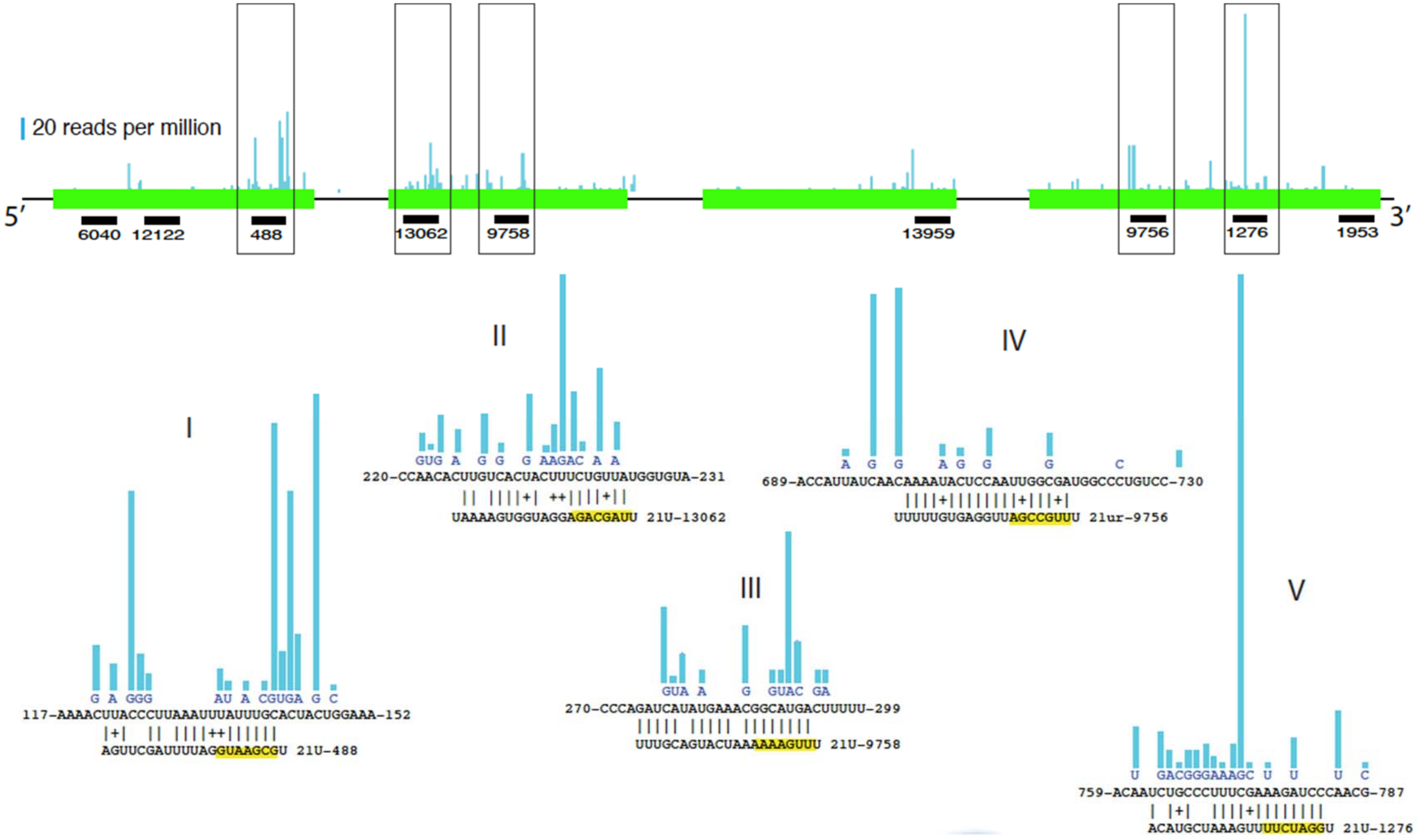
piRNA targeting rules in *C. elegans*



“seed region”
Up to 1 GU mismatch

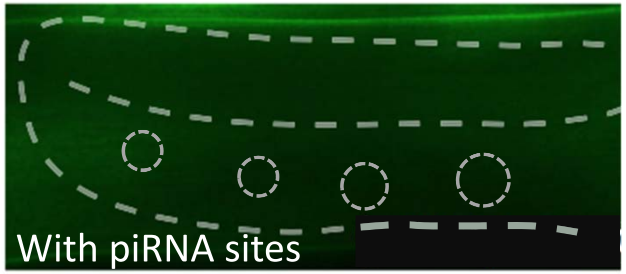
complementary sequences
Up to 3 mismatches

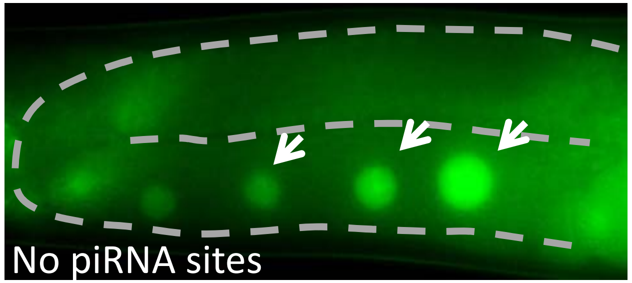
Can we avoid piRNA recognition?

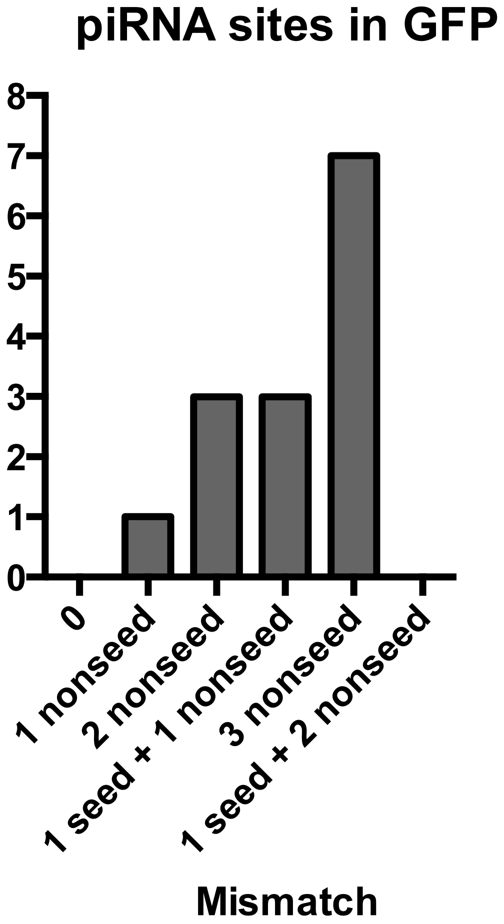


Removing piRNA sites from transgenes allows stable expression in germline cells

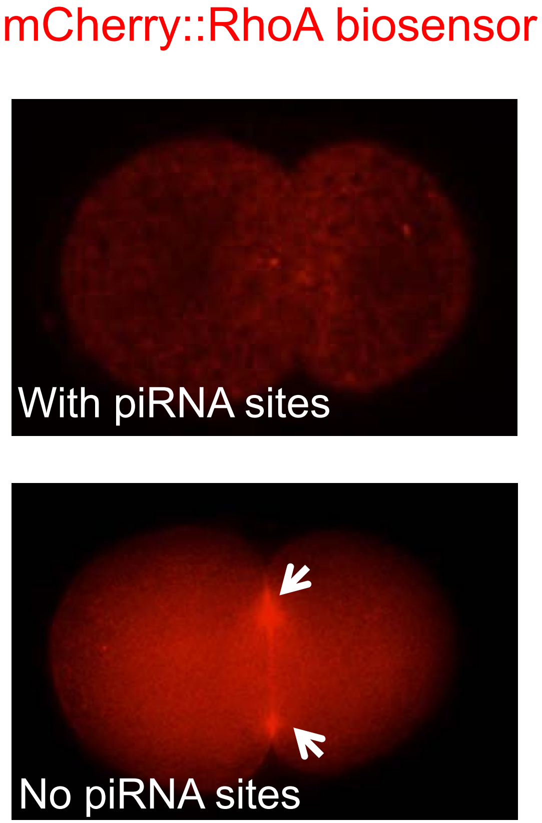
Transgene	% expression
GFP::CSR-1	6
GFP::RDE-3	17
GFP::CDK-1	0

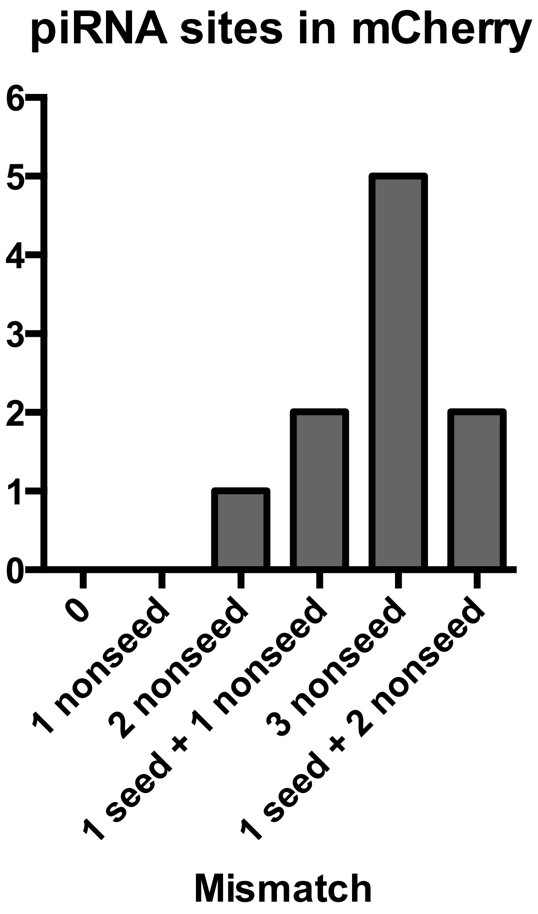




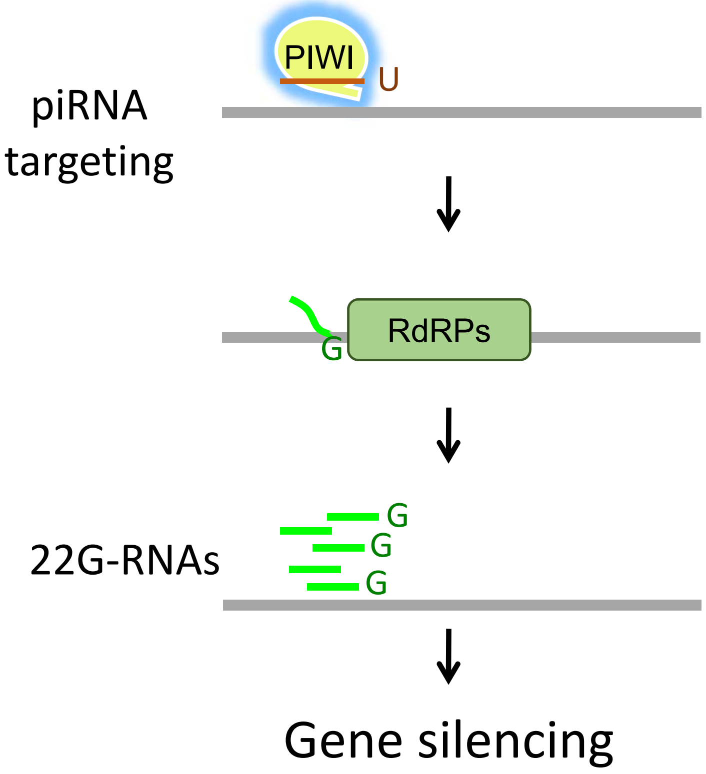


Removing piRNA sites allows mCherry::RhoA biosensor expressed





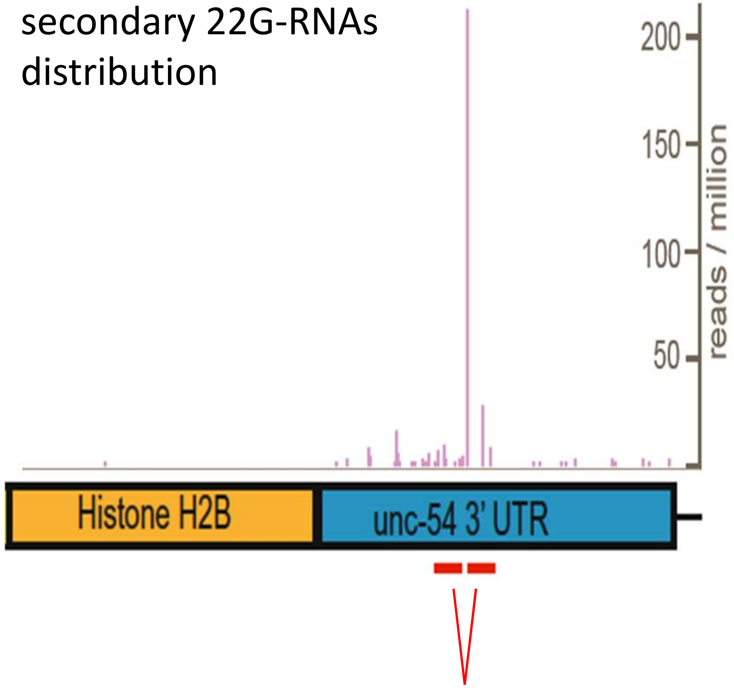
piRNA targeting locally produces 22G-RNAs



piRNA targeting

22G-RNAs

Gene silencing

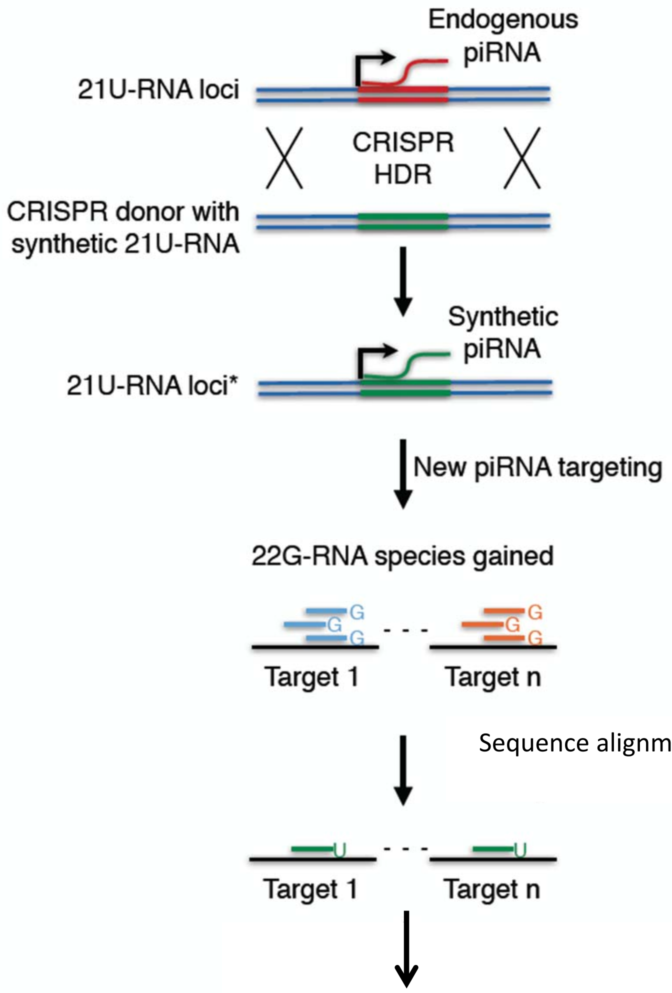


secondary 22G-RNAs distribution

piRNA complementary sites

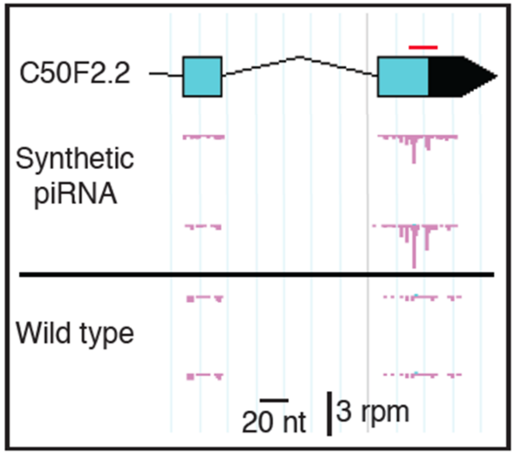
Bagijn MP et al., 2012
Lee HC et al., 2012

22G-RNAs as a proxy to identify piRNA target & rules

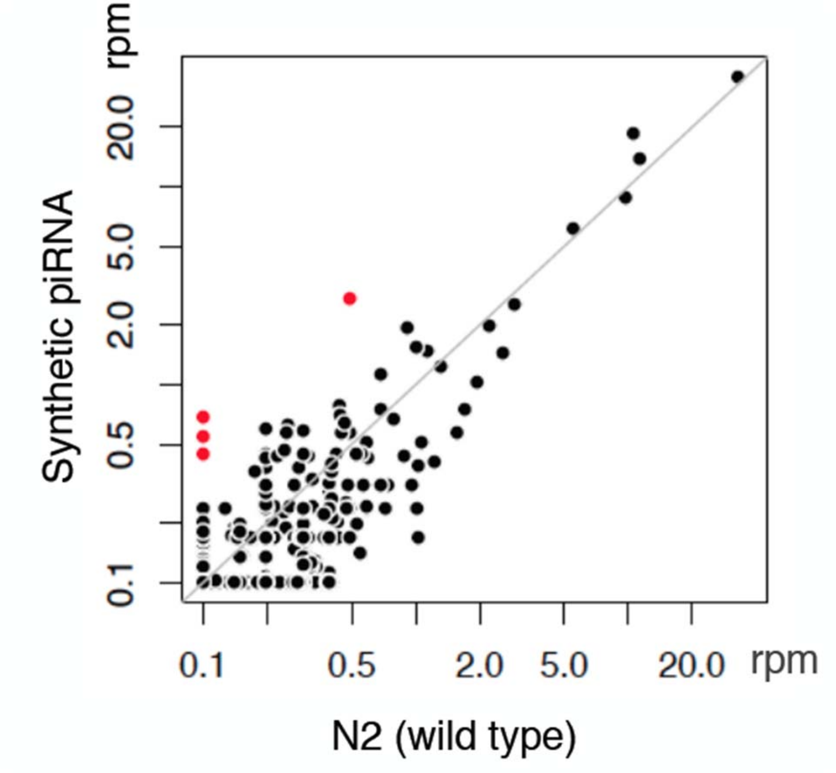


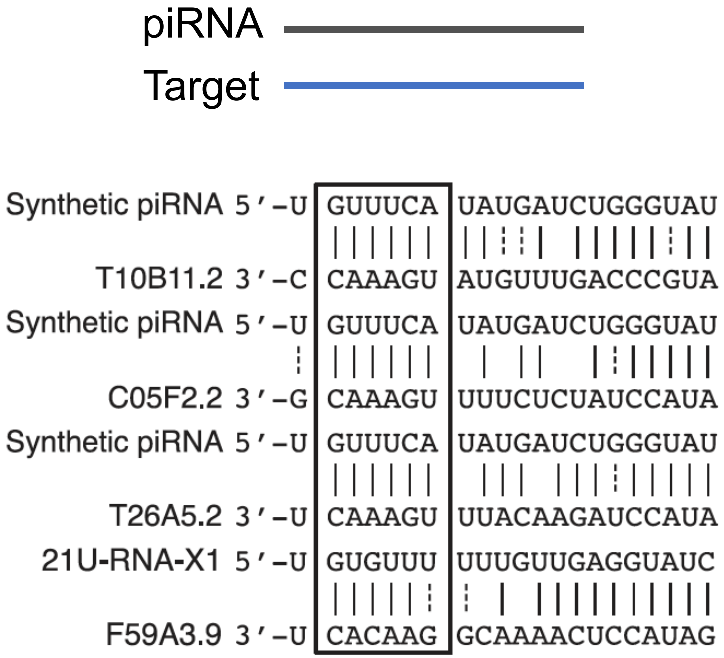
Sequence alignment between piRNA and targets

piRNA targeting rules



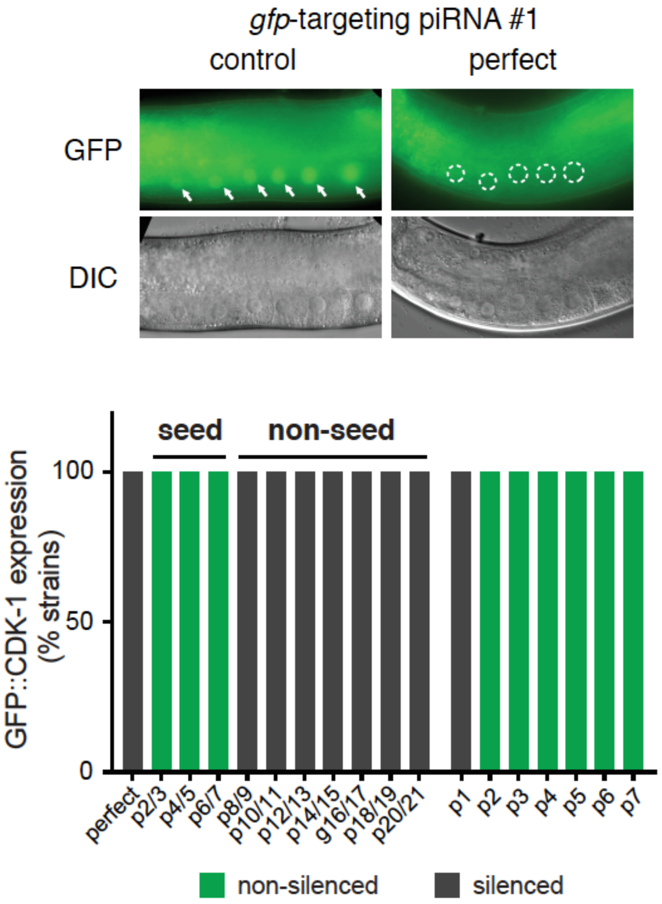
Synthetic piRNAs induce novel 22G-RNA sites and reveal piRNA targeting signature

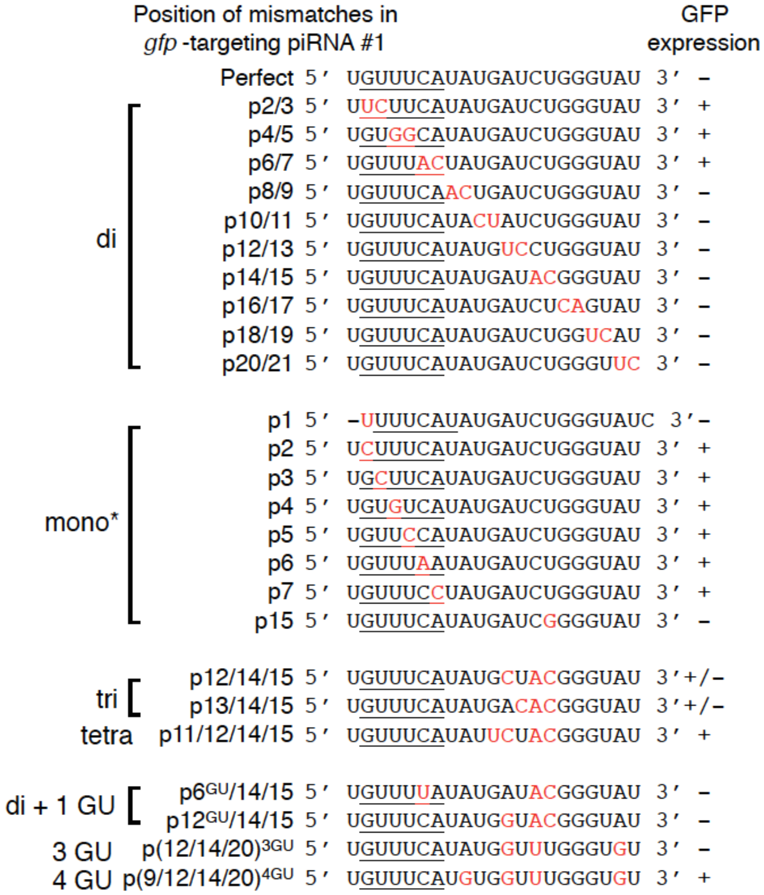




```
Synthetic piRNA 5'-U  GUUUCA  UAUGAUCUGGGUAU
                      ||||||  ||¦¦| |||||¦||
      T10B11.2 3'-C  CAAAGU  AUGUUUGACCCGUA
Synthetic piRNA 5'-U  GUUUCA  UAUGAUCUGGGUAU
                   ¦  ||||||   | ||  |¦|||||
       C05F2.2 3'-G  CAAAGU  UUUCUCUAUCCAUA
Synthetic piRNA 5'-U  GUUUCA  UAUGAUCUGGGUAU
                      ||||||   ||| |||¦|||||
       T26A5.2 3'-U  CAAAGU  UUACAAGAUCCAUA
   21U-RNA-X1 5'-U  GUGUUU  UUUGUUGAGGUAUC
                      |||||¦  ¦ | ||||||||||
       F59A3.9 3'-U  CACAAG  GCAAAACUCCAUAG
```

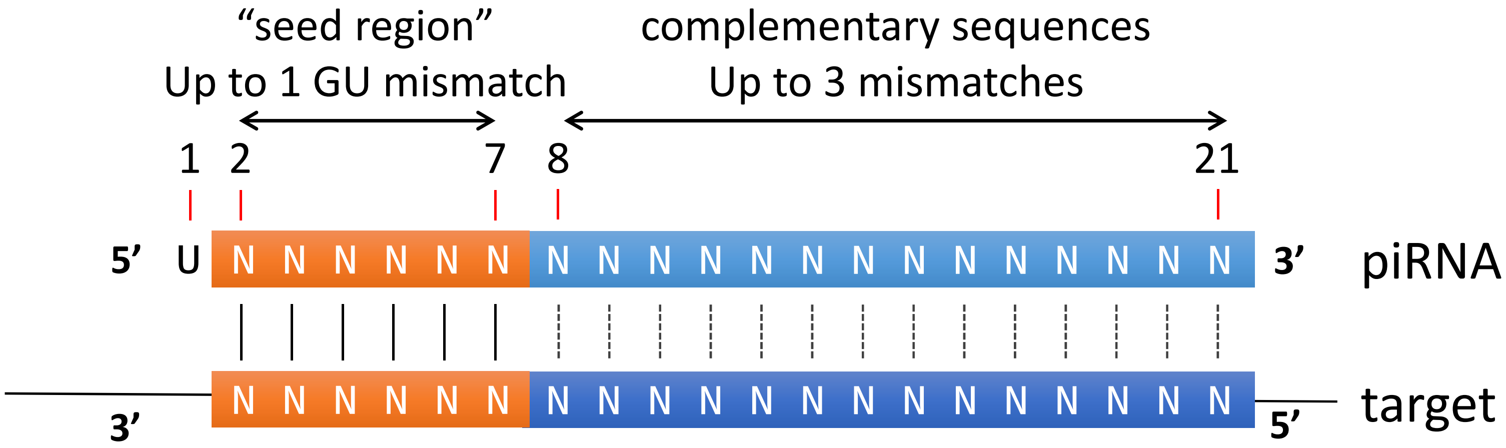
A piRNA reporter assay to investigate the piRNA targeting rules





	Position of mismatches in *gfp* -targeting piRNA #1				GFP expression
	Perfect	5′	UGUUUCAUAUGAUCUGGGUAU	3′	–
di	p2/3	5′	UUCUUCAUAUGAUCUGGGUAU	3′	+
	p4/5	5′	UGUGGCAUAUGAUCUGGGUAU	3′	+
	p6/7	5′	UGUUUACUAUGAUCUGGGUAU	3′	+
	p8/9	5′	UGUUUCAACUGAUCUGGGUAU	3′	–
	p10/11	5′	UGUUUCAUACUAUCUGGGUAU	3′	–
	p12/13	5′	UGUUUCAUAUGUCCUGGGUAU	3′	–
	p14/15	5′	UGUUUCAUAUGAUACGGGUAU	3′	–
	p16/17	5′	UGUUUCAUAUGAUCUCAGUAU	3′	–
	p18/19	5′	UGUUUCAUAUGAUCUGGUCAU	3′	–
	p20/21	5′	UGUUUCAUAUGAUCUGGGUUC	3′	–
mono*	p1	5′	–UUUUCAUAUGAUCUGGGUAUC	3′	–
	p2	5′	UCUUUCAUAUGAUCUGGGUAU	3′	+
	p3	5′	UGCUUCAUAUGAUCUGGGUAU	3′	+
	p4	5′	UGUGUCAUAUGAUCUGGGUAU	3′	+
	p5	5′	UGUUCCAUAUGAUCUGGGUAU	3′	+
	p6	5′	UGUUUAAUAUGAUCUGGGUAU	3′	+
	p7	5′	UGUUUCCUAUGAUCUGGGUAU	3′	+
	p15	5′	UGUUUCAUAUGAUCGGGGUAU	3′	–
tri	p12/14/15	5′	UGUUUCAUAUGCUACGGGUAU	3′	+/–
	p13/14/15	5′	UGUUUCAUAUGACACGGGUAU	3′	+/–
tetra	p11/12/14/15	5′	UGUUUCAUAUUCUACGGGUAU	3′	+
di + 1 GU	p6^{GU}/14/15	5′	UGUUUUAUAUGAUACGGGUAU	3′	–
	p12^{GU}/14/15	5′	UGUUUCAUAUGGUACGGGUAU	3′	–
3 GU	p(12/14/20)^{3GU}	5′	UGUUUCAUAUGGUUUGGGUGU	3′	–
4 GU	p(9/12/14/20)^{4GU}	5′	UGUUUCAUGUGGUUUGGGUGU	3′	+

piRNA targeting rules in *C. elegans*



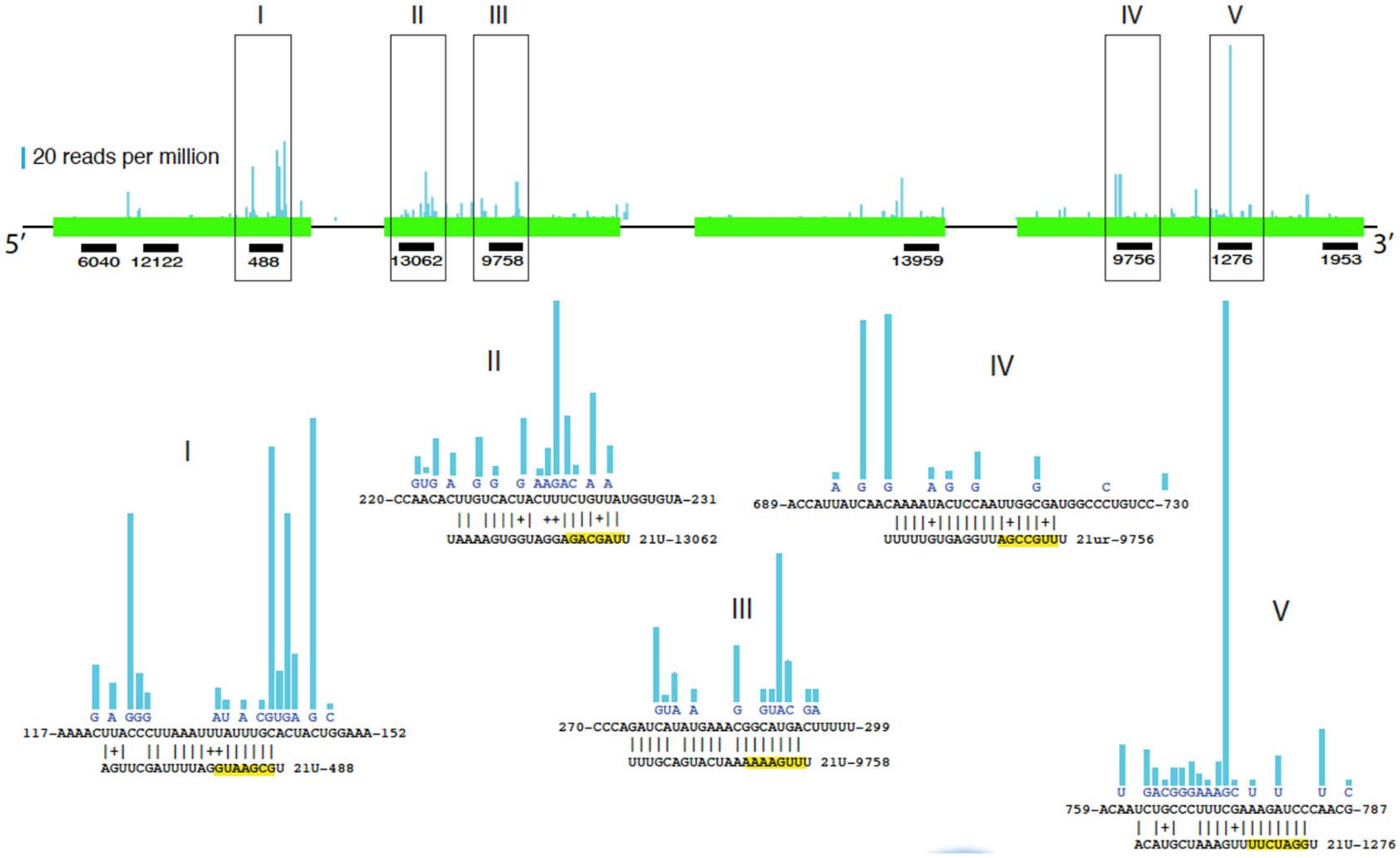
Transgenes carrying foreign nucleic acids are frequently silenced in the germline of *C. elegans*

Transgene	% expression
Cas9	*0*
I-Sce I	*0*
Mos transposase	*0*
gfp::csr-1	6
gfp::rde-3	17
gfp::cdk-1	0
flag::csr-1	83
flag::rde-3	100

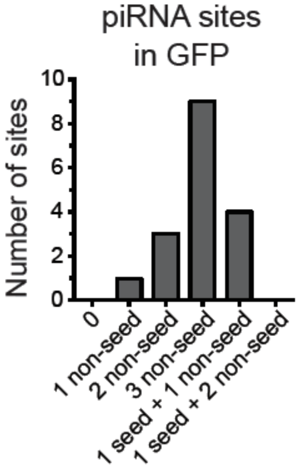
Shirayama M et al., 2012

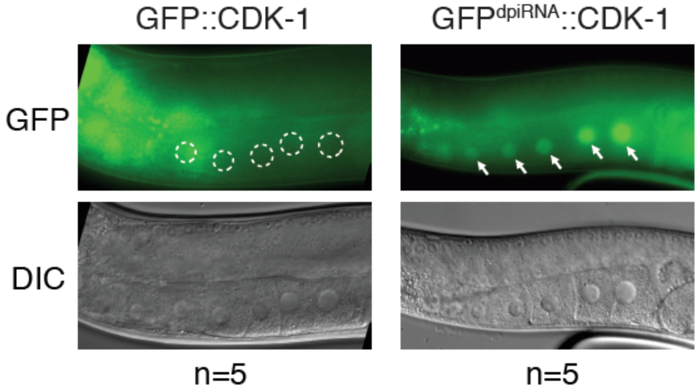
Do piRNA recognize foreign nucleic acids and silence them?

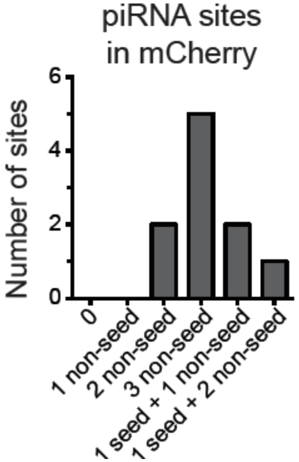
Can we avoid piRNA recognition?

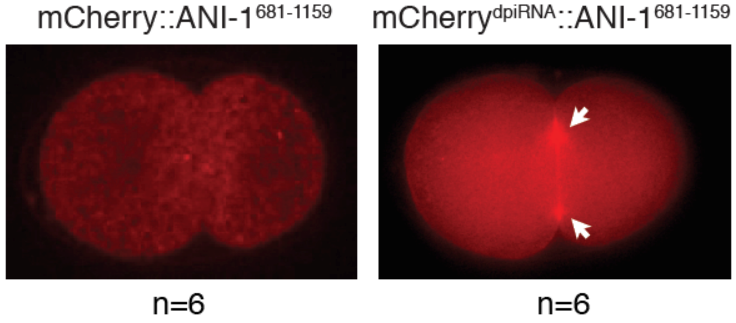


Silencing-prone transgenes can be expressed in the germline by avoiding piRNA targeting

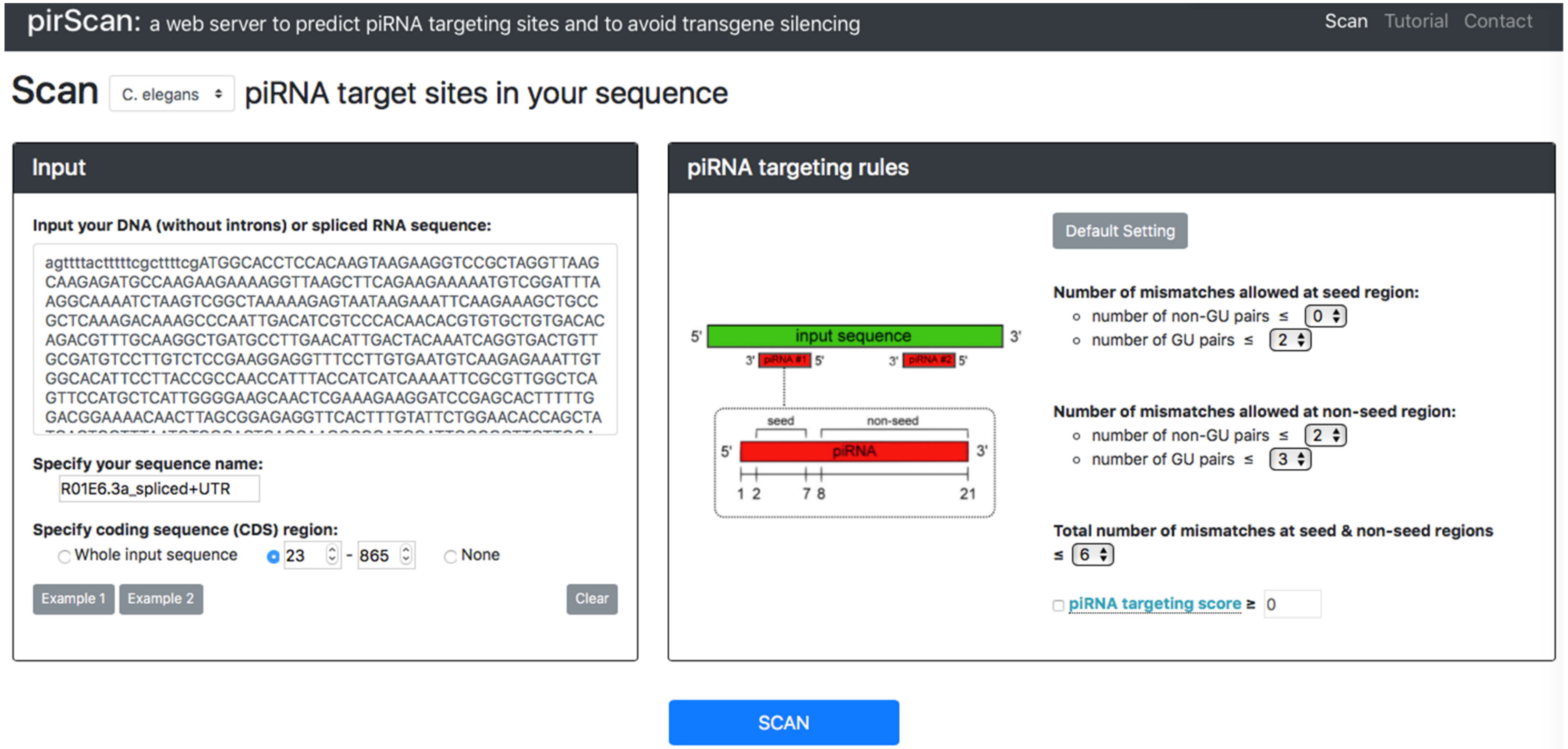








pirScan: a web server to predict piRNA targeting sites and to avoid transgene silencing



Scan C. elegans piRNA target sites in your sequence

Input

Input your DNA (without introns) or spliced RNA sequence:

```
agttttactttttcgcttttcgATGGCACCTCCACAAGTAAGAAGGTCCGCTAGGTTAAG
CAAGAGATGCCAAGAAGAAAAGGTTAAGCTTCAGAAGAAAAATGTCGGATTTA
AGGCAAAATCTAAGTCGGCTAAAAAGAGTAATAAGAAATTCAAGAAAGCTGCC
GCTCAAAGACAAAGCCCAATTGACATCGTCCCACAACACGTGTGCTGTGACAC
AGACGTTTGCAAGGCTGATGCCTTGAACATTGACTACAAATCAGGTGACTGTT
GCGATGTCCTTGTCTCCGAAGGAGGTTTCCTTGTGAATGTCAAGAGAAATTGT
GGCACATTCCTTACCGCCAACCATTTACCATCATCAAAATTCGCGTTGGCTCA
GTTCCATGCTCATTGGGGAAGCAACTCGAAAGAAGGATCCGAGCACTTTTTG
GACGGAAAACAACTTAGCGGAGAGGTTCACTTTGTATTCTGGAACACCAGCTA
```

Specify your sequence name: R01E6.3a_spliced+UTR

Specify coding sequence (CDS) region:
○ Whole input sequence ● 23 - 865 ○ None

piRNA targeting rules

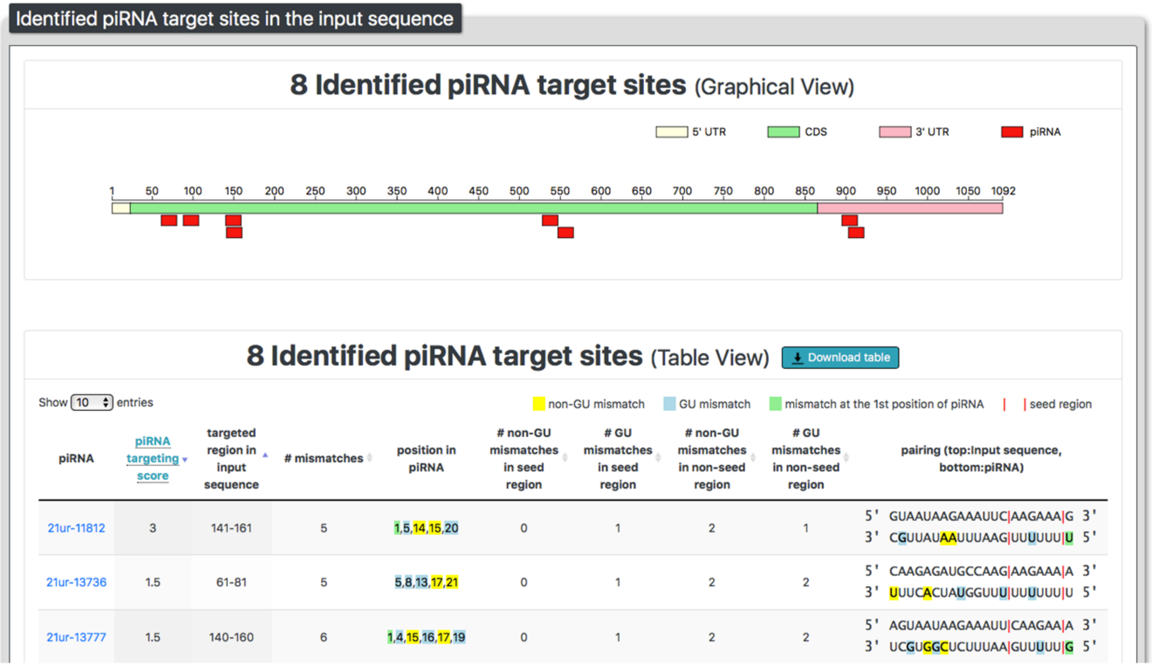
Number of mismatches allowed at seed region:
- number of non-GU pairs ≤ 0
- number of GU pairs ≤ 2

Number of mismatches allowed at non-seed region:
- number of non-GU pairs ≤ 2
- number of GU pairs ≤ 3

Total number of mismatches at seed & non-seed regions
≤ 6

☐ piRNA targeting score ≥ 0

pirScan predicts piRNA sites in the input seqequence and offer choices of silent mutations



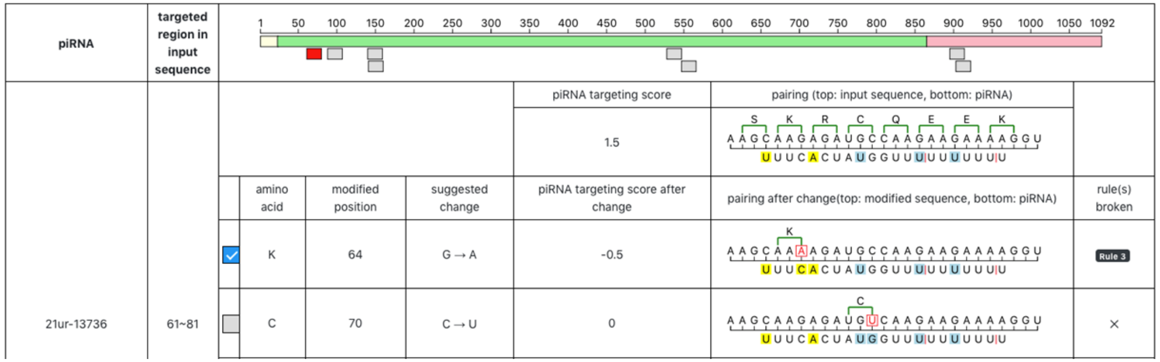
Identified piRNA target sites in the input sequence

8 Identified piRNA target sites (Graphical View)

8 Identified piRNA target sites (Table View)

🟨 non-GU mismatch 🟦 GU mismatch 🟩 mismatch at the 1st position of piRNA | | seed region

piRNA	piRNA targeting score	targeted region in input sequence	# mismatches	position in piRNA	# non-GU mismatches in seed region	# GU mismatches in seed region	# non-GU mismatches in non-seed region	# GU mismatches in non-seed region	pairing (top:Input sequence, bottom:piRNA)
21ur-11812	3	141-161	5	1,5,14,15,20	0	1	2	1	5' GUAAUAAGAAAUUC|AAGAAA|G 3' 3' CGUUAUAAUUUAAG|UUUUUU|U 5'
21ur-13736	1.5	61-81	5	5,8,13,17,21	0	1	2	2	5' CAAGAGAUGCCAAG|AAGAAA|A 3' 3' UUUCACUAUGGUUU|UUUUUU|U 5'
21ur-13777	1.5	140-160	6	1,4,15,16,17,19	0	1	2	2	5' AGUAAUAAGAAAUU|CAAGAA|A 3' 3' UCGUGGCUCUUUAA|GUUUUU|G 5'



piRNA	targeted region in input sequence					piRNA targeting score	pairing (top: input sequence, bottom: piRNA)	
						1.5	S K R C Q E E K A A G C A A G A G A U G C C A A G A A G A A A A G G U U U U C A C U A U G G U U U|U U U U U U|U	
			amino acid	modified position	suggested change	piRNA targeting score after change	pairing after change(top: modified sequence, bottom: piRNA)	rule(s) broken
		☑	K	64	G → A	-0.5	K A A G C A A A A G A U G C C A A G A A G A A A A G G U U U U C A C U A U G G U U U|U U U U U U|U	Rule 3
21ur-13736	61~81	☐	C	70	C → U	0	C A A G C A A G A G A U G U C A A G A A G A A A A G G U U U U C A C U A U G G U U U|U U U U U U|U	×

How do piRNAs recognize their targets?

siRNA Argonaute

near-perfect matches

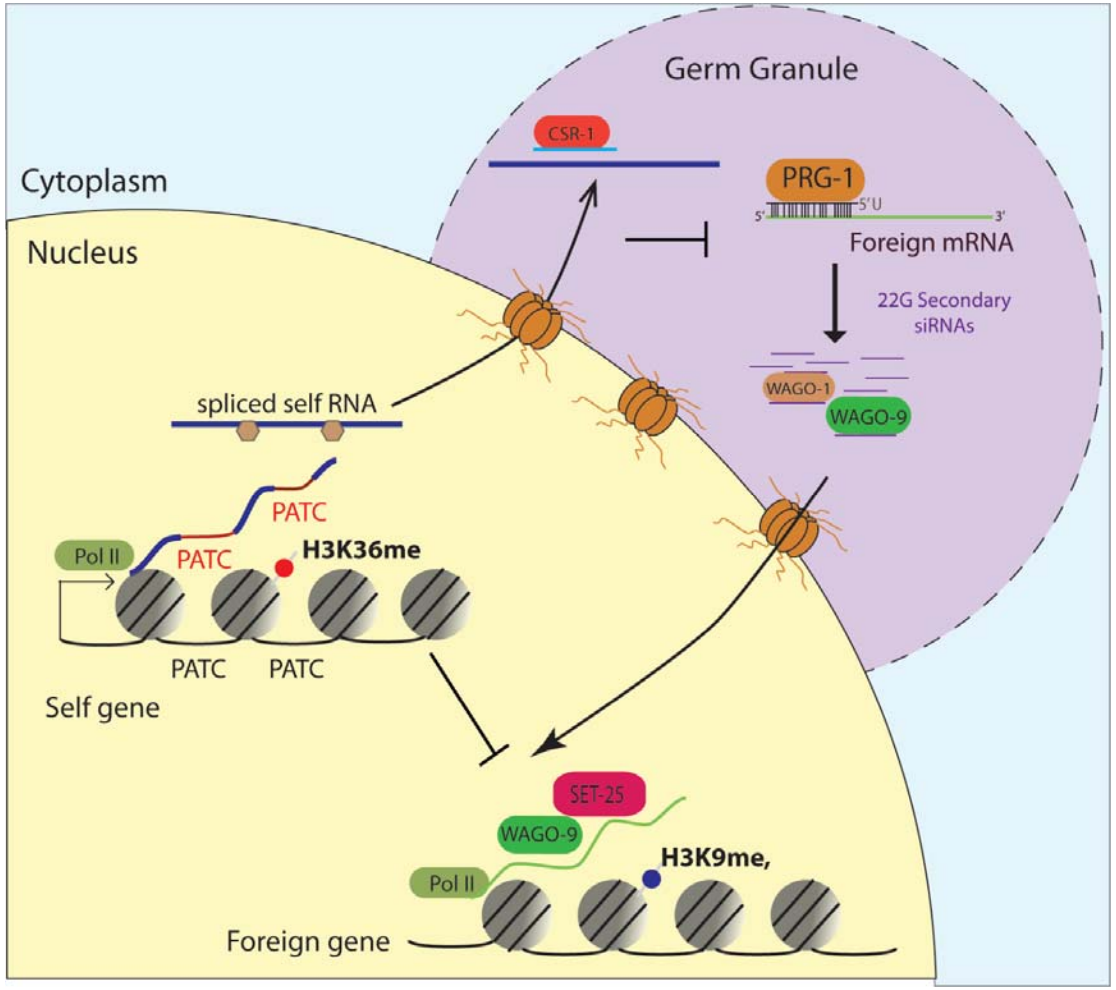
miRNA Argonaute

base pairing at
“Seed region”

PIWI Argonaute

seed pairing;
mismatch tolerant
non-seed pairing

A model for piRNA mediated genome defense



How do piRNAs recognize their targets?

siRNA Argonaute

near-perfect matches

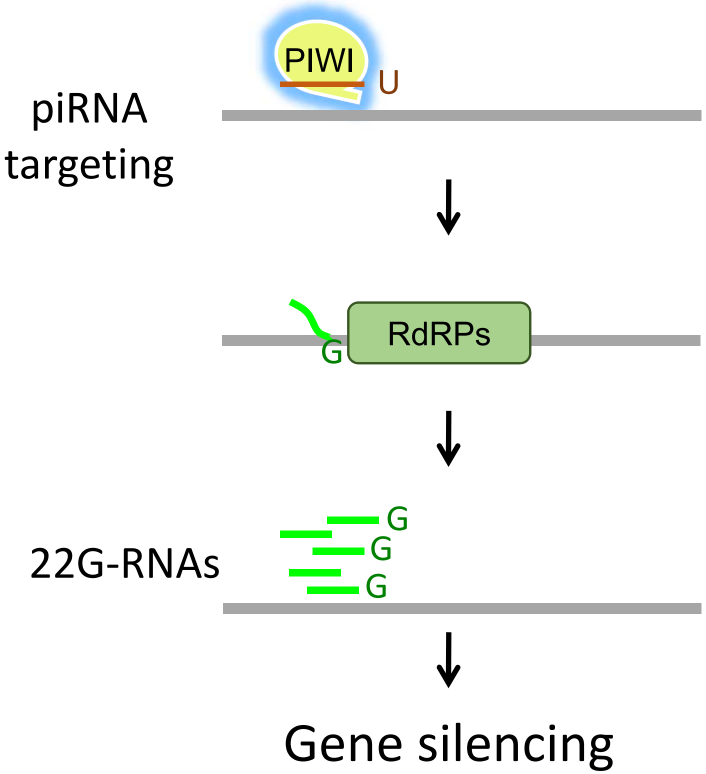
miRNA Argonaute

base pairing at
“Seed region”

PIWI Argonaute

?

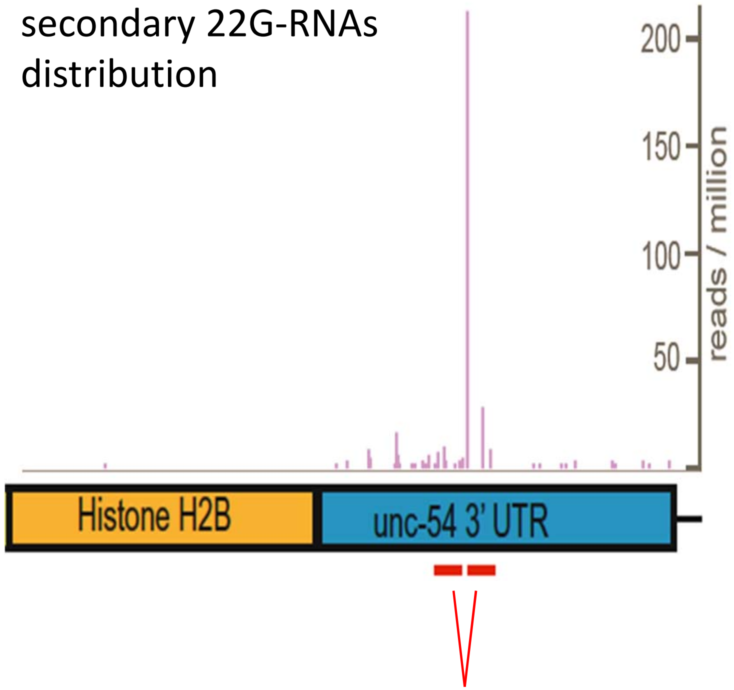
piRNA targeting locally produces 22G-RNAs



piRNA targeting

22G-RNAs

Gene silencing



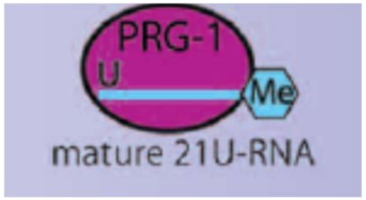
secondary 22G-RNAs distribution

piRNA complementary sites

Bagijn MP et al., 2012
Lee HC et al., 2012

Function of piRNA

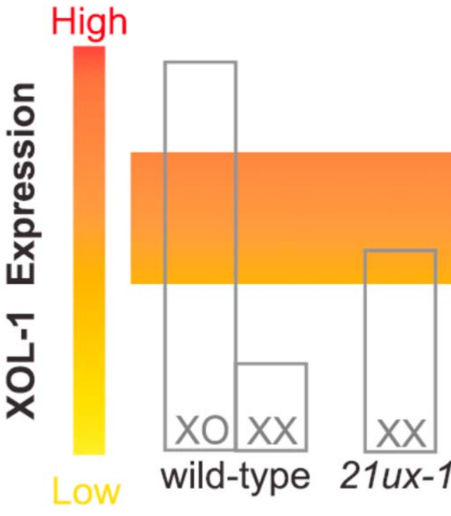
- piRNAs interact with **Piwi-related Argonaute proteins** (**PRG-1** in *C. elegans*) to **defend the genome against foreign nucleic acids** (e.g. **transposons** or **transgenes**).



- piRNA function extends beyond genome defense.
- piRNAs can also regulate the expression of endogenous genes in *C. elgans*.

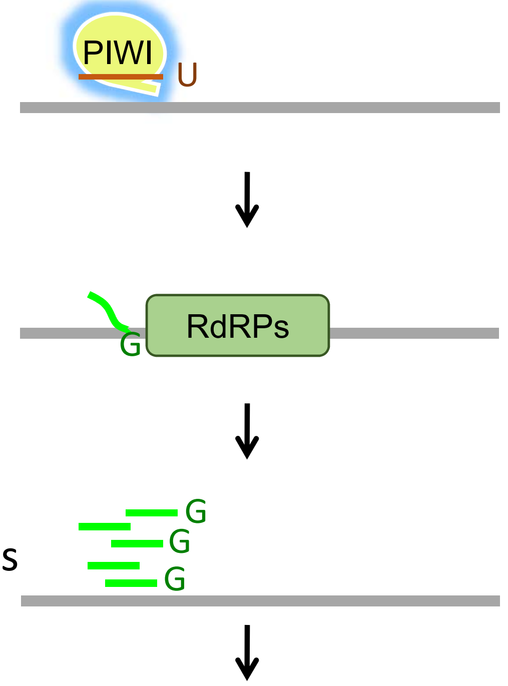
- The X chromosome-derived piRNA 21ux-1 **downregulates** XOL-1 (XO Lethal), a master regulator of X chromosome dosage compensation and sex determination in *C. elegans*.

- Mutations in 21ux-1 sensitize hermaphrodites to dosage compensation and sex determination defects.



Tang et al., 2018, Developmental Cell *44*, 762–770

22G-RNAs can serve as a “mark” for piRNA targeting sites *in vivo*



piRNA targeting

↓

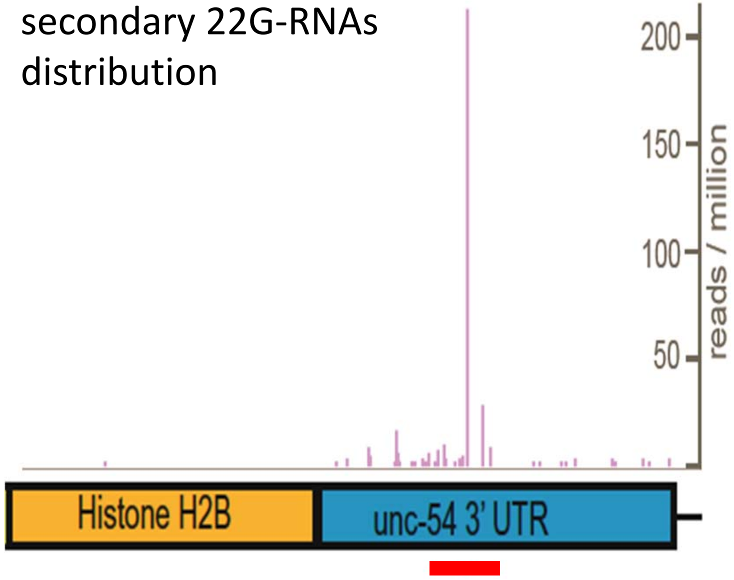
↓

22G-RNAs

↓

Gene silencing

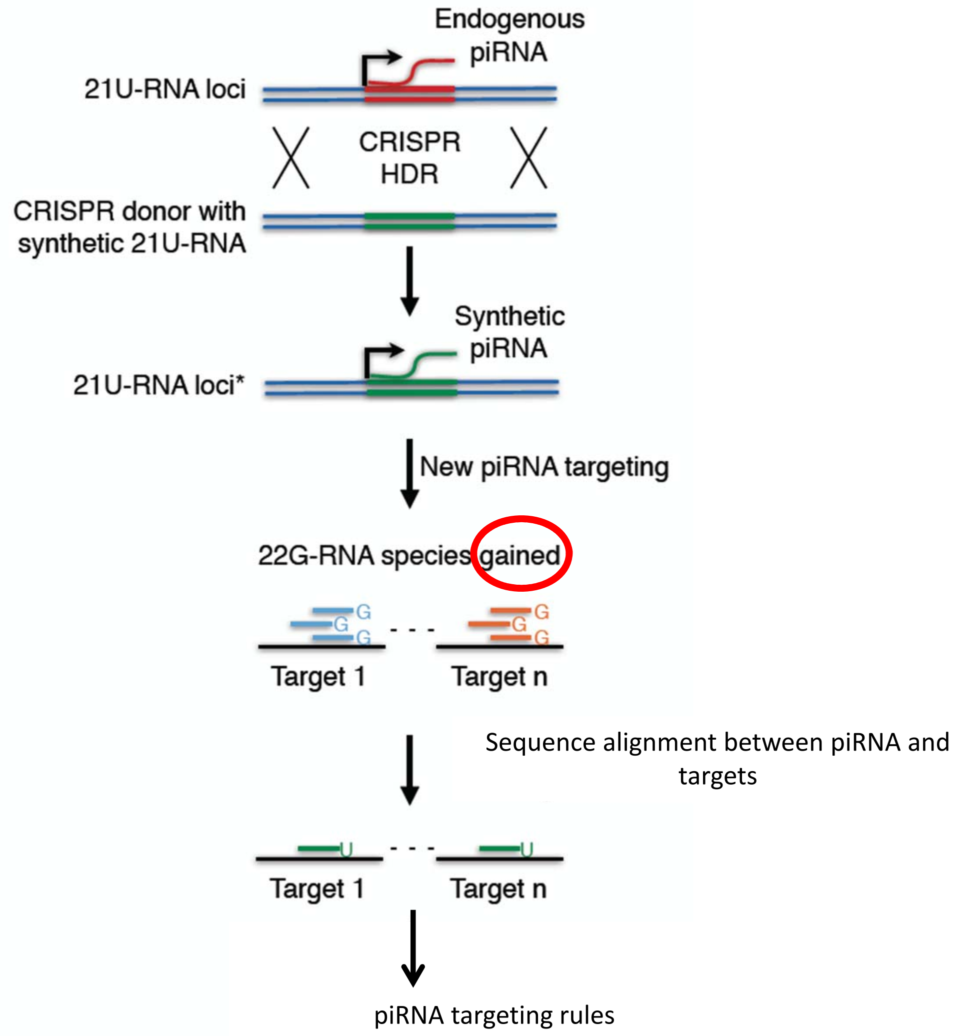
Bagijn MP et al., 2012
Lee HC et al., 2012



secondary 22G-RNAs distribution

piRNA complementary sites

Identify the mRNA targets of a single piRNA



Endogenous piRNA

21U-RNA loci

CRISPR HDR

CRISPR donor with synthetic 21U-RNA

Synthetic piRNA

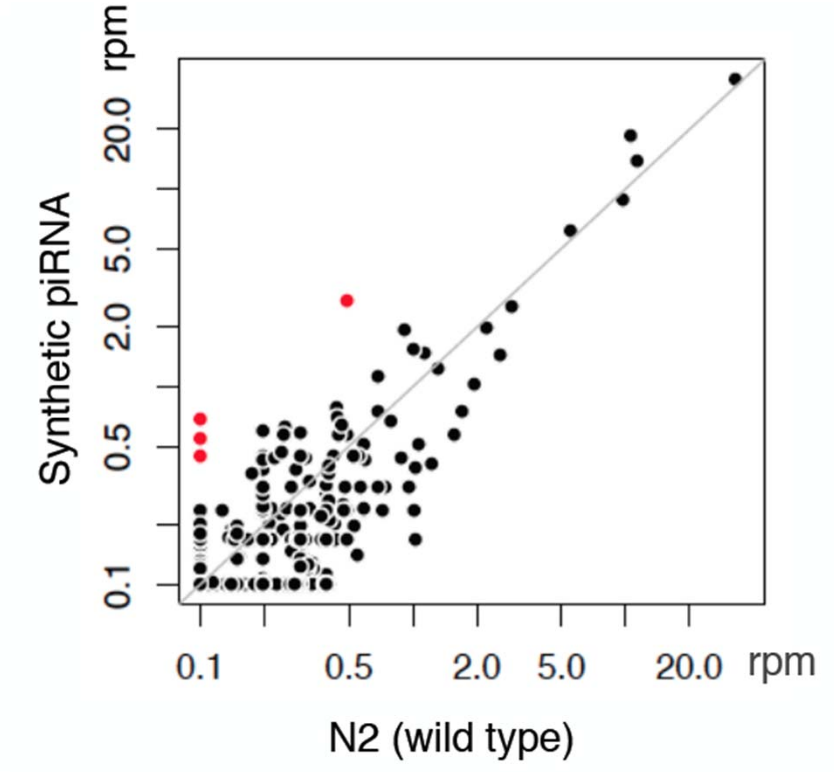
21U-RNA loci*

New piRNA targeting

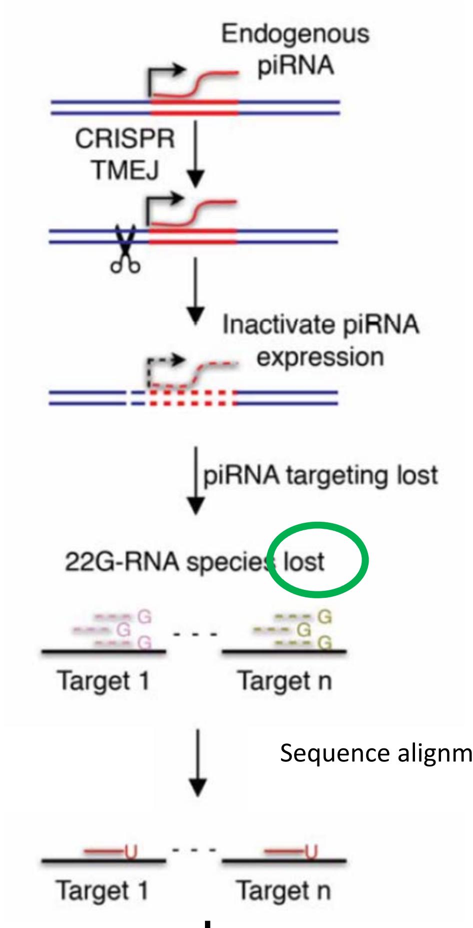
22G-RNA species gained

Sequence alignment between piRNA and targets

piRNA targeting rules

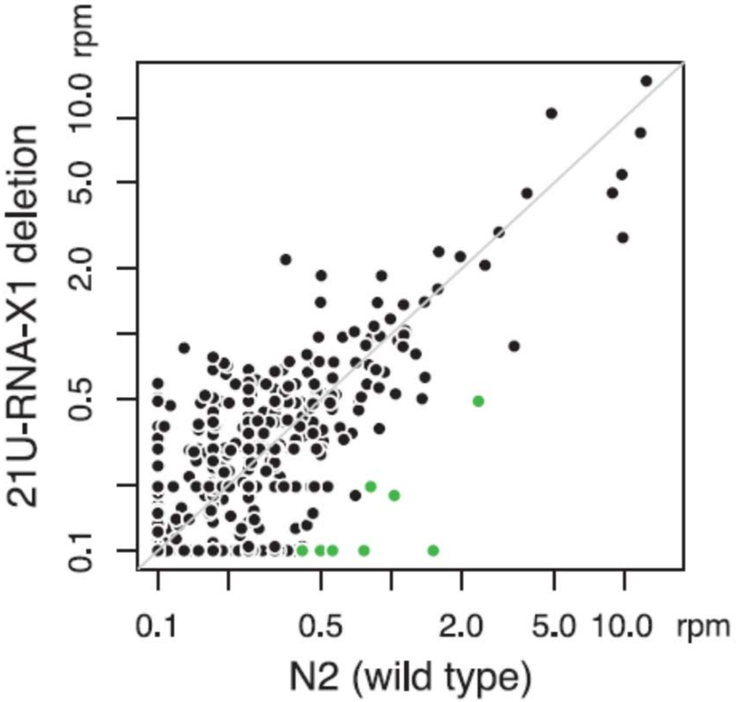


Identify the mRNA targets of a single piRNA

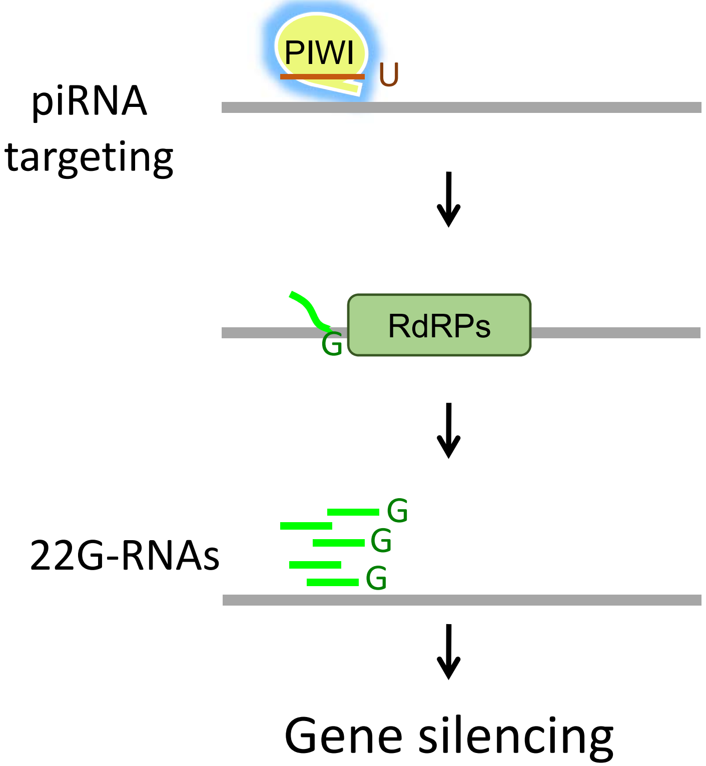


Sequence alignment between piRNA and targets

piRNA targeting rules



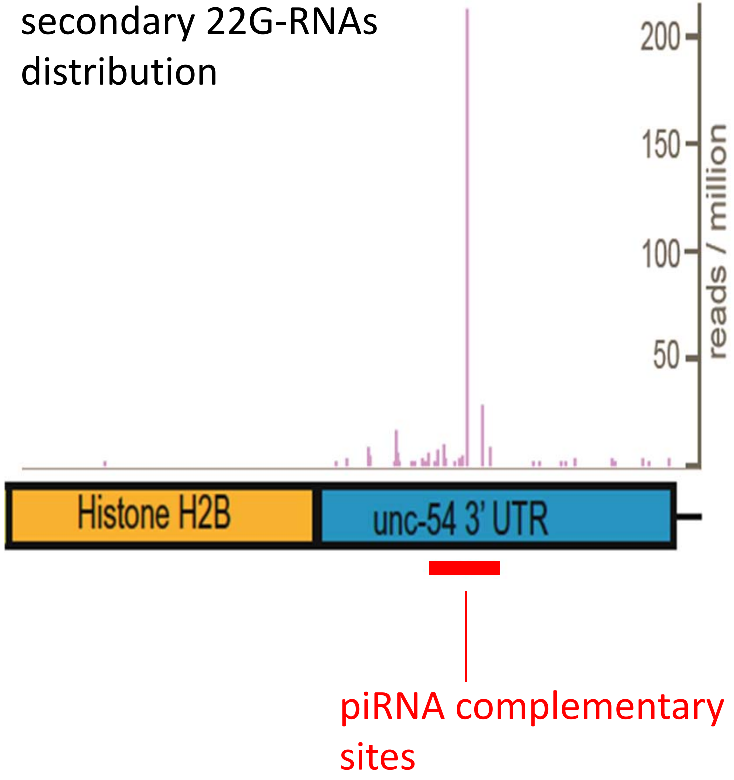
22G-RNAs can serve as a "mark" for piRNA targeting sites *in vivo*



piRNA targeting

22G-RNAs

Gene silencing



secondary 22G-RNAs distribution

piRNA complementary sites

Bagijn MP et al., 2012
Lee HC et al., 2012