Weak seed-pairing stability and high target-site abundance decrease the proficiency of *lsy-6* and other microRNAs

David M. Garcia[1–3,8], Daehyun Baek[1–5,8], Chanseok Shin[1–3,6], George W. Bell[1], Andrew Grimson[1–3,7] & David P. Bartel[1–3]

[1]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, USA. [2]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]School of Biological Sciences, Seoul National University, Seoul, Republic of Korea. [5]Bioinformatics Institute, Seoul National University, Seoul, Republic of Korea. [6]Department of Agricultural Biotechnology, Seoul National University, Seoul, Republic of Korea. [7]Present address: Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York, USA. [8]These authors contributed equally to this work. Correspondence should be addressed to D.P.B. (dbartel@wi.mit.edu) and D.B. (baek@snu.ac.kr).

Supplementary Information:

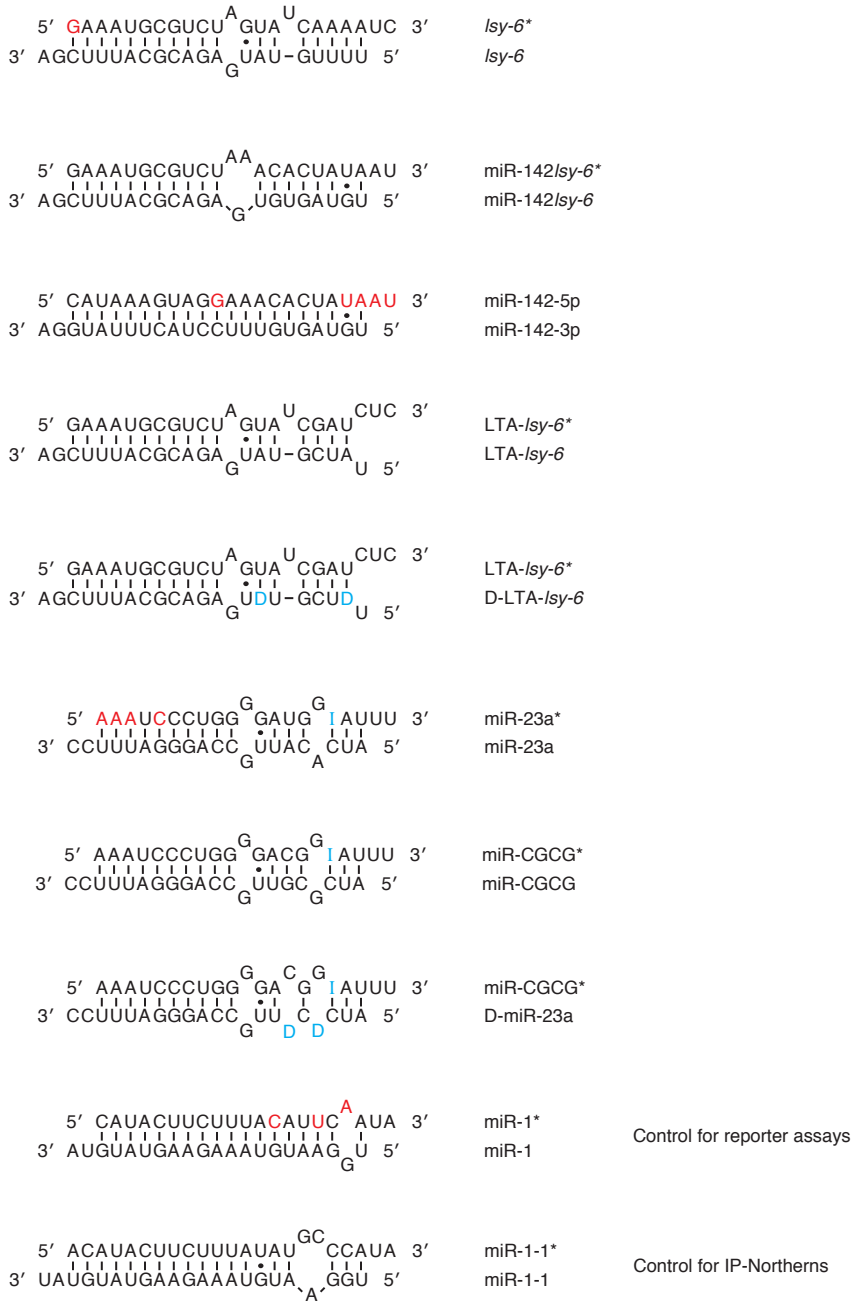
Supplementary Figures 1–5

Supplementary Tables 1–5

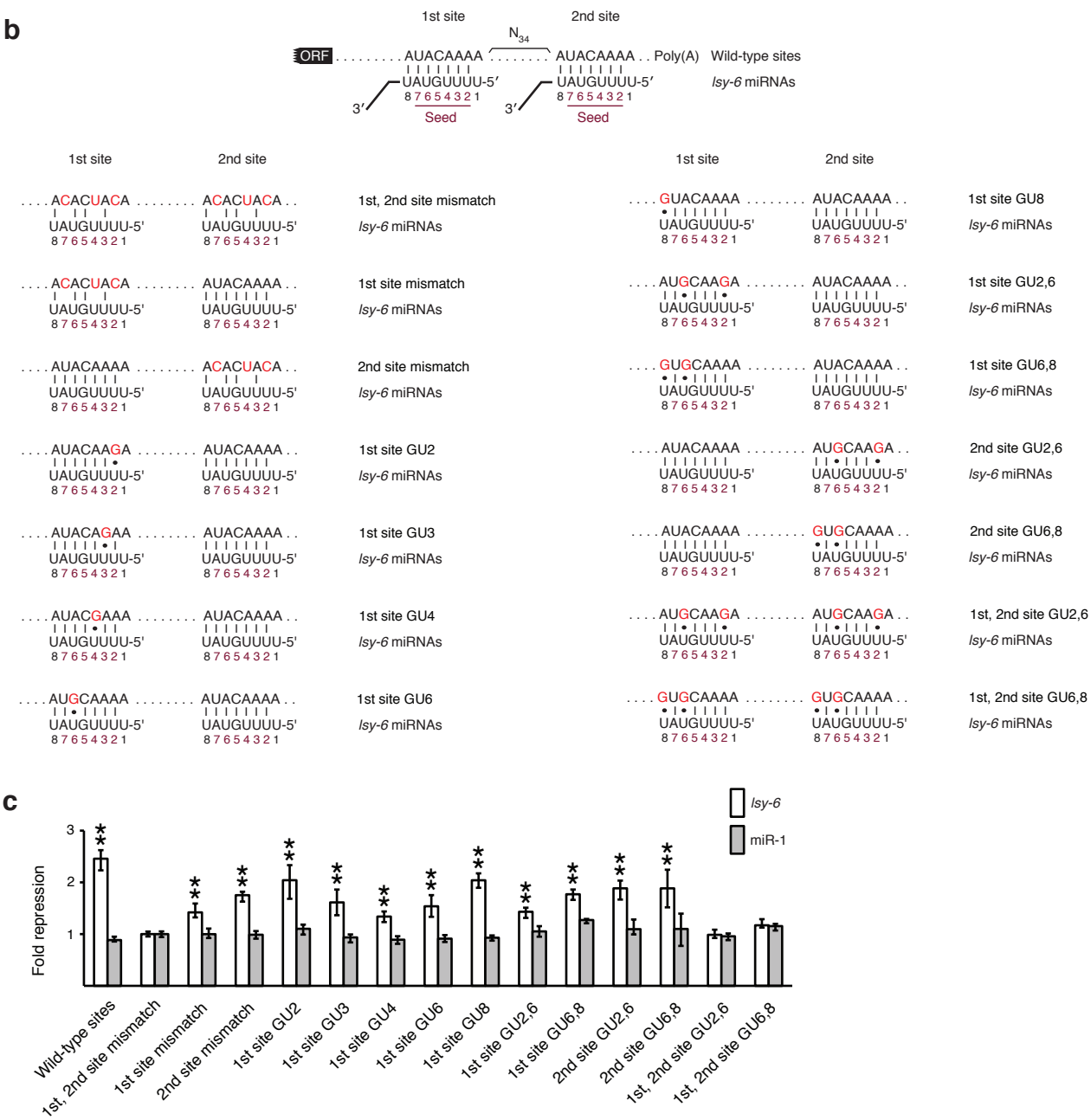
Supplementary References

Nature Structural & Molecular Biology: doi:10.1038/nsmb.2115

a

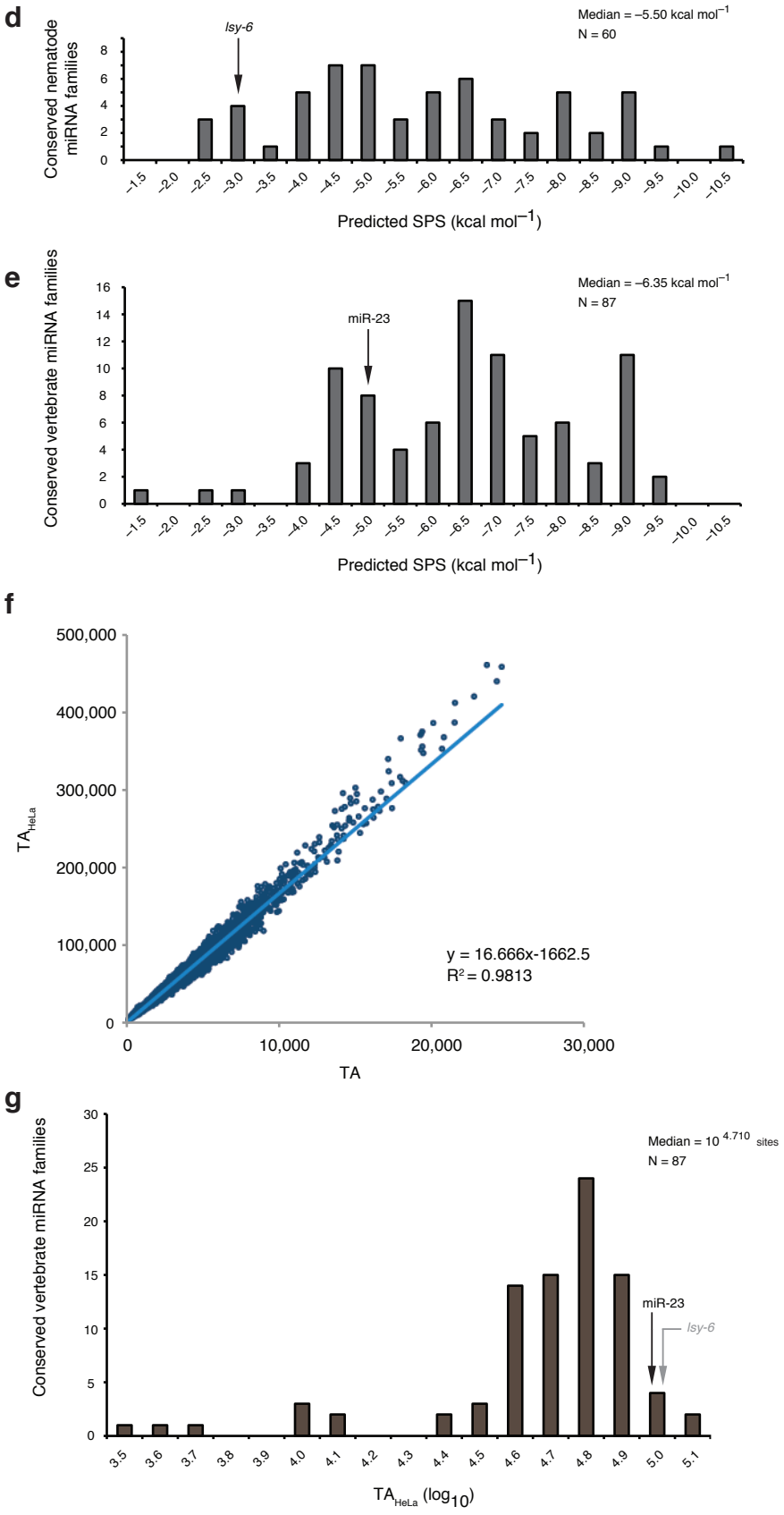


Supplementary Figure 1. Information and analyses related to **Figure 1**. (**a**) Predicted structures for miRNA duplexes transfected in this study. For miRNA mimics of endogenous sequences (*lsy-6*, miR-142-3p, miR-23a, miR-1, miR-1-1), miRNA* nucleotides that differed from their endogenous identities[36,56] are highlighted in red. These changes were designed to facilitate loading of the miRNA. Additionally, a guanine present within endogenous miR-142-5p was deleted (not shown). Non-canonical nucleotides used to either increase SPS (D = 2,6-di-aminopurine), or facilitate loading (I = Inosine), are highlighted in cyan.

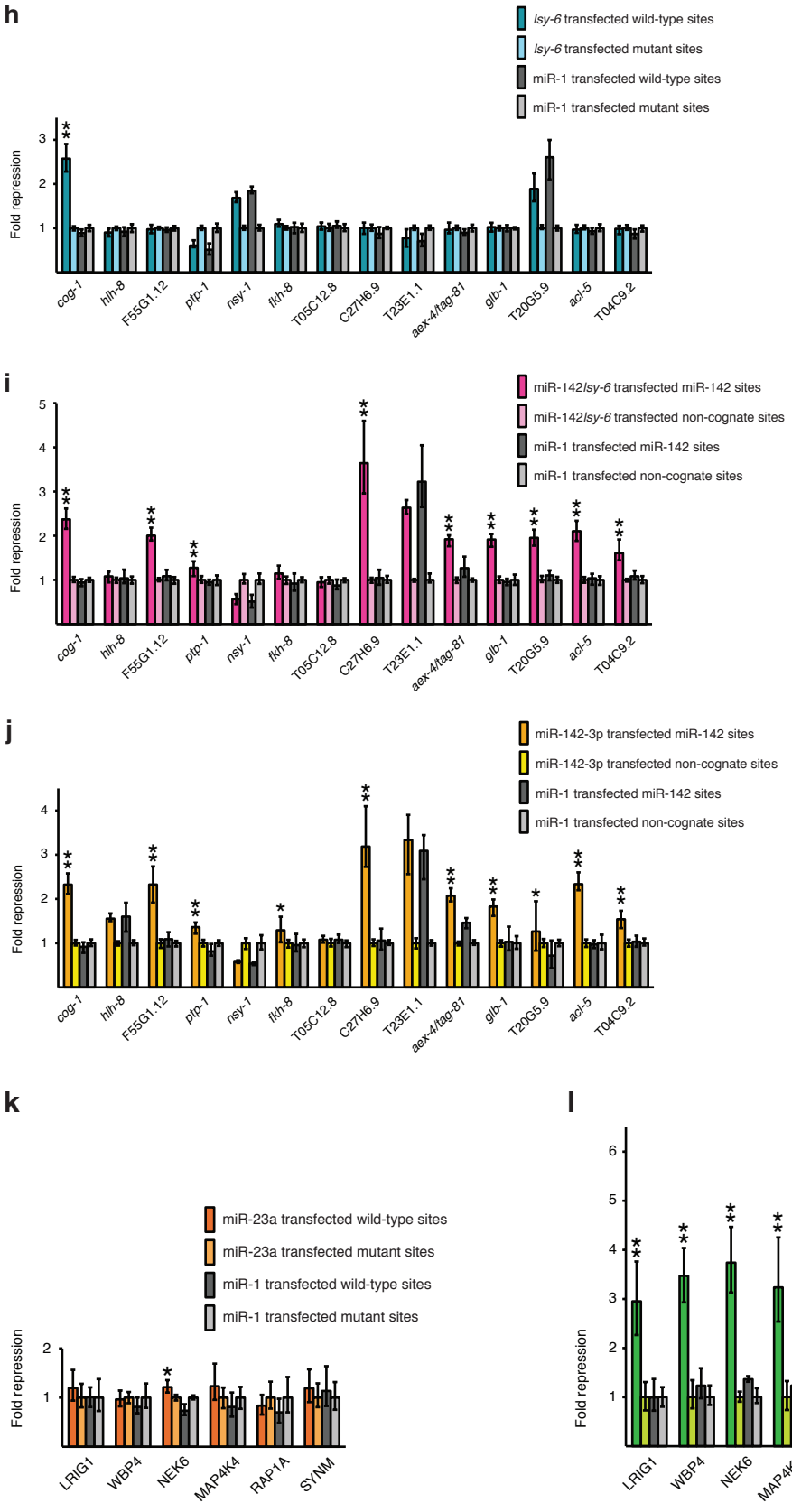


Supplementary Figure 1 continued. Response of the *cog-1* 3′UTR reporter to mutations in the *lsy-6* sites. (**b**) Wild-type and mutant sites containing mismatches or G:U wobbles to the indicated nucleotide(s) of *lsy-6*. Illustrations of mutant sites, with mutated positions shown in red, are simplified from the wild-type sites at top. (**c**) Repression of each construct by *lsy-6* was normalized to a construct with two mutated *lsy-6* sites, each containing two mismatches (1st, 2nd site mismatch). In parallel, activity was measured using a non-cognate miRNA, miR-1 (grey bars). Normalization was as panels **h–l** of this figure. Error bars and statistical significance is as in **Figure 1b**.

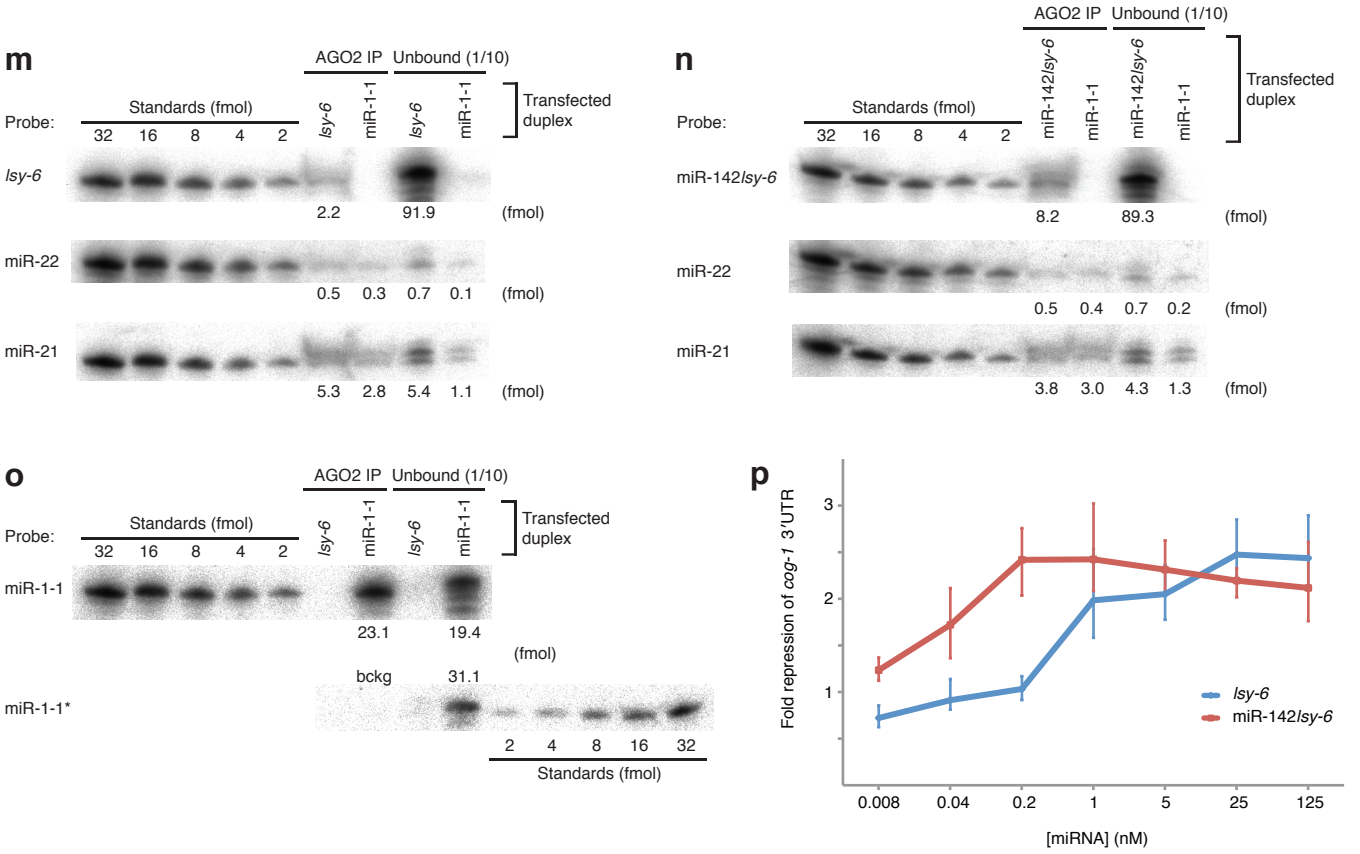
The original study using *in vivo* reporter assays in *C. elegans* concludes that repression of *cog-1* by *lsy-6* is not strongly diminished by the introduction of G:U wobbles into the seed match[23], which contrasts with conclusions from studies using reporters in mammalian cells[5,7] and *D. melanogaster*[5] as well as many other studies using comparative sequence analysis and large-scale experimental datasets[3]. A second study of the *lsy-6*:*cog-1* interaction concludes that some G:U wobble combinations diminish repression of *cog-1* by *lsy-6* in the *in vivo* reporter assay[37]. We used luciferase reporter assays in HeLa cells to examine the same G:U wobble changes as those examined in worms, as well as some additional changes (**Supplementary Table 1**). Introducing G:U wobbles into the upstream *lsy-6* site in *cog-1* was detrimental in all cases. G:U wobbles in the downstream *lsy-6* site also reduced repression, although the effect was less pronounced than for wobbles in the upstream site. Introducing two wobbles into both sites abolished repression.



Supplementary Figure 1 continued. (**d**) Distribution of predicted SPSs for 6mer miRNA sites to 60 conserved nematode miRNA families (**Supplementary Data 2**), as in **Figure 1c**. (**e**) Distribution of predicted SPSs for 6mer miRNA sites to 87 conserved vertebrate miRNA families (**Supplementary Data 2**), as in **Figure 1d**. (**f**) Relationship between human TA and TA_{HeLa} for all heptamers. The least-squares linear fit to the data is shown, with the equation for the line and its Spearman's R^2. (**g**) Distribution of TA_{HeLa}, counting 7mer-m8 3′UTR sites for 87 conserved vertebrate miRNA families, plotted as in **Figure 1f**. TA_{HeLa} values for all 16,384 heptamers are provided in **Supplementary Data 5**.



Supplementary Figure 1 continued. (**h–l**) Reporter results presented in **Figure 1** before normalizing to ratios obtained for the non-cognate miRNA, miR-1. In the main figures, cognate miRNA repression values are normalized to repression values by miR-1. This normalization method was useful because expression differences between the test and control constructs were sometimes observed in the absence of the cognate miRNA (e.g., *nsy-6* or *ptp-1* in **h**).



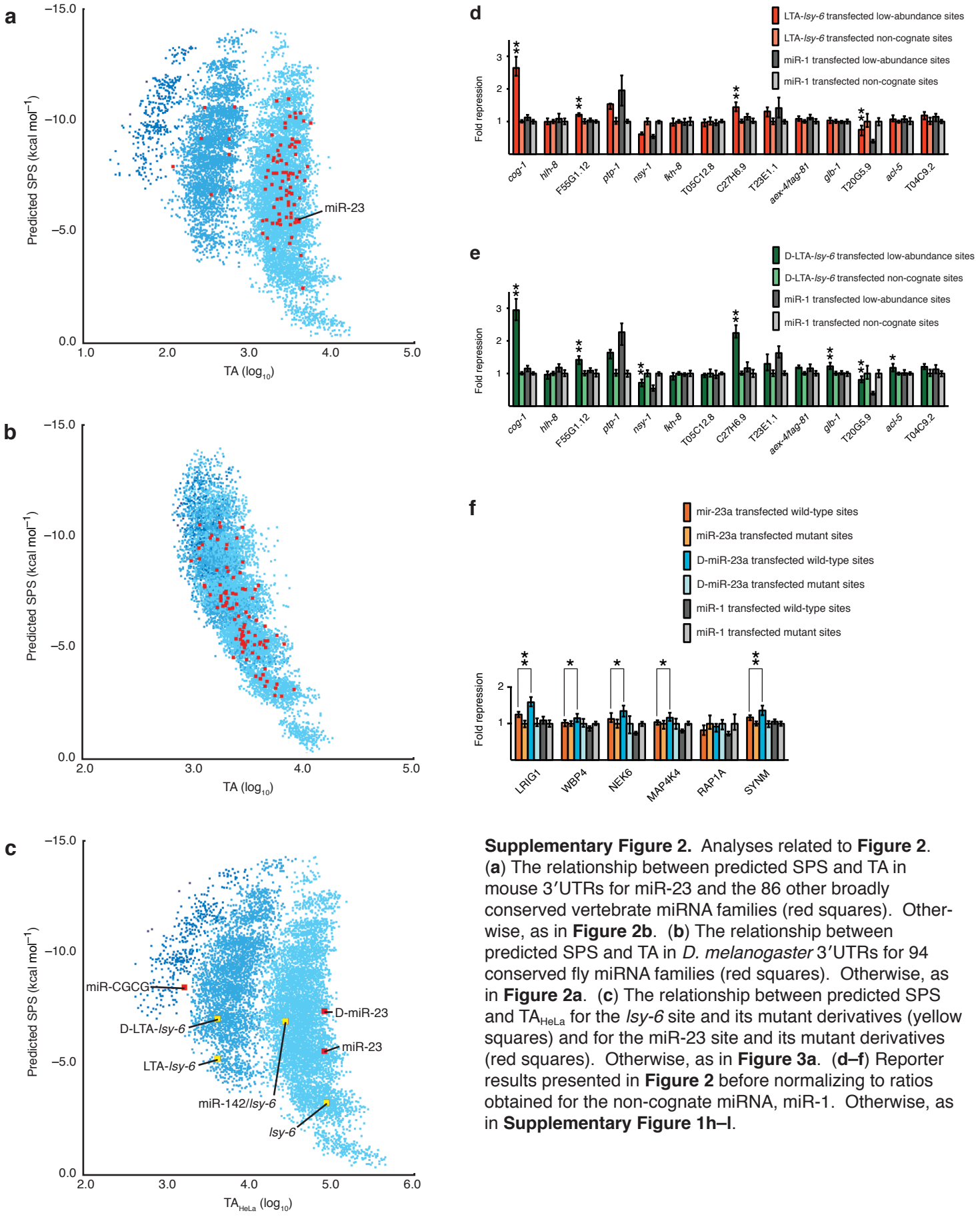
Supplementary Figure 1 continued. Accumulation of transfected miRNAs within the AGO2 silencing complex. (**m**) Quantitative RNA blot probing for *lsy-6* and endogenous controls (miR-21 and miR-22), comparing samples with synthetic RNA standards to samples with material that co-purified with AGO2 (AGO2 IP) and material that did not co-purify (unbound) after transfecting the indicated miRNA duplexes. miR-1-1 samples contained half of the material present in *lsy-6* samples. Because a large fraction of the transfected miRNA did not co-purify with AGO2, only one-

Supplementary Figure 1m–s continued.
tenth of unbound material corresponding to bound material was loaded on the gel. (**n**) Quantitative RNA blot probing for miR-142*lsy-6* chimera and endogenous controls, otherwise as in **m**. (**o**) Control blot probing for miR-1-1 and miR-1-1*, which demonstrated the specificity of the co-purification for loaded miRNA. Otherwise, as in **m**. (**p**) Repression of *cog-1* reporters containing cognate sites for either *lsy-6* or miR-142*lsy-6* chimera measured across a range of transfected miRNA concentrations. Data is plotted as in **Figure 1**, except error bars represent the second largest and second smallest values among 9 replicates from 3 independent experiments. For normalization, a non-cognate miRNA (miR-1) was co-transfected in parallel at the same concentrations as the cognate miRNAs. (**q,r**) Repeat of the experiment in panels **m** and **n**, transfecting less miR-142*lsy-6* chimera to account for its more efficient accumulation in the AGO2 silencing complex. (**s**) Response of reporters to transfection of miR-142*lsy-6* chimera at 0.2 nM, otherwise as in **Figure 1g**.

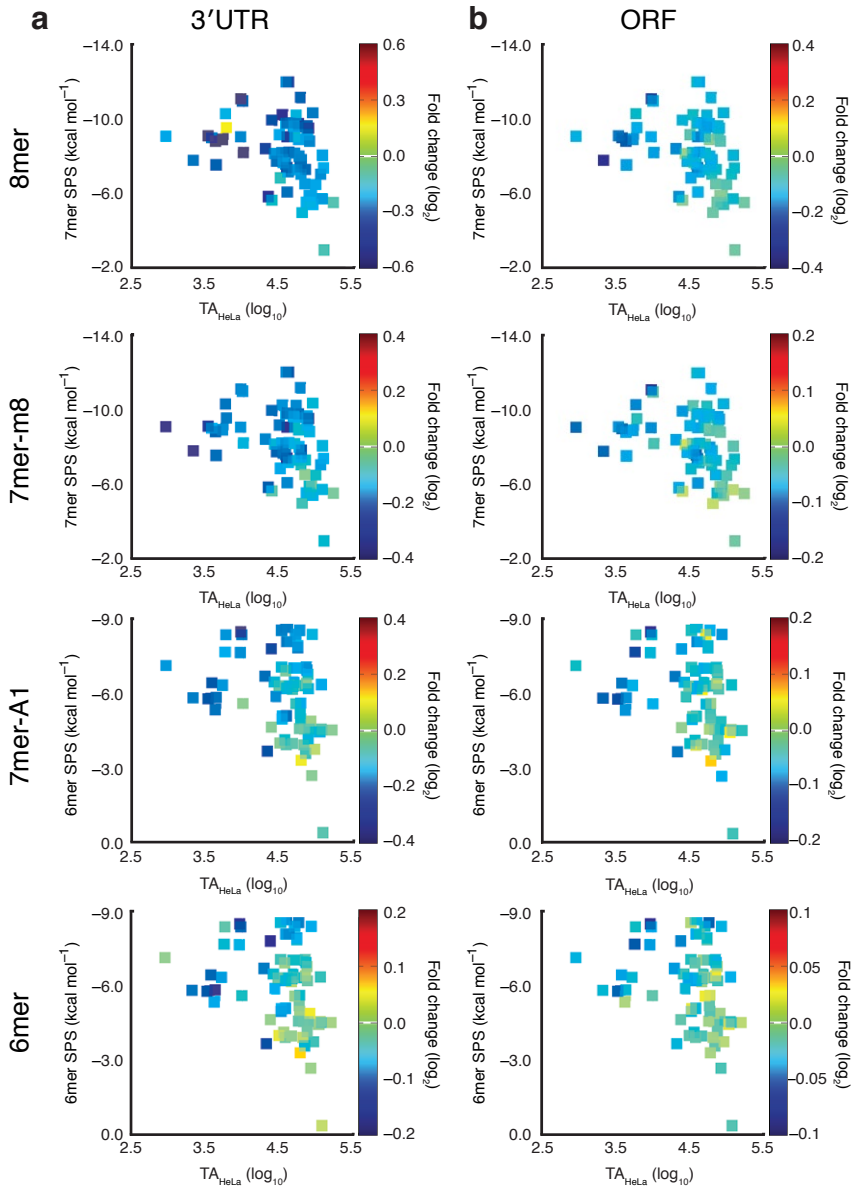
Analyses of the co-purfication results in panels **m** and **n** (geometric mean of ratios normalized to the endogenous internal controls, miR-22 and miR-21) indicated that miR-142*lsy-6* chimera accumulated in AGO2 at a level 4.4-fold higher than did *lsy-6*. This difference represented an estimate of relative accumulation in the silencing complex because levels in AGO1, AGO3, and AGO4 were not determined and because loaded miRNAs might have different degradation rates over the 24 hours after transfection. Because eight targets in **Figure 1** were not significantly repressed by *lsy-6* but were repressed between 1.3- and 3.5-fold by miR-142*lsy-6* chimera, an accumulation difference of less than 5-fold could not explain the difference in proficiency. Consistent with this interpretation were miRNA titration results (**p**), which indicated a rather shallow relationship between miRNA tranfection concentration and fold repression, such that 5-fold differences in miRNA concentration would not be expected to result in the binary differences observed between *lsy-6* and miR-142*lsy-6* chimera, particularly near the concentration used (25 nM).

To find transfection concentrations yielding equal the levels of AGO2-bound *lsy-6* and miR-142*lsy-6* chimera, AGO2 immunopurification was repeated after transfecting miR-142*lsy-6* chimera at concentrations matching those tested in panel **p**. Analyses of these results (panels **q** and **r**) suggested that transfection of miR-142*lsy-6* chimera at 0.2 nM resulted in accumulation of AGO2-bound miRNA to a level similar to that of *lsy-6* transfected at 25 nM. At even lower transfection concentrations, miR-142*lsy-6* chimera levels in AGO2 decreased further, consistent with the reduced repression of *cog-1* at these concentrations (panel **p**). Transfection of miR-142*lsy-6* chimera at 0.2 nM yielded greater reporter repression than that observed in **Figure 1b**, but less than that observed in **Figure 1g** (panel **s**). These results indicate that the relative level of miRNA in the silencing complex (presumed functions of miRNA turnover and loading efficiencies) was not the only factor contributing to proficiency, thereby supporting our conclusion that properties of the seed also played a role. Additional experiments will be needed to learn whether the less efficient accumulation of AGO2-bound *lsy-6* is attributable to poorer loading or faster turnover. If faster turnover of loaded *lsy-6* were a factor, then comparing the results of panel **s** with **Figure 1b** would underestimate the effects of SPS and TA, because the luciferase reporter assay results represented cumulative effects of the miRNA on targets since transfection, and at earlier times the levels of loaded *lsy-6* in **Figure 1b** would have been relatively higher than levels of loaded miR-142*lsy-6* in panel **s**.

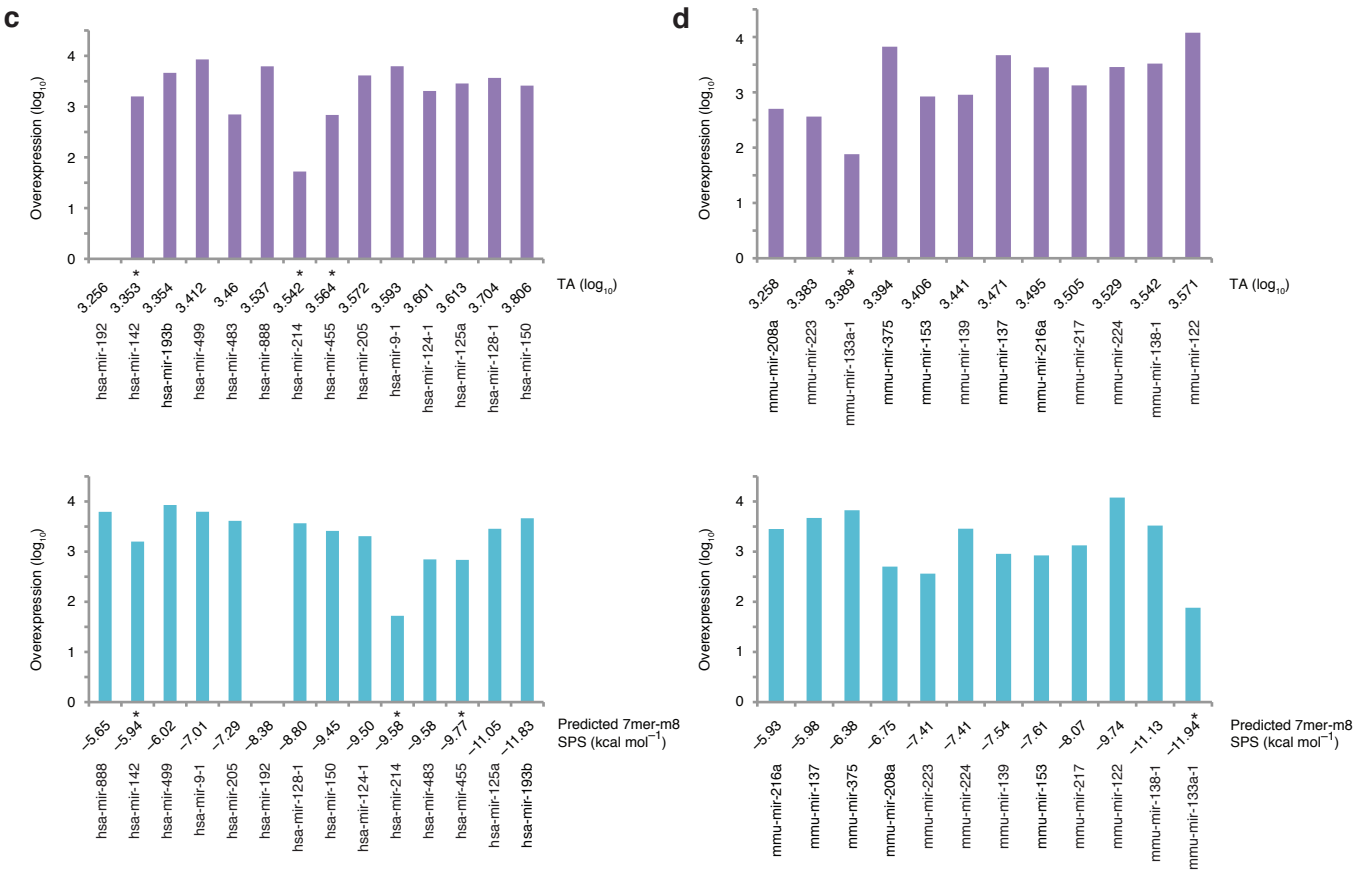
The methods for the immunopurification experiment were as follows: For each miRNA duplex, four (**m–o**) or three (**q,r**) 24-well plates of HeLa cells were transfected as described for the reporter assays (at 25 nM unless otherwise labeled). Half of the wells were co-transfected with pIS0 and pIS1 containing wild-type *lsy-6* sites, the other half with pIS0 and pIS1 containing mutated *lsy-6* sites, and cells were mixed during harvesting. After 24 hours, cells were washed once with 1X PBS and trypsinized, after which all remaining steps were carried out either at 4º C or on ice. Cells were harvested by resuspension in growth media, pelleted (200 x g for 5 minutes), washed with 1X PBS and re-pelleted, then lysed with 4.8 mL or 3.6 mL (50 μL per well) Ago Lysis Buffer (ALB) (25 mM Tris-Cl pH 7.4, 150 mM KCl, 0.5 mM EDTA, 0.5% NP-40, 0.5 mM DTT, one Roche EDTA-free Protease Inhibitor Cocktail tablet per 10 mL) for 1 hour. Cellular debris was spun out (200 x g for 5 minutes), and for each sample, supernatant was mixed with 15 μL of Anti-Human AGO2 antibody (Wako, clone 4G8). After 1 hour, 80 μL EZview Red Protein G Affinity Gel (Sigma) was added, and the mixture was incubated another 4 hours with rocking. Beads were spun down and supernatant ("Unbound") was set aside for later RNA isolation. Beads were washed two times in ALB and then two times in Minimal Cleavage Buffer (MCB) (400 mM KCl, 1mM $MgCl_2$, 10 mM Tris-Cl pH 7.4, 20% w/v Glycerol, 0.5mM DTT). Yeast total RNA was added to IP samples to a concentration of 200 ng per μL, and RNA from IP and Unbound samples was isolated using TRI reagent (Ambion). Small RNA blots were generated and probed as described (http://web.wi.mit.edu/bartel/pub/protocols.html). To enable quantification of RNA levels in the IP and unbound samples, dilution series of synthetic standards for the relevant RNAs were also loaded and used to generate a standard curve–AAGCUGCCAGUUGAAGAACUGU (miR-22); UAGCUUAUCAGACUGAUGUUGA (miR-21); *lsy-6*, miR-142*lsy-6*, miR-1-1, and miR-1-1* sequences are shown in **Supplementary Figure 1a**. Probe sequences: TCGAAATGCGTCTCATACAAAA (*lsy-6*); TCGAAATGCGTCT-CACACTACA (miR-142*lsy-6*); TACATACTTCTTTACATTCCA (miR-1-1); TATGGGCATATAAAGAAGTATGT (miR-1-1*); ACAGTTCTTCAACTGGCAGCTT (miR-22); TCAACATCAGTCTGATAAGCTA (miR-21).



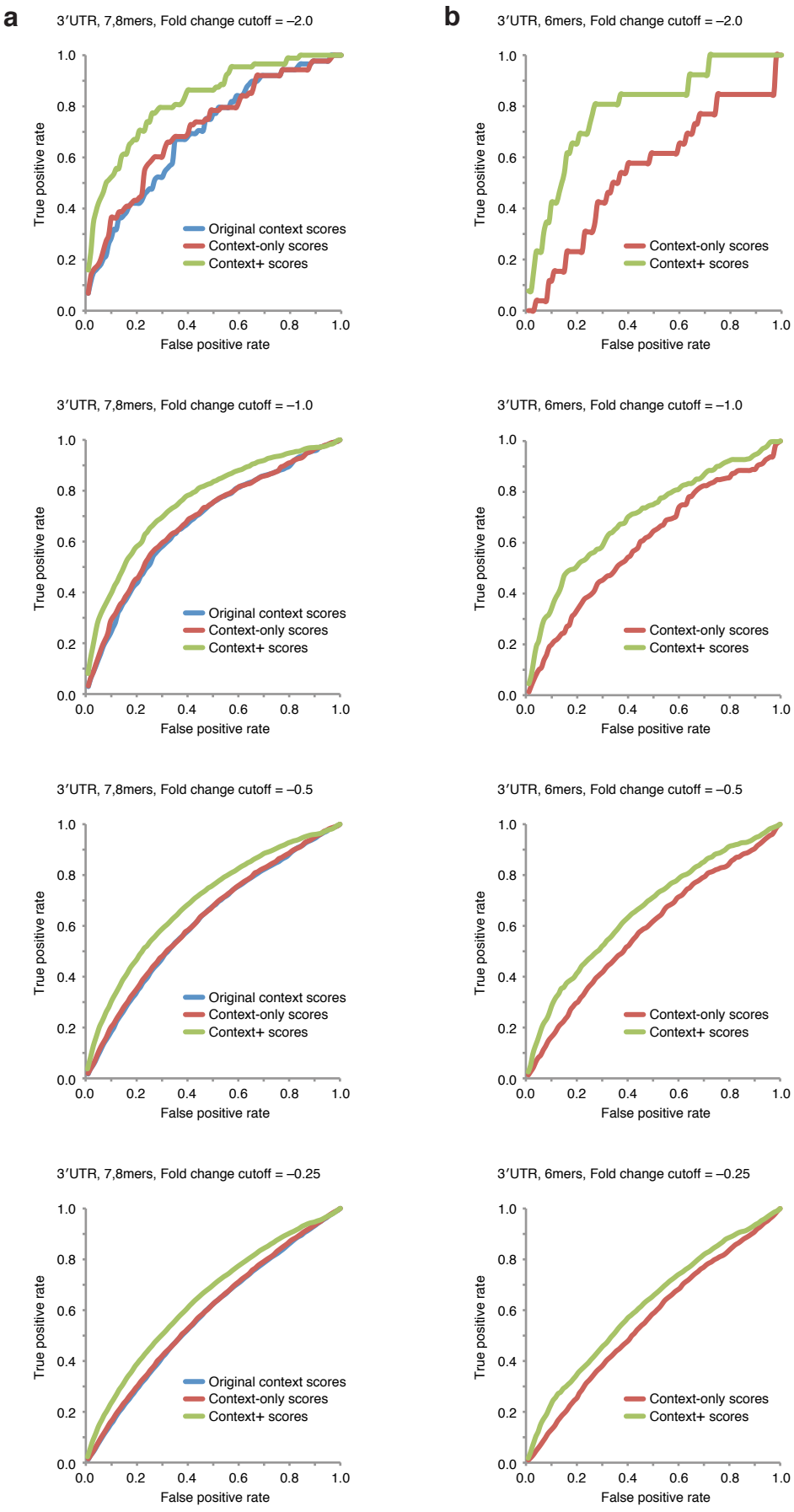
Supplementary Figure 2. Analyses related to **Figure 2**. (**a**) The relationship between predicted SPS and TA in mouse 3′UTRs for miR-23 and the 86 other broadly conserved vertebrate miRNA families (red squares). Otherwise, as in **Figure 2b**. (**b**) The relationship between predicted SPS and TA in *D. melanogaster* 3′UTRs for 94 conserved fly miRNA families (red squares). Otherwise, as in **Figure 2a**. (**c**) The relationship between predicted SPS and TA_{HeLa} for the *lsy-6* site and its mutant derivatives (yellow squares) and for the miR-23 site and its mutant derivatives (red squares). Otherwise, as in **Figure 3a**. (**d–f**) Reporter results presented in **Figure 2** before normalizing to ratios obtained for the non-cognate miRNA, miR-1. Otherwise, as in **Supplementary Figure 1h–l**.



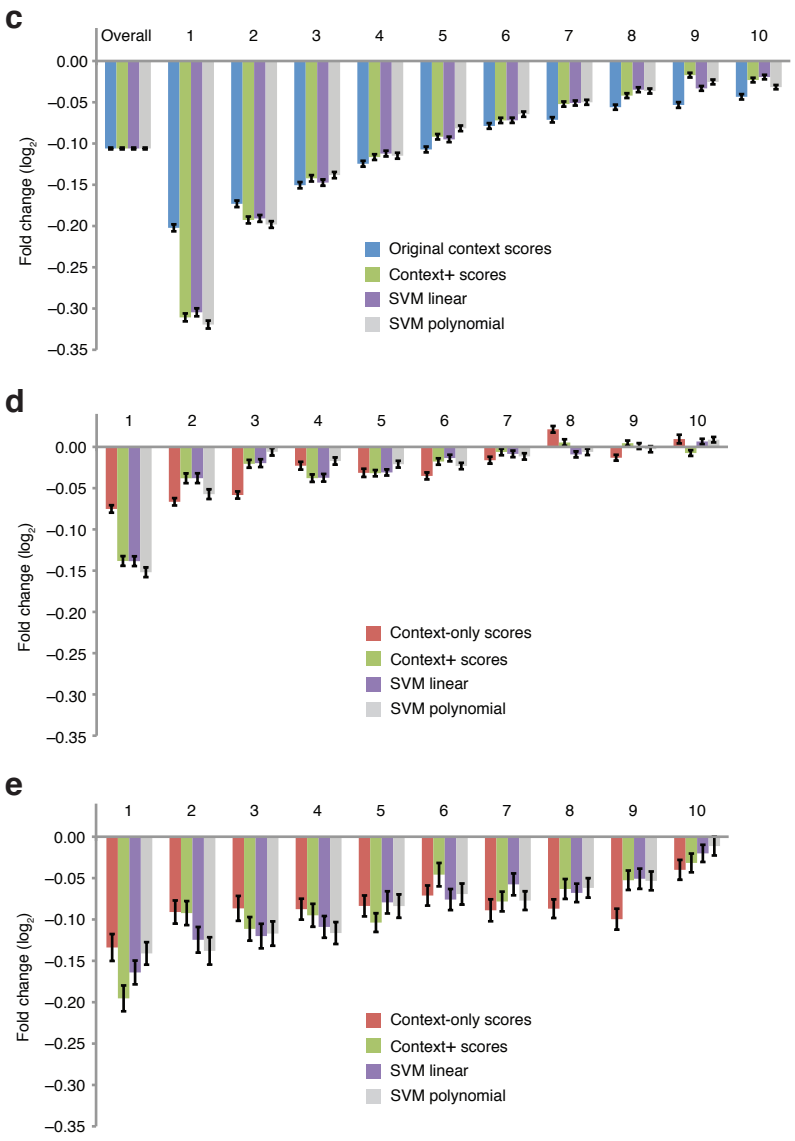
Supplementary Figure 3. Analyses related to **Figure 3**. Impact of TA and SPS on sRNA targeting proficiency of single 3′UTR sites (**a**) and single ORF sites (**b**) to the cognate sRNA, as measured using array data from 74 datasets that passed the motif-enrichment analysis (**Figure 3a**, red squares). Otherwise, as in **Figure 3b,c**.



Supplementary Figure 3 continued. Plots showing the relationship between predicted SPS or TA and the accumulation of mature miRNA after over-expressing the miRNAs from DNA vectors in HEK 293 cells[59]. (**c**) Results from analyses of 14 human miRNAs. Overexpression was calculated as the number of sequencing reads from the most dominant mature miRNA species minus the number of reads found in the mock-transfection control, after normalizing to the reads of endogenous miRNAs that were not overexpressed[59]. For miRNAs marked with asterisks, the most dominant mature miRNA sequence was offset by 1–2 nucleotides with respect to the miRBase annotations, and therefore the predicted SPS and TA values shown differed from those found in **Supplementary Data 2**. These plots show that miRNA accumulation does not decrease with weaker SPS or higher TA. (**d**) Results from analysis of 12 mouse miRNAs. Otherwise, as in **c**.

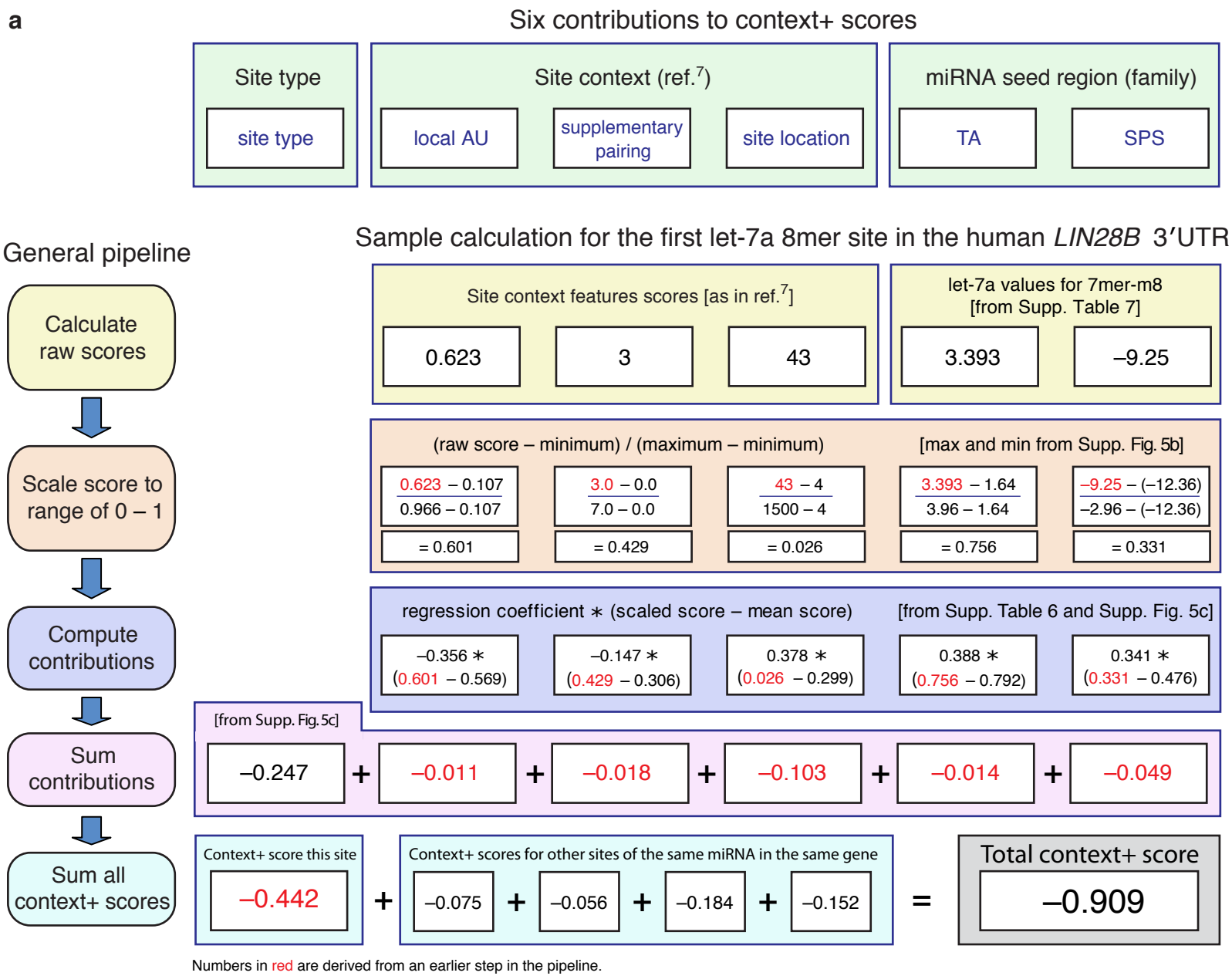


Supplementary Figure 4. Analyses related to **Figure 4**. This page shows ROC curves demonstrating improvements in sRNA target prediction after integrating TA and predicted SPS as features in context+ scores. (**a**) Analyses of mRNAs with 7–8-nucleotide sites in 3′UTRs, performed at four different fold-change cutoffs. (**b**) Analyses of mRNAs with 6mer 3′UTR sites but no larger sites, performed at four different fold-change cutoffs.



Supplementary Figure 4 continued. Performance of the context+ model and SVM regression models with either linear or polynomial kernel. (**c**) Predictions for mRNAs with canonical 7–8-nucleotide 3′UTR sites. Predicted interactions between mRNAs and cognate sRNA were distributed into 10 equally populated bins based on scores generated using the indicated models (key), with the first bin comprising interactions with the most favorable scores. Plotted for each bin is the mean mRNA change on the arrays (error bars, 95% confidence intervals). To perform SVM regression, SVM*light* version 6.02 was used with default parameters[58]. Performance of other SVM kernels (radial basis function and sigmoid tanh) was similar or worse (data not shown). (**d**) Prediction of responsive interactions involving mRNAs with only 3′UTR 6mers sites. Otherwise, as in **c**. (**e**) Prediction of responsive interactions involving mRNAs with at least one 8mer ORF site but no 3′UTR sites. Otherwise, as in **c**.

Supplementary Figure 5. Calculation of context+ scores for TargetScan 6.



Supplementary Figure 5 continued.

b Minimum and maximum values used to scale each parameter.

Site location and type	Local AU content		3′-Supplementary pairing		Site location		TA		SPS	
	Min	Max	Min	Max	Min*	Max	Min	Max	Min	Max
3′UTR 8mer	0.107	0.966	0.0	7.0	4	1500	1.64	3.96	–12.36	–2.96
3′UTR 7mer-m8	0.093	0.990	0.0	7.5	3	1500	1.64	3.96	–12.36	–2.96
3′UTR 7mer-A1	0.122	0.984	0.5	7.5	3	1500	1.64	3.96	–10.00	–0.40
3′UTR 6mer	0.071	0.989	0.0	7.0	3	1500	1.64	3.96	–10.00	–0.40
ORF 8mer	0.033	0.893	0.0	6.5	4	1000	1.64	3.96	–12.36	–2.96
ORF 7mer-m8	0.024	0.914	0.0	7.5	3	1000	1.64	3.96	–12.36	–2.96
ORF 7mer-A1	0.045	0.891	0.0	7.5	3	1000	1.64	3.96	–10.00	–0.40
ORF 6mer	0.024	0.918	0.0	7.5	3	1000	1.64	3.96	–10.00	–0.40

*Although sites within 15 nt of the stop codon were not included as UTR sites because they are in the path of the ribosome as it approaches the stop codon, 3′UTR sites could nonetheless be within 3–4 nucleotides of the polyadenylation site.

c The mean parameters to be used to compute the individual contribution of each determinant in TargetScan6, from analysis of 74 microarrays chosen after motif-enrichment analysis (see main text).

Site location and type	The mean parameter values					
	Fold change	Local AU content	3′-Supplementary pairing	Site location	TA	SPS
3′UTR 8mer	–0.247	0.569	0.306	0.299	0.792	0.476
3′UTR 7mer-m8	–0.120	0.509	0.285	0.289	0.796	0.457
3′UTR 7mer-A1	–0.074	0.555	0.236	0.303	0.794	0.450
3′UTR 6mer	–0.019	0.524	0.306	0.293	0.792	0.437
ORF 8mer	–0.078	0.554	0.334	0.640	0.761	0.438
ORF 7mer-m8	–0.035	0.499	0.288	0.641	0.751	0.422
ORF 7mer-A1	–0.027	0.554	0.290	0.635	0.767	0.415
ORF 6mer	–0.007	0.501	0.289	0.641	0.757	0.404

Supplementary Table 1. Predicted target genes investigated in this study.

Predicted *lsy-6* target genes investigated in this study. Conservation indicates sites present in orthologous UTRs of *C. elegans*, *C. briggsae,* and *C. remanei*. More negative context scores indicate sites predicted to be in more favorable contexts for miRNA recognition[7]. A new tool that precisely maps the 3′ ends of transcripts was applied to *C. elegans*[36], enabling us to check the 3′UTR annotations of these targets. These data indicated that for some of the predicted targets the UTRs end before reaching the *lsy-6* sites. However, this information did not change our conclusions regarding the targeting proficiency of the *lsy-6* miRNA because many of the predicted sites not retained in the worm UTRs must have been retained in those UTRs in HeLa cells — otherwise, repression would not have been observed in **Figure 1g,h**.

Target gene	Site	Sequence name	*C. elegans* site type	Conserved	Context score
cog-1	1	R03C1.3(A)	8mer	Yes	–0.43
cog-1	2	R03C1.3(A)	8mer	Yes	–0.46
hlh-8		C02B8.4	7mer-m8	Yes	–0.26
F55G1.12		F55G1.12	8mer	As a 7mer-A1	–0.51
ptp-1	1	C48D5.2A	7mer-A1	No	–0.14
ptp-1	2	C48D5.2A	8mer	No	–0.52
nsy-1		F59A6.1	7mer-A1	No	–0.12
fkh-8		F40H3.4	7mer-A1	Yes	–0.21
T05C12.8		T05C12.8	7mer-m8	Yes	–0.30
C27H6.9*	1	C27H6.9	8mer	As a 7mer-A1	–0.50
C27H6.9*	2	C27H6.9	8mer	No	–0.43
T23E1.1		T23E1.1	7mer-m8	No	–0.19
aex-4/tag-81		T14G12.2	7mer-m8	Yes	–0.27
glb-1		ZK637.13	7mer-A1	Yes	–0.15
T20G5.9		T20G5.9	7mer-A1	Yes	–0.15
acl-5		R07E3.5	7mer-A1	Yes	+0.01
T04C9.2		T04C9.2	7mer-m8	Yes	–0.23

*Listed as C27H6.3 in ref 23.

Predicted miR-23 target genes investigated in this study. Conservation status and site context scores (calculated for miR-23a) from TargetScan 5.1[7]. More negative scores indicate sites predicted to be in more favorable contexts for miRNA recognition[7].

Target gene	Site	Human site type	Conserved	Context score
LRIG1	1	7mer-A1	Yes	–0.19
LRIG1	2	8mer	Yes	–0.29
WBP4	1	7mer-A1	Yes	–0.28
WBP4	2	7mer-A1	Yes	–0.26
NEK6	1	8mer	Yes	–0.36
NEK6	2	8mer	Yes	–0.43
MAP4K4	1	7mer-m8	Yes	–0.20
MAP4K4	2	8mer	Yes	–0.44
RAP1A	1	7mer-A1	Yes	–0.17
RAP1A	2	7mer-A1	Yes	–0.18
DMN	1	7mer-A1	No	–0.11
DMN	2	8mer	No	–0.32

Supplementary Table 1 continued. Sequences of UTR fragments assayed. Listed are the plasmid name in brackets, gene name, RefSeq accession number (where applicable, longest isoform is cited), and UTR sequence tested with miRNA seed sites underlined. For *lsy-6* and miR-23 targets, the full-length sequence shown is wild-type. For mutant constructs, only the miRNA site (underlined, mutations in uppercase) is shown; the remainder of the UTR sequence was identical to wild-type. For all *cog-1* UTRs assayed in **Supplementary Figure 1b,c**, full-length sequences are shown.

lsy-6 target UTRs:

[pDMG1a] *cog-1*; NM_001027093
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtactttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaaaaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaaatttcccatttaattgtacgttttttttctcttcaaattgg
attctaatgacataaatttattagattaa
[pDMG1b] miR-142 sites: aCacTaCa, aCacTaCa; otherwise as for pDMG1a
[pDMG1c] LTA sites: atacGaTa, atacGaTa; otherwise as for pDMG1a

[pDMG2a] *hlh-8*; NM_076966
tttcgaatatggaaaaaactggccagctcctaatttatttgataatgtatgcttctcaatcaacatagtcctcatgatatagtgccttattctcatttttgatgtatcaaatctgtctaa
taatatctactcgatttcaattttcgtttgctcaaaacttaaaaatttgcttgagaattttgcaagagactattagtaagcccattgttaaatgaaacaagtttatcccgcaaacaga
aactatgtgtgaatatgatcaaactataatacaaacacgtaaaaaaaaaattttgaatcataattatcatttgaccactaagccatgcaatgatgaaccaatttcaacttgacattc
acacccagtagtggtatcaattgactcttttacccagtcatcgccattctgtctcatcacatcgatcgtcctcattattggcttgcattctccgaatcctaaaaaaagtgtgggtc
agcggcgtgatggatgggcgtctatgaaaaaaaacgagcccatcggagcccaaatg
[pDMG2b] miR-142 site: aCacTaC; otherwise as for pDMG2a
[pDMG2c] LTA site: atacGaT; otherwise as for pDMG2a

[pDMG3a] F55G1.12; NM_068805
attgattttatttattttaatttaatgaatctcgccggaatgtctgatttgttgcttggtttggtttgaaaattatatacaaaaatagtggattaaatgat
[pDMG3b] miR-142 site: aCacTaCa; otherwise as for pDMG3a
[pDMG3c] LTA site: atacGaTa; otherwise as for pDMG3a

[pDMG4a] *ptp-1*; NM_065331.2
gcttttatccaaaaaatacatatatcgtttttgtttcttcaaattcttcttcccatcgaactctcatgaatcacggatcccgcgaggtgctagctatttttgccttttttctttcttcttttttt
ttattgcatagttaattagctattgttttcctacacaaaactagtcaatgttttaagtaattaaatatcatcatttaatatttcaacaaaaaatctatctcaatgggtcacccgatgtgat
tttcgtaccaattgttccccatcactacgtcataattgtaccacccccccccatctttcatgtacgaaaaatcgcccaaacttgtatgtaaaaaaaacaaaaaagtcctctcaaa
tcatcacaaactttccttcttttttcatataaaatgttacagtctgtgtttccattgtacaaaaaaaaagtgtgatcggggaaaaagaaacggggctacatgatcgggaaagtgtg
aacagtttgcttgatttcggaaatcaccaggtttcaaaatttctaaataaaattggaagggaagggaagggaagggaaaagagaaatatatataaatatatagaacccaaga
aaaaatggaacaaaaaaaacggagaatgaggtgtttagatgaacgaaaaaatgcaaatttttagagtttngtcgatccagcgaataaaatcgaaacttngaaaaaaaaaa
agagacactgcctattagaaaacaaaaaaaaacatttcacaaaaaaaattagatgggggtgagggacgaagagtagatcagaaaattgtgaaaaagaaaatatttttacat
cggttttccataacaaaacggtacataaaatgatggagagaatcgaggggaatgatcaaggaatgggacatggcctgccatgagacaaaagacgnttcncaatacaca
antagagcgggggatagaagantagaagaagcttattccagtgnttatggggctatttatatgatgtagaaaaatacaaaaatgtatttttatacatnttccccc
[pDMG4b] miR-142 sites: CacTaCa, aCacTaCa; otherwise as for pDMG4a
[pDMG4c] LTA sites: tacGaTa, atacGaTa; otherwise as for pDMG4a

[pDMG5a] *nsy-1*; NM_062524.6
ttttctgtagtttctctgttctctctctctctctctctctggtcatttttctctcttctagtttttctgtctctttctctcatttttattgtgatatctttttctctctctctctctctctctctctaat
cctctgtttcgtgtacaaagttttcagtttcttaatttgttctctgaaaatcatcaaaccccctccaaaatttgcttgcgtgtagaacttttcattacaaacaaaaaccaaattgctagt
gtctttcctatccaactacattatagagatttctaatctcatgtcaattgtttcatgtatcattctacaaaactcacacacatccgaatcatcataactatcataataataagttttatta
accaaaaataaataaatatatatttcatgatcgttattctccgagcacttccgcatcatcttcggcaacccataagtaaatcatttgtcctgctgcagtcatgagtcgttctttttttc
ggatgagccgtcagagaatgaatcacttttgaaggatgttcgagtttcgacacgatgtgtgaatcgaggagactgtaatttattctttctttgaaactaatattgttacctaccta
tacacataaacgaatccgtcttctgagccacttgccacgtgctcgatggattgtaaaactcggcagtctaatttgtattccgtgttttgatgacctttatagctgaaaatttataaa
tttatagctccaattttcatacaatattcaaaaatatgtaccttgctagcaattttccactagatttatcaataagccgaacaattccacccattactccagccaacaggcaattact
gtctggggtaaagctgacactgtttactgaatctcccatgtagtcgactgtcatctganaattagaaatttctaataacttgaaaattctttttttttttcaattacatttccatctctaa
tactatagactctgtaatttccatcagcacttccagcaacaatctcatggccatta
[pDMG5b] miR-142 site: CacTaCa; otherwise as for pDMG5a
[pDMG5c] LTA site: tacGaTa; otherwise as for pDMG5a

[pDMG6a] *fkh-8*; NM_062834.2
attgttaatctaaaggttcaaaaactcacatatttttcacacagtgtccaaatttcatttgtacaaaaatacattgttagttagttttcattttcatattcatttttcgtaaacattcaa
[pDMG6b] miR-142 site: CacTaCa; otherwise as for pDMG6a
[pDMG6c] LTA site: tacGaTa; otherwise as for pDMG6a

[pDMG7a] T05C12.8; NM_063322.2
tacaactaaatgtggcaaagcttcttcattgtttgaataatgaaaaacgaatacaaacaactttgaaatcaaaacattaaaacttacaacatttcgttcaataaagtatcatcaaa
agaagaaaacaaaagctgaacatgagaatttgggataaggagcagcagatcggaattatgtgagaagcacgcggaaaacagggatatataaacggggtaaacgggaaaat
[pDMG7b] miR-142 site: aCacTaC; otherwise as for pDMG7a
[pDMG7c] LTA site: atacGaT; otherwise as for pDMG7a

[pDMG8a] C27H6.9; NM_001129395.1
ttggaaaatgtgatgtttttctataaataaatattctcacaactctttttcatgttttatataatacaaaatgcacatcaagcagaaaaatttcaacataaagtttacaccagaagtga
atttagggatgaagaggaaccaaattacgtaaatacaaaagtatcgaaacatgatagat
[pDMG8b] miR-142 sites: aCacTaCa, aCacTaCa; otherwise as for pDMG8a
[pDMG8c] LTA sites: atacGaTa, atacGaTa; otherwise as for pDMG8a

[pDMG9a] T23E1.1; NM_067895.3
aattgagatcaaattgttcttttatatgtatgtactgaaaacaataaagaatttttgaaattaaaaatttaaagtcttcactcacacccgcctgggaacccccctcttctagccctga
aaacgccttaaattgcacacggagcaagtaaggagtggatgccttgtaggcttaggctcggacttaggcctaggctcaggattaggtttaggcttaggcttagactgggc
gggggaagagagcaaaaataagttccagaaaattcaagaattaaaaaaaggaaataagcctcctaattaggcgaggaggctggcgagaggcgagttttcaatccataat
atccgtgttaagctatttttttttaataaactcttcgaaaatatctactttccctgcaccagtttttctcttccaaaatgttccaaatatgtattgttgagtggcgtaagcaaaacaaag
tcaagtctctagtgaatacaaacacacgctcttcatttttt
[pDMG9b] miR-142 site: aCacTaC; otherwise as for pDMG9a
[pDMG9c] LTA site: atacGaT; otherwise as for pDMG9a

[pDMG10a] *aex-4/tag-81*; NM_076240.5
cttcacaaaaagtgtggtgcgcgcattccacgggctacgaacacatgggtaaactgtacattttcaaatattgttgaaaacttttaatttttcaattttaaattcaaactttatgtttt
aagcaataaaatgatgatttaatccgttatacaaagacatggaaaagttacagttagttttttttttaagcggtcgttatttatagggggttcgtttaatggtgtcacatactgctttg
cgt
[pDMG10b] miR-142 site: aCacTaC; otherwise as for pDMG10a
[pDMG10c] LTA site: atacGaT; otherwise as for pDMG10a

[pDMG11a] *glb-1*; NM_066573.5
ttgagcctttatattgtatttgaatgagctttgagtattataatgattatctctcttggaaacgtttttgtacaaaataaacaaag
[pDMG11b] miR-142 site: CacTaCa; otherwise as for pDMG11a
[pDMG11c] LTA site: tacGaTa; otherwise as for pDMG11a

[pDMG12a] T20G5.9; NM_066860.2
gcaacgattaaatatagattctacctctctgtttcatttcatgtgcgatagtttcagataattatttatttttatatttttgtattttatgaacgggttcgatacttgtctttttttcggttggaat
gtacaaaaatacacagaatacacgaattga
[pDMG12b] miR-142 site: CacTaCa; otherwise as for pDMG12a
[pDMG12c] LTA site: tacGaTa; otherwise as for pDMG12a

[pDMG13a] *acl-5*; NM_001047817.1
agtttttgatgtacaaaactagccaatttttttgtatcagatcttttattgattgtttacgtttgaacggttccatttgccaaa
[pDMG13b] miR-142 site: CacTaCa; otherwise as for pDMG13a
[pDMG13c] LTA site: tacGaTa; otherwise as for pDMG13a

[pDMG14a] T04C9.2; NM_065904.1
cgcgataacttttgtttcggctcctatacaaattttggttatttttgttggtcgctccaacatttttttcgtcctcatccgagccatgacttctcttctccttttcctctatttcgtctcaaact
tccgttctttttttacctaatcatcattattagccccatccttatcatcttctggaacccacatcgtcatcttcggtttctttctttttgaggcacaagcaacaactacttttctcgcatctt
ctctcctccagcttctcttctttatgagccgggttaggggctcttcgaaaattgtttccactcggctgccttcgtgttttcgacgtgcccgaacttgctcaaaaccgaagctcac

gcatcgttaggtaacgaaagaattacacgtaggagggacgcactgccgtttgattcttatctgtcatcgtcaggattgttgcaacctcttcaatcctccggatgtgcgttgatt
cccgcacgattagacaatttgttgtg
[pDMG14b] miR-142 site: aCacTaC; otherwise as for pDMG14a
[pDMG14c] LTA site: atacGaT; otherwise as for pDMG14a

miR-23 target UTRs

[pAG247] LRIG1; NM_015541.2
gataaaagcaaatgtggccttctcagtatcattcgattgctatttgagacttttaaattaaggtaaaggctgctggtgttggtacctgtggatttttctatactgatgttttcgttttg
ccaatataatgagtattacattggccttgggggacagaaaggaggaagttctgacttttcagggctaccttatttctactaaggacccagagcaggcctgtccatgccattcc
ttcgcacagatgaaactgagctgggactggaaaggacagcccttgacctgggttctgggtataatttgcacttttgagactggtagctaaccatcttatgagtgccaatgtgt
catttagtaaaacttaaatagaaacaaggtccttcaaatgttcctttggccaaaagctgaagggagttactgagaaaatagttaacaattactgtcaggtgtcatcactgttcaa
aaggtaagcacatttagaattttgttcttgacagttaactgactaatcttacttccacaaaatatgtgaatttgctgcttctgagaggcaatgtgaaagagggagtattacttttat
gtacaaagttatttatttatagaaattttggtacagtgtacattgaaaaccatgtaaaatattgaag
[pCS247] miR-CGCG sites: aCgCgaa, aaCgCgaa; otherwise as for pAG247

[pAG249] WBP4; NM_007187.3
catgcttttaggacagaatggagacttatacacccaaagtttatctgtgtttgtttgtaagtattatgatgctaaaaatttagatttattctaaatgtatttgatgtgaattaaaataaa
tatttttttcatgtgaaatttatttttggttcctaaaatggaagcctaccacattgcattgtaatacagtgtattatgttcagtgtctaaaaactgctaattaagtcataatttaagatgcta
tgtatctgttatttaaaacatggagaaacagggccttttattccattcatattcataagagcatatttatcctgcattgaaaatgcattacttttgcacattgatattaactgttgtccaa
caaataagtatcggagtacgtgagaatattccc
[pCS249] miR-CGCG sites: aCgCgaa, aCgCgaa; otherwise as for pAG249

[pAG250] NEK6; NM_001145001.2
aacagctaagaccacagggttcagcaggttccccaaaaggctgcccagccttacagcagatgctgaaggcagagcagctgagggaggggcgctggccacatgtcac
tgatggtcagattccaaagtcctttctttatactgttgtggacaatctcagctgggtcaataagggcaggtggttcagcgagccacggcagccccctgtatctggattgtaat
gtgaatctttagggtaattcctccagtgacctgtcaaggcttatgctaacaggagacttgcaggagaccgtgtgatttgtgtagtgagcctttgaaaatggttagtaccgggtt
cagtttagttcttggtatcttttcaatcaagctgtgtgcttaatttactctgttgtaaagggataaagtggaaatcatttttttccgtggagtggtgattctgctaacatttttatctacg
ttttataacttggtgagtgacgatgagagccctgcacctggccagagtgtcacaggcaaaaggcatcgggaagcaggagcatcttcttggcagccaggctgggccatctt
ctcctggacacctgctgtgtaccaggaacttcgtcacctccttgaatgctggcggttcatttcatgatcagtgttaagcattttcctccatgggaaggaagcatgggatatagа
aaagcgaagggctgtcctttacaaattctggttctgcaacttcctagcgtgactttgggcttgggcaagtttcttagccgttctgagccttcatttcctcatctgtacaatgagatt
aatagtacctatcatctaccttcaggattgctgacagacagaatttgaaataaaatatgcaagttagctaatacaaaaagtagatgatccaaaaatggtagccactcacccttc
acaaactgaagtccatggaccacggaagtcgagaattaatgtacacctgtatcatgtgtaggaaaccagaaatgtgttccttatttcttgttcccaaacaggattaactgtgaa
gactaatttataaatgtgaacctaagaaaactccacctctgaaggaaatcatttgaattttgtttttgtacgtaaagttaaccttccaattgtctgagctgtcgtcactgacttcatg
acagtctggccctccagacaagagcagcgctggcatcgggcaggtgattcctgacacct
[pCS250] miR-CGCG sites: aaCgCgaa, aaCgCgaa; otherwise as for pAG250

[pAG252] MAP4K4; NM_145686.2
tttgggattgagcatcatactggaaagcaaacacctttcctccagctccagaattccttgtctctgaatgactctgtcttgtgggtgtctgacagtggcgacgatgaacatgcc
gttggttttattggcagtgggcacaaggaggtgagaagtggtggtaaaaggagcggagtgctgaagcagagagcagatttaatatagtaacattaacagtgtatttaattg
acatttctttttgtaatgtgacgatatgtggacaaagaagaagatgcaggtttaagaagttaatatttataaaatgtgaaagacacagttactaggataacttttttgtgggtgg
ggcttgggagatggggtggggtgggttaaggggtcccattttgtttctttggatttggggtgggggtcctggccaagaactcagtcatttttctgtgtaccaggttgcctaaat
ca
[pCS252] miR-CGCG sites: aaCgCga, aaCgCgaa; otherwise as for pAG252

[pAG253] RAP1A; NM_001010935.1
gccagattacaggaatgaagaactgttgcctaattggaaagtgccagcattccagacttcaaaaataaaaaatctgaagaggcttctcctgttttatatattatgtgaagaatttt
agatcttatattggtttgcacaagttccctggagaaaaaaattgctctgtgtatatctcttggaaaataagacaatagtatttctcctttgcaatagcagttataacagatgtgaaa
atatacttgactctaatatgattatacaaaagagcatggatgcatttcaaatgttagatattgctactataatcaaatgatttcatattgatctttttatcatgatcctccctatcaagc
actaaaaagttgaaccattatactttatatctgtaatgatactgattatgaaatgtcccctgaa
[pCS253] miR-CGCG sites: aCgCgaa, aCgCgaa; otherwise as for pAG253

[pAG260] SYNM; NM_145728.2
cagacagagatgtgctgattttgttttagctgtaacaggtaatggtttttggatagatgattgactggtgagaatttggtcaaggtgacagcctcctgtctgatgacaggacag
actggtggtgaggagtctaagtgggctcagtttgatgtcagtgtctgggctcatgacttgtaaatggaagctgatgtgaacaggtaattaatattatgacccacttctatttactt

tgggaaatatcttggatcttaattatcatctgcaagtttcaagaagtattctgccaaaagtatttacaagtatggactcatgagctattgttggttgctaaatgtgaatcacgcgg
gagtgagtgtgcccttcacactgtgacattgtgacattgtgacaagctccatgtcctttaaaatcagtcactctgcacacaagagaaatcaacttcgtggttggatggggccg
gaacacaaccagtctttttgtatttattgttactgagacaaaacagtactcactgagtgtttttcagtttcctactggtggttttga
[pCS260] miR-CGCG sites: aCgCgaa, aaCgCgaa; otherwise as for pAG260

cog-1 UTR sequences assayed in **Supplementary Figure 1b,c**.

[pDMG1a] wild-type sites
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaaaaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaatttcccatttaattgtacgttttttttctcttcaaattgg
attctaatgacataaatttattagattaa

[pDMG1b] 1st, 2nd site mismatch
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttaCacTaCaaccaaactccctttttaccgttaaaccatgcccaaaCacTaCaaatttcccatttaattgtacgttttttttctcttcaaaat
tggattctaatgacataaatttattagattaa

[pDMG1d] 1st site mismatch
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttaCacTaCaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaatttcccatttaattgtacgttttttttctcttcaaattg
gattctaatgacataaatttattagattaa

[pDMG1e] 2nd site mismatch
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaaaaaccaaactccctttttaccgttaaaccatgcccaaaCacTaCaaatttcccatttaattgtacgttttttttctcttcaaattg
gattctaatgacataaatttattagattaa

[pDMG1f] 1st site GU2
cttttaagcgtctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttccttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaaGaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaatttcccatttaattgtacgttttttttctcttcaaattg
gattctaatgacataaatttattagattaa

[pDMG1g] 1st site GU3
cttttaagcgtctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaGaaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaatttcccatttaattgtacgttttttttctcttcaaattgg
attctaatgacataaatttattagattaa

[pDMG1h] 1st site GU4
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacGaaaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaatttcccatttaattgtacgttttttttctcttcaaattgg
attctaatgacataaatttattagattaa

[pDMG1i] 1st site GU6
cttttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatGcaaaaaccaaactccctttttaccgttaaaccatgcccaaatacaaaaaatttcccatttaattgtacgttttttttctcttcaaattgg
attctaatgacataaatttattagattaa

[pDMG1j] 1st site GU8
ctttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttGtacaaaaaccaaactccctttaccgttaaaccatgcccaaatacaaaaaaatttcccatttaattgtacgttttttctcttcaaattgg
attctaatgacataaatttattagattaa

[pDMG1k] 1st site GU2,6
ctttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatGcaaGaaccaaactccctttaccgttaaaccatgcccaaatacaaaaaaatttcccatttaattgtacgttttttctcttcaaattg
gattctaatgacataaatttattagattaa

[pDMG1l] 1st site GU6,8
ctttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttGtGcaaaaaccaaactccctttaccgttaaaccatgcccaaatacaaaaaaatttcccatttaattgtacgttttttctcttcaaattgg
attctaatgacataaatttattagattaa

[pDMG1m] 2nd site GU2,6
ctttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaaaaaccaaactccctttaccgttaaaccatgcccaaatGcaaGaaatttcccatttaattgtacgttttttctcttcaaattg
gattctaatgacataaatttattagattaa

[pDMG1n] 2nd site GU6,8
ctttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatacaaaaaccaaactccctttaccgttaaaccatgcccaaGtGcaaaaaatttcccatttaattgtacgttttttctcttcaaattg
gattctaatgacataaatttattagattaa

[pDMG1o] 1st, 2nd site GU2,6
ctttaagcgttctacctctccccctcccttcaaccgagtgtattattcccccaatttgtttgcaatttttttcctgaagccctctaagaaaatccaaaatcatgacctacttccgtcttt
acacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatttt
ttttctttttttttccaaatcatcgtcacttatGcaaGaaccaaactccctttaccgttaaaccatgcccaaatGcaaGaaatttcccatttaattgtacgttttttctcttcaaatt
ggattctaatgacataaatttattagattaa

[pDMG1p] 1st, 2nd site GU6,8
ctttaagcgttctacctctccccctcccttcaaccgagnngtattattcccccaatttgtttgcaatttttttcctgaagccctctaanaaaatccaaaatcatgacctacttccgtc
tttacacctgattacctgaataccaacaccccacacagatgccatgatctctcgtcttttctcgtacttttgtataatttttttttcttaatttttttgcatgttttcccatagttatagccatt
ttttttttctttttttttccaaatcatcgtcacttGtGcaaaaaccaaactccctttaccgttaaaccatgcccaaGtGcaaaaaatttcccatttaattgtacgttttttctcttcaaat
tggattctaatgacataaatttattagattaa

Supplementary Table 2. Relationship between mean mRNA repression and either TA or predicted SPS for the indicated site types, from analysis of microarrays chosen after motif-enrichment analysis.

Site location and type	Multiple linear regression			Simple linear regression			
	Multiple R^2	*P* value		TA_{HeLa}		SPS	
		TA_{HeLa}	SPS	R^2	*P* value	R^2	*P* value
3′UTR 8mer	0.189	0.032	0.012	0.113	0.0034	0.134	0.0013
3′UTR 7mer-m8	0.320	9.3×10^{-5}	0.013	0.258	3.8×10^{-6}	0.156	5.0×10^{-4}
3′UTR 7mer-A1	0.442	4.6×10^{-5}	2.2×10^{-5}	0.280	1.3×10^{-6}	0.294	6.0×10^{-7}
3′UTR 6mer	0.345	2.3×10^{-4}	0.0013	0.241	8.9×10^{-6}	0.206	4.8×10^{-5}
ORF 8mer	0.350	2.7×10^{-6}	0.087	0.323	1.3×10^{-7}	0.112	0.0036
ORF 7mer-m8	0.306	7.4×10^{-5}	0.032	0.259	3.7×10^{-6}	0.132	0.0014
ORF 7mer-A1	0.298	1.8×10^{-5}	0.14	0.276	1.5×10^{-6}	0.089	0.0099
ORF 6mer	0.287	0.0031	0.0017	0.179	1.7×10^{-4}	0.193	9.1×10^{-5}
5′UTR 8mer	0.006	0.52	0.81	0.006	0.54	0.000	0.97
5′UTR 7mer-m8	0.000	0.91	0.97	0.000	0.91	0.000	0.99
5′UTR 7mer-A1	0.022	0.42	0.49	0.016	0.29	0.013	0.33
5′UTR 6mer	0.016	0.33	0.47	0.009	0.42	0.003	0.65

Supplementary Table 3. Multiple linear regression statistics for miRNA target prediction for context+ scores, using 11 microarray datasets previously used to build the TargetScan context score model.

Site location and type	Multiple linear regression intercept and coefficients (*P* value)					
	Intercept	Local AU content	3′-Supplementary pairing	Site location	TA_{HeLa}	SPS
3′UTR 8mer	–0.674 (0.003)	–0.447 (2×10^{-7})	–0.006 (1)	0.312 (1×10^{-7})	0.431 (0.1)	0.416 (1×10^{-5})
3′UTR 7mer-m8	–0.309 (0.02)	–0.443 (2×10^{-22})	–0.186 (0.01)	0.213 (4×10^{-13})	0.300 (0.06)	0.310 (2×10^{-12})
3′UTR 7mer-A1	–0.596 (1×10^{-7})	–0.226 (3×10^{-8})	–0.111 (0.07)	0.119 (3×10^{-6})	0.681 (6×10^{-7})	0.163 (0.002)
3′UTR 6mer	–0.350 (7×10^{-10})	–0.164 (5×10^{-16})	–0.023 (0.4)	0.084 (2×10^{-12})	0.431 (2×10^{-10})	0.106 (7×10^{-6})
ORF 8mer	–0.317 (0.02)	–0.191 (2×10^{-4})	–0.048 (0.5)	0.117 (2×10^{-7})	0.289 (0.07)	0.134 (0.007)
ORF 7mer-m8	–0.110 (0.2)	–0.139 (1×10^{-5})	–0.042 (0.4)	0.052 (9×10^{-5})	0.149 (0.1)	0.019 (0.5)
ORF 7mer-A1	–0.077 (0.2)	–0.077 (0.01)	–0.050 (0.2)	0.052 (3×10^{-5})	0.089 (0.3)	0.042 (0.2)
ORF 6mer	–0.104 (0.01)	–0.059 (0.002)	–0.016 (0.6)	0.025 (8×10^{-4})	0.144 (.004)	0.007 (0.7)

Supplementary Table 4. Multiple linear regression statistics for miRNA target prediction for context-only scores, using 11 microarrays previously used to build the TargetScan context score.

Site location and type	Multiple linear regression intercept and coefficients (*P* value)			
	Intercept	Local AU content	3′-Supplementary pairing	Site location
3′UTR 8mer	–0.150 (0.03)	–0.376 (1×10^{-5})	–0.076 (0.6)	0.290 (1×10^{-6})
3′UTR 7mer-m8	0.061 (0.08)	–0.395 (7×10^{-18})	–0.230 (0.002)	0.198 (3×10^{-11})
3′UTR 7mer-A1	0.019 (0.5)	–0.188 (3×10^{-6})	–0.163 (0.008)	0.100 (9×10^{-5})
3′UTR 6mer	0.045 (0.002)	–0.143 (9×10^{-13})	–0.043 (0.1)	0.074 (8×10^{-10})
ORF 8mer	–0.022 (0.6)	–0.189 (2×10^{-4})	–0.057 (0.4)	0.113 (6×10^{-7})
ORF 7mer-m8	0.019 (0.4)	–0.138 (1×10^{-5})	–0.043 (0.3)	0.052 (8×10^{-5})
ORF 7mer-A1	0.009 (0.7)	–0.073 (0.01)	–0.054 (0.2)	0.051 (3×10^{-5})
ORF 6mer	0.017 (0.4)	–0.054 (0.005)	–0.014 (0.6)	0.025 (8×10^{-4})

Supplementary Table 5. Context+ parameters to be used for improved target predictions in TargetScan 6. Analysis is with 74 filtered representative array datasets (**Supplementary Data 1**).

Site location and type	Multiple linear regression intercept and coefficients (*P* value)					
	Intercept	Local AU content	3′-Supplementary pairing	Site location	TA	SPS
3′UTR 8mer	–0.583 (7×10^{-25})	–0.356 (1×10^{-16})	–0.147 (0.03)	0.378 (2×10^{-45})	0.388 (1×10^{-10})	0.341 (6×10^{-17})
3′UTR 7mer-m8	–0.243 (6×10^{-23})	–0.366 (1×10^{-74})	–0.139 (2×10^{-5})	0.212 (4×10^{-63})	0.243 (4×10^{-20})	0.207 (3×10^{-28})
3′UTR 7mer-A1	–0.298 (2×10^{-28})	–0.187 (1×10^{-17})	–0.048 (0.1)	0.164 (6×10^{-39})	0.239 (5×10^{-16})	0.220 (2×10^{-26})
3′UTR 6mer	–0.114 (1×10^{-19})	–0.084 (7×10^{-15})	–0.048 (0.002)	0.094 (3×10^{-51})	0.106 (7×10^{-15})	0.098 (1×10^{-22})
ORF 8mer	–0.260 (1×10^{-18})	–0.147 (5×10^{-8})	–0.035 (0.3)	0.122 (1×10^{-24})	0.203 (2×10^{-11})	0.095 (1×10^{-4})
ORF 7mer-m8	–0.095 (1×10^{-11})	–0.074 (6×10^{-7})	–0.033 (0.1)	0.056 (5×10^{-19})	0.071 (2×10^{-7})	0.043 (8×10^{-4})
ORF 7mer-A1	–0.164 (2×10^{-21})	–0.014 (0.4)	–0.041 (0.07)	0.063 (2×10^{-21})	0.130 (1×10^{-14})	0.040 (0.007)
ORF 6mer	–0.054 (3×10^{-10})	0.004 (0.7)	–0.035 (0.005)	0.028 (2×10^{-14})	0.037 (7×10^{-6})	0.023 (0.004)

Supplementary Data 1, 2, 4 and 5 are provided separately as excel files.

Supplementary Data 3 is provided separately as an .fa file.

Supplementary References:

56. Landgraf, P. et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401-14 (2007).
57. Doench, J.G. & Sharp, P.A. Specificity of microRNA target selection in translational repression. *Genes Dev* **18**, 504-11 (2004).
58. Joachims, T. Optimizing Search Engines Using Clickthrough Data. *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)* (2002).
59. Chiang, H.R. et al. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* **24**, 992-1009 (2010).