Functional relevance of "seed" and "non-seed" sequences in microRNA-mediated promotion of *C. elegans* developmental progression

HUIBIN ZHANG, KAREN L. ARTILES, and ANDREW Z. FIRE

Stanford University School of Medicine, Stanford, California 94305, USA

ABSTRACT

The founding heterochronic microRNAs, *lin-4* and *let-7*, together with their validated targets and well-characterized phenotypes in *C. elegans*, offer an opportunity to test functionality of microRNAs in a developmental context. In this study, we defined sequence requirements at the microRNA level for these two microRNAs, evaluating *lin-4* and *let-7* mutant microRNAs for their ability to support temporal development under conditions where the wild-type *lin-4* and *let-7* gene products are absent. For *lin-4*, we found a strong requirement for seed sequences, with function drastically affected by several central mutations in the seed sequence, while rescue was retained by a set of mutations peripheral to the seed. *let-7* rescuing activity was retained to a surprising degree by a variety of central seed mutations, while several non-seed mutant effects support potential noncanonical contributions to *let-7* function. Taken together, this work illustrates both the functional partnership between seed and non-seed sequences in mediating *C. elegans* temporal development and a diversity among microRNA effectors in the contributions of seed and non-seed regions to activity.

Keywords: *C. elegans*; *let-7*; *lin-4*; microRNA; seed

INTRODUCTION

Since their first discovery in *C. elegans* almost two decades ago, microRNAs have emerged as an important class of small regulatory RNA molecules involved in many processes in diverse organisms (Bartel 2004; Bushati and Cohen 2007). Despite the hundreds of new microRNAs identified and extensive characterization (Lagos-Quintana et al. 2001; Lau et al. 2001; Lee and Ambros 2001), there are few well-validated downstream microRNA targets and most microRNAs' functions remain unclear. Comprehensive studies on microRNA mutants have revealed that the absence of a single microRNA rarely results in a severe developmental phenotype, hampering efforts to uncover microRNA functions (Miska et al. 2007; Alvarez-Saavedra and Horvitz 2010). MicroRNAs downregulate gene expression by binding to sites with partial complementarity in target messenger RNAs (mRNA) (Bartel 2004; Bushati and Cohen 2007), with the lack of perfect complementarity in the interactions between microRNA and target mRNA further compounding the target mRNA identification problem.

Comparative sequence analyses provided the first insights into the fundamentals of microRNA target recognition. Comparison of known microRNAs and their targets across orthologous genes in related genomes revealed the high conservation of target 3′ end, suggesting the significance of the corresponding microRNA seed sequence (Lewis et al. 2003; Stark et al. 2003). Development of target prediction tools based on conservation of microRNA seed::target pairing and experimental validation of predicted targets further support the importance of the conserved 5′ region of the microRNA in target identification (Lewis et al. 2003; Stark et al. 2003). Different types of preferentially conserved target sites and their efficacies were also defined through various studies using both computational and experimental methods, refining our knowledge on microRNA specificity (Brennecke et al. 2005; Lewis et al. 2005; Grimson et al. 2007; Friedman et al. 2009; Shin et al. 2010; Jan et al. 2011). Close analyses of conserved sequences flanking microRNA seed::target complementary sites revealed features of site context, such as a strong preference for an adenosine across from the first microRNA nucleotide (Lewis et al. 2005; Grimson et al. 2007).

These initial observations and further work have led to the usage of the sequence at the 5′ end of microRNAs as one of several key criteria in searches of the transcriptome for potential mRNA targets, often placing high weight on target matches to the "seed" sequence between second and eighth nucleotide of the microRNA (Lewis et al. 2003; Stark et al. 2003; Sethupathy et al. 2006). Due to the short (7 nt) length of the seed sequence, simple alignment with the transcriptome would yield thousands of candidates; hence, other factors such as conservation and thermodynamics are also taken into consideration (Sethupathy et al. 2006). Even so, such methods predict up to a few hundred target genes for each microRNA family (Lewis et al. 2005; Rajewsky 2006). Experimental validation of these predictions including mutational studies, mRNA/protein analyses following pull down of silencing complexes and assessing functional consequences of microRNA perturbation lend support to the importance of the seed sequence in assessing candidate microRNA targets (Lewis et al. 2003; Doench and Sharp 2004; Kiriakidou et al. 2004; Kloosterman et al. 2004; Krutzfeldt et al. 2005; Beitzinger et al. 2007; Baek et al. 2008; Selbach et al. 2008).

Despite various lines of experimental evidence highlighting the importance of the seed sequence, it has been shown that imperfect seed matches can also confer functionality (e.g., Ha et al. 1996; Vella et al. 2004a; Stern-Ginossar et al. 2007). Besides the canonical seed match target sites, target sites with extensive base-pairing to the 3′ end of microRNA or centered base-pairing can compensate for weak 5′ base-pairing to mediate gene regulation (Yekta et al. 2004; Brennecke et al. 2005; Grimson et al. 2007). However, such sites make up a small fraction of microRNA target sites (Friedman et al. 2009). Mutational studies on non-seed pairing regions show that other regions are also important in target downregulation (e.g., Vella et al. 2004b; Didiano and Hobert 2006, 2008). In fact, only one out of 14 predicted seed match targets is downregulated by endogenous *lsy-6* when used in *lsy-6* sensors, demonstrating that seed matches are poor predictors of function in this case (Didiano and Hobert 2006). This may be explained by weak seed-pairing ability of *lsy-6* (Garcia et al. 2011). Recent work has shown that weak seed-pairing stability and high target-site abundance can account for *lsy-6*'s low proficiency (Garcia et al. 2011). Although immunoprecipitation studies of proteins associated with silencing complexes show an enrichment of seed matches in associated mRNAs, other pairing conformations are also significantly enriched (Zhang et al. 2009; Zisoulis et al. 2010). Furthermore, a recent study utilizing RIPchipSRM demonstrates that a substantial portion of *mir-58* target mRNAs have imperfect seed matches (Jovanovic et al. 2012).

The seemingly disparate outcomes from the various studies on the microRNA seed are in part due to the differences in experimental models and resultant interpretations. Much of the mammalian work relies heavily on tissue culture systems or repor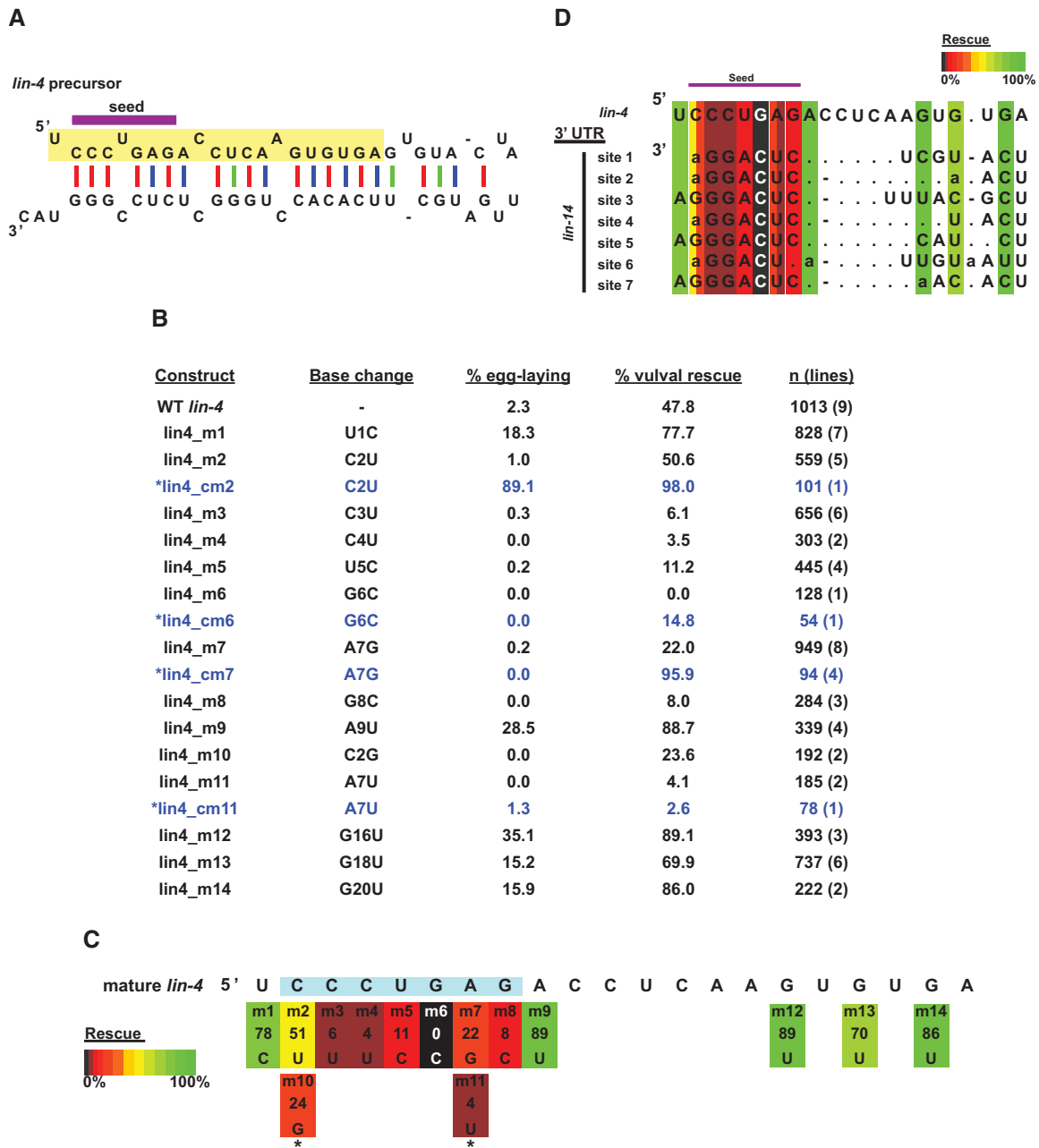ter assays that provide very clear assays but may lack certain features and conditions under which physiological microRNA-mediated regulation takes place. Even for in vivo work, microRNA functions must frequently be assayed using sensor reporters or by measuring changes in levels of mRNAs and proteins, which are challenging to connect biologically to microRNA function. The heterochronic *C. elegans* microRNAs, *lin-4* and *let-7*, have well-characterized mutant phenotypes and genetically validated microRNA target genes (Lee et al. 1993; Reinhart et al. 2000). Hence, these two microRNAs have provided a uniquely amenable system in which to assay functional consequences of imperfect seed matches under physiological conditions. In this work, we have introduced seed sequence variants of *lin-4* and *let-7* into microRNA mutant animals and evaluated their abilities to provide functional rescue.

RESULTS

lin-4 rescue is drastically affected by central seed mutations while peripheral seed mutations retain *lin-4* activity

In order to test the functionality of the microRNA seed, *lin-4* variants (lin4_m1 to lin4_m14) were made with mutations in the 5′ end of *lin-4* and accompanying compensatory mutations on the microRNA star strand to preserve microRNA precursor structure (Fig. 1A,C). These expression constructs derived from a 695-nt genomic fragment that has been used in a variety of studies to rescue *lin-4* null mutant animals (Lee et al. 1993; Zhang et al. 2011). Using microinjection, transgenic lines carrying each lin4_m construct in the *lin-4 (e912)* null background were created. Adults from each line were then assayed for *lin-4* activity by examination of vulval structures formed during development and by functional analysis of egg-laying behavior that requires these structures (see below).

lin-4 is important in the L1-to-L2 transition; animals lacking *lin-4* are heterochronically retarded, reiterating L1-specific events at subsequent larval stages (Chalfie et al. 1981). Terminal differentiation is absent in certain cell types (Chalfie et al. 1981). As a result of lineage and differentiation defects, these mutants are vulvaless (leading to an inability to extrude eggs) and lack adult alae, a set of longitudinal ridges in the cuticle formed by hypodermal seam cells (Chalfie et al. 1981; Ambros and Horvitz 1984). The highly consistent nature of these phenotypes allows a sensitive and definitive assay for *lin-4* function based on capacity for phenotypic rescue (Lee et al. 1993; Zhang and Fire 2010; Zhang et al. 2011). Due to differences in transgene expression levels between transgenic lines, a range of vulval phenotypes associated with *lin-4* rescue was observed. In assaying transgenic animals for *lin-4* activity, they were categorized into one of the following four vulval phenotypes: (i) vulvaless, (ii) protruding vulva (pvl), (iii) bursting vulva (bursting), or (iv) egg-laying (Supplemental Fig. 1). We considered vulval phenotypes

A

lin-4 precursor

D

B

Construct	Base change	% egg-laying	% vulval rescue	n (lines)
WT *lin-4*	-	2.3	47.8	1013 (9)
lin4_m1	U1C	18.3	77.7	828 (7)
lin4_m2	C2U	1.0	50.6	559 (5)
*lin4_cm2	C2U	89.1	98.0	101 (1)
lin4_m3	C3U	0.3	6.1	656 (6)
lin4_m4	C4U	0.0	3.5	303 (2)
lin4_m5	U5C	0.2	11.2	445 (4)
lin4_m6	G6C	0.0	0.0	128 (1)
*lin4_cm6	G6C	0.0	14.8	54 (1)
lin4_m7	A7G	0.2	22.0	949 (8)
*lin4_cm7	A7G	0.0	95.9	94 (4)
lin4_m8	G8C	0.0	8.0	284 (3)
lin4_m9	A9U	28.5	88.7	339 (4)
lin4_m10	C2G	0.0	23.6	192 (2)
lin4_m11	A7U	0.0	4.1	185 (2)
*lin4_cm11	A7U	1.3	2.6	78 (1)
lin4_m12	G16U	35.1	89.1	393 (3)
lin4_m13	G18U	15.2	69.9	737 (6)
lin4_m14	G20U	15.9	86.0	222 (2)

C

FIGURE 1. *lin-4* rescue is drastically affected by central seed mutations while peripheral seed mutations retain activity. (*A*) Partial *lin-4* precursor structure with the mature *lin-4* microRNA highlighted (yellow box). Red and blue lines represent Watson–Crick base pairs (C:G and A:U, respectively), with green lines representing wobble base pairs (G:U). (*B*) Percentages of transgenic animals exhibiting vulval morphology and egg-laying rescue for each lin4_m construct are shown. Results from CRISPR/Cas9-derived *lin-4* mutant animals are in blue (lin4_cm). "n" refers to total number of animals assayed for each lin4_m/lin4_cm mutation, with the subsequent value in parentheses referring to the total number of independent lines assayed. For completeness, we have included a construct (lin4_m6) tested in a single line; although subject to doubt accompanying any individual transgenic line, we note (i) a strong consistency between multiple independently derived lines from other constructs tested in this study, (ii) effects on rescue were consistent with those of other proximal alterations, and (iii) supporting data from endogenous *lin-4* mutants (lin4_cm6). (*C*) *lin-4* mutations and their rescue activities. The seed sequence is highlighted by the light blue box. The different colored boxes *below lin-4* sequence represent different levels of *lin-4* vulval rescue activities (gray = 0.0%, yellow = 50.0%, and green = 100%) corresponding to the lin4_m mutation depicted in the box at the same position. The numbers *within* the colored boxes are the average percentages of transgenic mutant animals showing vulval rescue. The two mutations marked with an asterisk (*) are mutations that result in loss of pairing with known target binding sites, in contrast to the mutations depicted *above* that result in wobble pairing at those positions. (*D*) Binding sites in *lin-14* 3′ UTR (*bottom*) are aligned against *lin-4* mature microRNA (*top*) (Lee et al. 1993; Lewis et al. 2005; Bartel 2009; Jan et al. 2011). Dots indicate absent bases at those positions; dashes indicate one or more non-complementary bases. Lowercase letters represent complementary bases as a result of mutations made in lin4_m constructs. First miRNA nucleotides and alignments of 5′ ends of binding sites have been included for completeness although their functionalities have not been demonstrated. Colored bars represent levels of *lin-4* vulval rescue ([gray] 0.0%, [yellow] 50.0%, [green] 100%) for mutations at specific positions. A position that is highlighted by two colored bars has two different mutations at the same position and each color represents the rescue activity of each specific mutation.

that show any signs of vulval formation (pvl, bursting and egg-laying; all absent in *lin-4* null mutant populations) as indicative of functional *lin-4* rescue.

As shown in Figure 1B and C, the mutants that were tested in the non-seed region (lin4_m1, m9, m12-m14) provided robust rescue of the vulvaless phenotype in a *lin-4* null mutant genetic background, with an average of 69.9% to 89.1% of transgenic *lin-4(e912)* animals exhibiting vulval formation (pvl, bursting and egg-laying). These observed rescue values are comparable to (and in some cases marginally better than) those of transgenic lines carrying a wild-type (WT) *lin-4* DNA construct (wild-type average 47.8%, standard deviation 25.6, standard error 8.5, Supplemental Fig. 2ai). The rescue results demonstrate flexibility in the non-seed regions of *lin-4* in promoting developmental progression.

For mutations in the *lin-4* seed region, with the exception of lin4_m6 (0.0%), some *lin-4* activity was observed (3.5%–50.6%, Fig. 1B,C). lin4_m2 (50.6%) provided comparable *lin-4* rescue to WT *lin-4* construct (47.8%), while *lin-4* activity was drastically reduced in lin4_m3-m5, m7-m8 and m10-m11 (3.5% to 23.6%). In particular, we note that the central seed mutations (lin4_m3-m6) were more severely affected than the peripheral seed mutations (lin4_m2, m7-m8) (Fig. 1C). The *lin-4* activity observed in lin4_m2 may be explained by an examination of the interactions between *lin-4* and its predicted binding sites in *lin-14* (Lee et al. 1993) in Figure 1D. Although the C-to-U change would have resulted in loss of pairing to *lin-14* sites 3, 5, and 7, this mutation also results in a gain of additional pairing for sites 1, 2, 4, and 6 (Fig. 1D). Therefore, the compensatory gain of better *lin-14* target sites could explain the functionality of lin4_m2 microRNA, despite having a seed mutation. Comparing lin4_m2 and lin4_m10 (Fig. 1C), we note that a C-to-G change at the same position reduced *lin-4* rescue, in agreement with our above observation, as this nucleotide change would result in loss of pairing for all seven *lin-14* sites (Fig. 1D). By comparing lin4_m7 to lin4_m11 (Fig. 1C,D), we find that a mutation that allows wobble pairing with described target sites (so pairing may be weakened but not broken) retains a greater degree of activity than a mutation that would be expected to fully abrogate base-pairing. This suggests that wobble pairing in the seed region can suffice for a considerable contribution to *lin-4* function.

To ensure that an observed limitation of *lin-4* rescue with some mutant constructs was not due to low lin4_m expression, the small RNA populations in transgenic *lin-4(e912)* L4 animals were sequenced and ratios of *lin-4* to other microRNAs compared. Transgenic lines with low *lin-4* activities (lin4_m3-m5, m7) showed levels of *lin-4* transcripts that were comparable (0.48%–6.91%, Supplemental Fig. 2b) to WT animals (0.97%) or transgenic lines with robust *lin-4* rescue (lin4_m1-m2, 0.55%–2.40%).

The partial character of rescue conferred by transgenes with the *lin-4* wild-type construct (average 47.8%, Fig. 1B) raised the concern that our observations may be affected by variations sometimes seen with extrachromosomal array experiments, caused by differences in copy number, expression level, and regulatory elements. In order to determine if our observations from extrachromosomal arrays reflect sequence requirements at the level of the endogenous gene, we used a CRISPR/Cas9 oligonucleotide-mediated conversion strategy (Zhao et al. 2014), combined with a homologous repair (HR) marker coconversion strategy (Arribere et al. 2014), to create *lin-4* seed mutants at the endogenous locus. Four endogenous *lin-4* mutants (lin4_cm2, cm6, cm7 and cm11) were made and assayed for *lin-4* activities (Fig. 1B). We found that the CRISPR/Cas9 mutants behaved similarly to their corresponding transgene-based *lin-4* mutants; both lin4_cm6 and cm11 mutants have little *lin-4* activity (Fig. 1B). As further support for the lack of *lin-4* activity in these mutants, additional lin4_cm6 and lin4_cm11 animals (807 and 304, respectively) were screened and none were found to have egg-laying abilities. lin4_cm2 and lin4_cm7 mutants showed a higher degree of rescue than their transgene-based counterparts (lin4_m2 and m7, Fig. 1B). As with the above-mentioned partial rescue by wild type *lin-4* transgenes, the improved rescue by lin4_cm2 and lin4_cm7 endogenous gene mutants may reflect less-than-precise expression levels and patterns for extrachromosomal arrays (Kelly et al. 1997).

Sequence changes within the microRNA would result in changes in hybridization between the microRNA and target mRNA. An initial metric to categorize such changes is calculated free energy of interaction; it has previously been shown that translational repression by microRNAs can be associated with free energy of binding in the microRNA 5′ region (Doench and Sharp 2004). To evaluate potential relationships between energies of interaction and observed rescue activities, we used the RNAhybrid tool (Rehmsmeier et al. 2004) to estimate minimum free energies of hybridization between each of the *lin-4* mutant microRNAs and 3′ UTRs of several known *lin-4* target mRNAs (Fig. 1D, Supplemental Fig. 2c,d; Lee et al. 1993; Moss et al. 1997; Abrahante et al. 2003). Two sets of simulations were performed, first using the full-length microRNA variant and subsequently using just the microRNA seed region. Although by necessity, such calculations are estimates, they can provide some sense of simple energy-of-interaction models for given effector::target interactions as a predictor for microRNA function. As shown in Supplemental Figure 2d, we see an evident relationship (albeit imperfect) between the best target interaction energies and *lin-4* function, with mutant activities most closely associated with predicted pairing in the seed. These comparisons are consistent with supporting the expected "seed" contribution to *lin-4* function (Lewis et al. 2003; Doench and Sharp 2004; Kiriakidou et al. 2004; Kloosterman et al. 2004; Krutzfeldt et al. 2005; Beitzinger et al. 2007; Baek et al. 2008; Selbach et al. 2008), while other regions (including bases immediately flanking the seed) apparently have flexibility in their sequence (Fig. 1C).

let-7 rescue is retained by central seed mutations while non-seed mutant effects support potential noncanonical contributions to *let-7* function

The *C. elegans let-7* system provides a second case in which we can assay biological function of a microRNA by its ability to promote normal developmental progression, with *let-7* playing an important role in *C. elegans* heterochronic development but at later larval transitions than *lin-4* (Reinhart et al. 2000; Vadla et al. 2012). Animals lacking *let-7* function burst at the vulva during L4-to-adult transition, causing lethality (Reinhart et al. 2000). Adult alae formation is also delayed in *let-7* mutant animals (Ambros 1989; Reinhart et al. 2000). As with *lin-4*, the highly penetrant phenotypes (100% of *let-7* null animals lack adult alae at L4 moult (Reinhart et al. 2000) and 100% ($n = 89$) burst during L4-to-adult transition) provide a sensitive biological assay for *let-7* microRNA function. We constructed a series of *let-7* seed and non-seed variants (let7_m1 to let7_m13, Fig. 2A, D), each derived from a 2.5-kb genomic fragment that rescued lethality in *let-7* mutant animals (Reinhart et al. 2000). As with the *lin-4* variants tested above, each *let-7* variant was also constructed with accompanying (compensatory) mutations on the star strand to ensure a structured *let-7* precursor was formed. We obtained transgenic lines expressing let7_m constructs and assayed in a *let-7* null mutant background [*let-7 (mn112)*, (Meneely and Herman 1979; Reinhart et al. 2000)] for suppression of the bursting phenotype. Transgenic animals were categorized into one of three vulval phenotypes: (i) egg-laying, (ii) pvl, or (iii) burst (Supplemental Fig. 1). Both egg-laying and pvl transgenic animals were considered as exhibiting *let-7* activities.

Surprisingly, all of the initial *let-7* mutant constructs that we tested appeared to provide some functional rescue of *let-7* mutant defects, with particularly prevalent rescue of the vulval bursting phenotype (Fig. 2B,Di; Supplemental Fig. 3a). A range of bursting phenotype rescue activities was observed, from 70.5% to 100.0% of transgenic animals surviving to adulthood (Fig. 2B,Di; Supplemental Fig. 3a). As a negative control, we made a construct (let7_seedmut) that had a 5-nt mutation in the *let-7* seed sequence (with accompanying compensatory mutations on the star strand, Fig. 2Ci). Transgenic animals expressing let7_seedmut failed to rescue the bursting phenotype (Fig. 2B; Supplemental Fig. 3a). These results suggest that while the *let-7* seed sequence was important in mediating function (as demonstrated by the negative control, Fig. 2B; Supplemental Fig. 3a), single and double seed mutations in *let-7* were well tolerated (Fig. 2B,Di).
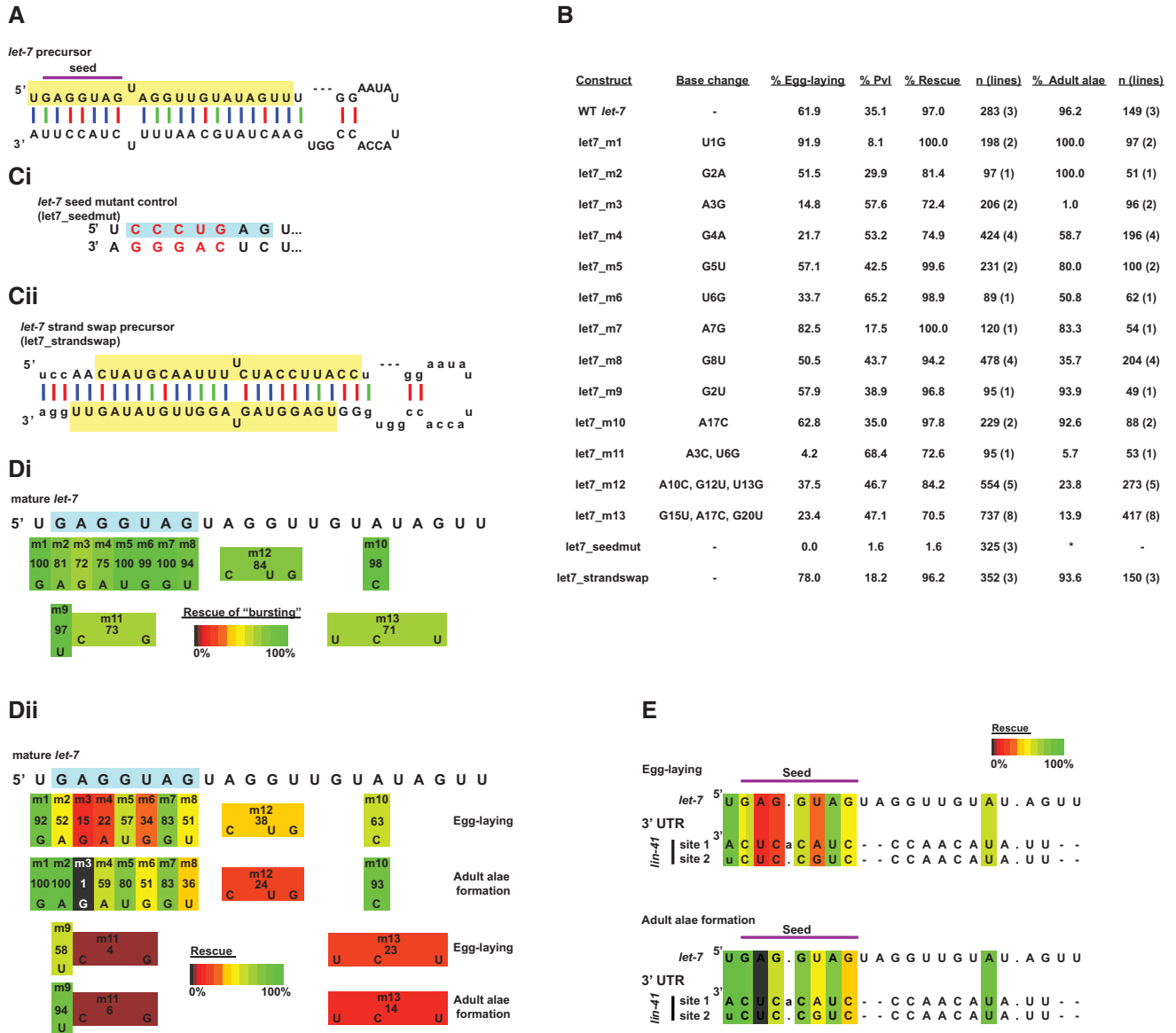
To differentiate functionality between the let7_m constructs, we looked at two additional rescue phenotypes involving specific cellular structures forming during late larval development. Requiring that adult animals become competent for egg laying (and not just avoid explosion at the L4-to-adult transition) provides a somewhat more strin-gent assay for proper function of the lineages and cells giving rise to the vulval machinery. In addition, we assayed for adult alae formation in the transgenic young adults, testing for an ability of lateral hypodermal cells to produce a stage-specific cuticular structure in animals that have recently undergone the L4-to-adult molt (Cox et al. 1981). As shown in Figure 2B,Dii and Supplemental Figure 3a, mutations in the *let-7* seed region affected rescue of egg-laying to varying degrees. In particular, rescue was greatly reduced in let7_m3 and m4 (14.8% and 21.7%, respectively), while other seed mutations retained substantial rescue of egg-laying (33.7% to 82.5%). For adult alae formation in young adults (Fig. 2B, Dii; Supplemental Fig. 3b), we found that with the exception of let7_m3, the *let-7* seed mutants robustly rescued adult alae formation (35.7% to 100.0%). Compared to flanking *let-7* seed mutations, let7_m5 has higher *let-7* activities (Fig. 2Dii). Examining predicted interactions between *let-7* and potential *lin-41* binding sites (Reinhart et al. 2000), we find that the G-to-U change in let7_m5 results in loss of pairing in the middle of the seed region in site 2 but retains pairing in site 1 (Fig. 2E); in this sense, it may not be surprising that let7_m5 is less disruptive. In aggregate, the egg laying and alae results both corroborated the observations from bursting rescue that seed mutations provide *let-7* function and illustrated the diversity in functionality retained by the different seed mutations (Fig. 2D).

To investigate the functional contributions of the non-seed region, single and multiple mutations (let7_m1, m10, m12-m13, Fig. 2D) were made in the non-seed region and assayed for *let-7* function as before. Although two single-site non-seed mutations (let7_m1, m10) provided robust *let-7* rescue of both egg-laying and adult alae formation phenotypes (Fig. 2B,Dii), a pair of multisite non-seed mutations (let7_m12, m13) greatly reduced rescue in both egg-laying and adult alae formation (Fig. 2B,Dii). These results are consistent with a substantial contribution of the *let7* non-seed region to *let-7* function (Vella et al. 2004a).

Energy calculations similar to those described above for *lin-4* are shown in Supplemental Figure 3e. Predicted interactions between *let-7* and several known target genes (*lin-41*, *lin-14*, *hbl-1*, *daf-12*, and *lin-28*) are found in Figure 2E and Supplemental Figure 3d (Reinhart et al. 2000; Abrahante et al. 2003; Lewis et al. 2003; Grosshans et al. 2005). As expected from the activities of the individual mutants, rescue activities of *let-7* variants were not well correlated with either minimum free energy (mfe) calculated from *let-7* seed region or full-length *let-7* microRNA (Supplemental Fig. 3e).

The availability of a physiological assay for *let-7* function provides an opportunity to address additional questions regarding the relationship between synthetic structure and eventual function. Using this assay, we address the question of whether the order of *let-7* and of the corresponding *let-7** microRNA in the precursor was critical for function. A let7_strandswap construct was designed to produce *let-7* microRNA from the 3′ arm of the microRNA precursor

A

let-7 precursor

Ci

Cii

Di

Dii

B

Construct	Base change	% Egg-laying	% Pvl	% Rescue	n (lines)	% Adult alae	n (lines)
WT *let-7*	-	61.9	35.1	97.0	283 (3)	96.2	149 (3)
let7_m1	U1G	91.9	8.1	100.0	198 (2)	100.0	97 (2)
let7_m2	G2A	51.5	29.9	81.4	97 (1)	100.0	51 (1)
let7_m3	A3G	14.8	57.6	72.4	206 (2)	1.0	96 (2)
let7_m4	G4A	21.7	53.2	74.9	424 (4)	58.7	196 (4)
let7_m5	G5U	57.1	42.5	99.6	231 (2)	80.0	100 (2)
let7_m6	U6G	33.7	65.2	98.9	89 (1)	50.8	62 (1)
let7_m7	A7G	82.5	17.5	100.0	120 (1)	83.3	54 (1)
let7_m8	G8U	50.5	43.7	94.2	478 (4)	35.7	204 (4)
let7_m9	G2U	57.9	38.9	96.8	95 (1)	93.9	49 (1)
let7_m10	A17C	62.8	35.0	97.8	229 (2)	92.6	88 (2)
let7_m11	A3C, U6G	4.2	68.4	72.6	95 (1)	5.7	53 (1)
let7_m12	A10C, G12U, U13G	37.5	46.7	84.2	554 (5)	23.8	273 (5)
let7_m13	G15U, A17C, G20U	23.4	47.1	70.5	737 (8)	13.9	417 (8)
let7_seedmut	-	0.0	1.6	1.6	325 (3)	*	-
let7_strandswap	-	78.0	18.2	96.2	352 (3)	93.6	150 (3)

E

FIGURE 2. *let-7* rescue is retained by central seed mutations while non-seed mutant effects support potential noncanonical contributions to *let-7* function. (*A*) Partial *let-7* precursor structure with the mature *let-7* microRNA highlighted (yellow box). Red and blue lines represent Watson–Crick base pairs (C:G and A:U, respectively), with green lines representing wobble base pairs (G:U). (*B*) Percentages of transgenic animals exhibiting rescue of bursting phenotype, egg-laying phenotype, and adult alae formation for each let7_m construct. "*n*" refers to total number of transgenic animals assayed for each let7_m construct, with the subsequent value in parentheses referring to the total number of independent transgenic lines assayed. (*) High-frequency bursting of transgenic let7_seedmut animals precludes the assay for adult alae formation in young adults. (*C*) Structures of two complex *let-7* variants. (i) The let7_seedmut construct consists of a multiple seed mutation with five substitutions (shown in red letters) within the seven base seed; black letters are unchanged from WT *let-7* sequence. The seed sequence is highlighted by the light blue box. (ii) The let7_strandswap mutation swaps the *let-7* microRNA sequence from the 5′ to the 3′ position in the precursor. Yellow boxes highlight the expected small RNA produced from either arm ([*top*] wild-type *let-7**, [*bottom*] wild-type *let-7*). The lowercase letters are the unchanged WT *let-7* sequence (*D*) Rescue activities of let7_m constructs. The seed sequence is highlighted by the light blue box. The different colored boxes *below let-7* sequence represent different levels of *let-7* rescue ([gray] 0.0%, [yellow] 50.0%, [green] 100%) corresponding to the let7_m mutation depicted in the box at the same position. (i) Rescue of vulval bursting phenotype. (ii) Rescue of egg-laying and adult alae formation phenotypes. The *top* row of colored boxes refers to rescue of egg-laying while the *bottom* row refers to rescue of adult alae formation in young adults. The numbers within the colored boxes are the average percentages of transgenic mutant animals showing specific rescue activity for let7_m constructs. (*E*) Binding sites in *lin-41* 3′ UTR (*bottom*) are aligned against *let-7* mature microRNA (*top*) (Reinhart et al. 2000; Lewis et al. 2005; Bartel 2009; Jan et al. 2011; Ecsedi et al. 2015). Dots indicate absent bases at those positions; dashes indicate one or more noncomplementary bases. Lowercase letters represent complementary bases as a result of mutations made in let7_m constructs. First miRNA nucleotides and alignments of 5′ ends of binding sites have been included for completeness although their functionalities have not been demonstrated. Colored bars represent levels of *let-7* activities ([gray] 0.0%, [yellow] 50.0%, [green] 100%; [*top* set] rescue of egg-laying, [*bottom* set] rescue of adult alae formation) for mutations at specific positions. The multiple mutations are not depicted here.

(Fig. 2Cii). As shown in Figure 2B, 96.2% of transgenic worms producing *let-7* from the 3′ arm were rescued for *let-7* bursting phenotype. Alae formation and egg-laying were also observed at near-wild-type levels in transgenic animals carrying the arm-swapped *let-7*. To ensure that wild-type mature *let-7* microRNA was produced from the let7_strandswap construct, the small RNA populations in transgenic *let-7 (mn112)* L4 animals were sequenced and ratios of *let-7* to other microRNAs compared (Supplemental Fig. 3c). We found that substantial amounts of wild-type mature *let-7* microRNA was made (Supplemental Fig. 3c) and as in wild-type animals, the major small RNA species sequenced (from the let7_strandswap construct) was *let-7* and not *let-7** (175 counts versus 3 counts, respectively). These results demonstrate that mature *let-7* microRNA can be produced from either arm of the *let-7* precursor with both providing *let-7* function.

DISCUSSION

In this work, we have sought to examine the functional relevance of imperfect microRNA seed matches. Prior studies on the microRNA seed sequence have arrived at somewhat disparate conclusions, depending on the experimental system and assays used to determine microRNA activity (Doench and Sharp 2004; Brennecke et al. 2005; Didiano and Hobert 2006; Beitzinger et al. 2007; Baek et al. 2008; Zhang et al. 2009; Zisoulis et al. 2010; Jovanovic et al. 2012). The lack of consensus is likely, in part, to reflect multiple mechanisms of microRNA-based repression. While tissue culture assays have proven invaluable in dissecting mechanisms for individual modes of regulation (and have been the method of choice for many past studies), their generalization to organism-scale regulatory mechanisms has been hampered by a lack of obvious knockout phenotypes and definitive target genes. Likewise, reporter constructs have provided extremely valuable data, but suffer a limitation that different reporter assays may each mimic only a portion of complex physiological regulation. To accompany and augment such studies, whole-organism assays can provide a valuable data set in the small number of systems where genetic tools and defined knockout phenotypes allow a physiological assay for microRNA function. Such studies are by nature lower throughput than in vitro or reporter assays, requiring both the construction of substantial numbers of transgenic animal lines and their careful phenotypic characterization. Given the contributions and limitations of each type of assay, we expect that long-term understanding of microRNA function will come from a combination of all of the above approaches.

The goal of this work has been to provide analysis of microRNA::target specificity through modifications of the microRNA with a constant (and physiological) pool of mRNA targets. In order to take advantage of an experimental model that would provide a well-characterized physiological assay for microRNA function, we have chosen to study the effects of microRNA sequence on the phenotypic rescue of *lin-4* and *let-7* mutants in *C. elegans*. Both *lin-4* and *let-7* have small numbers of genetically validated target mRNAs (including the heterochronic regulators *lin-14* and *lin-41* [Lee et al. 1993; Reinhart et al. 2000]), with mutations in these targets shown to suppress the corresponding microRNA mutant phenotypes (Lee et al. 1993; Moss et al. 1997; Reinhart et al. 2000; Abrahante et al. 2003; Grosshans et al. 2005).

The *lin-4* and *let-7* systems have already served to inform discussions of microRNA::target interactions. It was of interest in the early studies of *lin-4* (Lee et al. 1993) that partial complementarity could be observed between microRNA and target regions, with perfect base-pairing often present toward the 5′ side of the microRNA [later called the seed region (Lewis et al. 2003)]. Consistent with a substantial seed contribution to microRNA function, the first microRNA point mutant, *lin-4 (ma161)*, has a single base change within the seed region, resulting in a drastic decrease in mature mutant *lin-4* (Supplemental Fig. 2b) and delayed developmental progression (Lee et al. 1993). As the lack of *lin-4* function could have been due to insufficient levels of mutant *ma161* microRNA rather than inherent lack of microRNA function, our transgenic assays have confirmed that overexpression of a *lin-4* variant, *lin-4_m4*, with the same base change (together with a compensatory mutation in the precursor) failed to rescue *lin-4* developmental defects (Fig. 1, Supplemental Fig. 2b).

Genetic analysis of *let-7* provided an initial exemplification of flexibility in the seed region. *let-7 (n2853)*, an allele with a point mutation in the seed sequence, is temperature-sensitive (growth at higher temperatures causes lethality): at a lower growth temperature, the *n2853* seed mutant is functional and capable of normally promoting developmental progression (Reinhart et al. 2000). Although the failure to grow at high temperature could have reflected a failure of the mature seed-mutant microRNA to rescue, observations of microRNA levels indicated that the mutant microRNA was not accumulated to normal levels at the high temperature (Reinhart et al. 2000). Supporting the analysis in the specific case of *let-7 (n2853)* function, we note that the transgenic assays from this work allowed us to confirm that multicopy expression of a precursor-compensated seed mutant at the *let-7 (n2853)* site (Fig. 2D, let7_m5) and in fact of the precise *let-7 (n2853)* precursor (Supplemental Fig. 3f, let7_m15) allow full rescue of the mutant phenotype. We have also created endogenous *let-7* mutations at the same locus and demonstrated that these single-copy mutants behaved like the temperature-sensitive *let-7(n2853)* mutants, regardless of compensatory precursor pairing (let7_cm14 and let7_cm15, Supplemental Fig. 3f). It should be noted that it is still unclear why *let-7 (n2853)* is temperature-sensitive. We do not know if the lack of function is directly due to the imperfect seed match between *let-7 (n2853)* and targets or other unknown factors, such as microRNA processing, that affects the stability of *let-7 (n2853)* microRNA.

The transgene rescue experiments in this paper provide a considerably extended view of seed and non-seed sequence requirements in specifying *lin-4* and *let-7* function, yielding very different pictures for the two microRNAs. *lin-4* assays demonstrated a definitive requirement for the full seed region (second to eighth nucleotide) in biological function. Despite this requirement, we see clear evidence for functionality of microRNAs with some variation in this region. In particular, we note that central seed mutant versions of *lin-4* were capable of supporting events in vulval development and rare egg laying, events that were universally absent in the *lin-4* null mutant background. While central seed mutations greatly reduced *lin-4* function, a peripheral seed mutation (second nucleotide) retained substantial *lin-4* function in our phenotypic assays. Close examination of *lin-4*::*lin-14* interactions revealed that this mutation could strengthen pairing in several weaker binding sites, compensating for the loss of pairing in previously strong target sites. These observations suggest that flexibility in the seed region is partly dependent on the presence of weak binding sites (that have potential to become good binding sites) in target genes. None of the non-seed mutant *lin-4* constructs that we tested showed a dramatic effect on microRNA function. This result certainly does not rule out non-seed contributions by *lin-4*, and such contribution might indeed be expected to involve more flexible base-pairing interactions that may resist individual mutant effects. Despite such expectations, the picture of *lin-4* function that emerges from this analysis is one in which physiologically relevant *lin-4*::target interaction relies on a substantial contribution to specificity and energy from seed base-pairing.

In contrast to our *lin-4* results, functional data from the *let-7* mutants provides evidence for substantial flexibility in the physiological microRNA::target interaction. All constructs with single-site mutations in *let-7* seed region rescued the *let-7* bursting phenotype (Fig. 2B,Di). A multisite seed mutant failed to rescue this *let-7* function, indicating the flexibility in seed requirement was not due to a complete lack of seed contribution. Although previous data supported some flexibility in seed region requirements, we were still quite surprised at the generality of rescue of bursting by the point mutants. The conserved extensive pairing of *let-7* 3′ end with its target binding sites provides an explanation for our observations. 3′ compensatory sites with strong 3′ pairing have been shown to be effective in mediating microRNA function and the top three scoring predicted 3′ compensatory sites in *C. elegans* are *let-7* sites (Brennecke et al. 2005; Jan et al. 2011). Two of these sites mediate *lin-41* repression and Ecsedi et al. (2015) have recently demonstrated that sole repression of *lin-41* was sufficient to rescue *let-7* bursting phenotype.

Beyond the non-seed pairing potential by *let-7*, a possible additional contribution to the unexpectedly broad rescue by *let-7* derivatives with impaired seed pairing could be the presence of endogenous *let-7* family microRNAs (*mir-48*, *mir-84*, and *mir-241*) that have identical seed sequences to *let-7*. The

let-7 family microRNAs have overlapping spatial and temporal expression patterns and have been shown to mediate certain functions together (Esquela-Kerscher et al. 2005; Li et al. 2005; Hayes et al. 2006). Although the *let-7* family microRNAs and *let-7* have been shown to be functionally nonredundant (Abbott et al. 2005), overexpression of *mir-84* or *mir-48* in *let-7* mutants suppresses lethality, indicating that *let-7* family microRNAs are able to substitute for *let-7* function if expressed at the appropriate time in the right tissues (Li et al. 2005; Hayes et al. 2006). Despite any possible augmentation of mutant *let-7* function through the action of its microRNA sisters, the lack of *let-7* is lethal and the seed mutants are able to rescue this lethality; thus at least one critical target interaction must be retained by the single-site seed-modified *let-7* microRNAs. *let-7* drives several late larval-to-adult developmental shifts in *C. elegans*, and we note that two additional phenotypes (egg laying and adult alae formation) show a greater requirement for specific sequences in the *let-7* seed (Fig. 2B,Dii). Still, these phenotypes can be partially rescued by mutations that change central bases in the seed, confirming a degree of intrinsic flexibility in *let-7* target interactions beyond those for the bursting phenotype.

Our non-seed rescue results, although with a limited set of constructs, provide a picture of additional contributions to *let-7* function, with both proximal and distal non-seed sequences contributing to egg-laying and alae phenotypes. As our data does not distinguish between deficiencies in target recognition and RISC complex formation and function, we do not know if the exact role of the non-seed region lies in base-pairing with target mRNA or in RISC complex formation and function. This is also true for mutations in the seed region. Contributions of seed and non-seed sequences to phenotype are consistent with observations with reporter constructs in *C. elegans* (Vella et al. 2004a,b), although somewhat at odds with studies utilizing microinjection of concentrated miRNA duplexes into zebrafish embryos (Kloosterman et al. 2004). The variety of observations specific to *C. elegans let-7* in this work and those of Vella et al. (2004a,b) indicate that this may be a particularly complex microRNA in the diversity of its interactions and functions. Consistent with this possibility are studies in which specialized RNA (Cevec et al. 2010) and RNP complexes (Chan et al. 2008) have been observed in association with this microRNA. *let-7* is a highly conserved microRNA, with substantial (and somewhat consistent) diversity in both seed and non-seed regions. In the future, it will be of great interest to compare the functional rescue observations in this paper with functional phenotypic data in other invertebrate systems and in vertebrates.

Taken together, our analysis of *lin-4* and *let-7* functions is strongly illustrative of cooperation between seed and non-seed sequences in microRNA function. Although there is no doubt that the seed sequence is an important determinant for function (evident from the *lin-4* results), the demonstration of functionality of numerous seed variants raises some

interesting and more general questions in microRNA biology. Are there rules that determine seed dependence and non-seed contributions for different microRNA::target systems? Might there be classes of microRNA targets regulated by separate seed and non-seed matches to different microRNAs? Might there be noncanonical microRNA::target interactions where no match to the seed is evident? Are there means to build seed and non-seed contributions into more precise and accurate predictions of microRNA::target relationships?

MATERIALS AND METHODS

Transgene-based expression of seed mutant microRNA constructs

Details of plasmids and strains used in the transgene-based experiments can be found in Tables 1 and 2, respectively.

lin-4 *seed mutation (lin4_m) constructs*

The wild-type (WT) *lin-4 (+)* construct (pHZ018) consists of a 695-nucleotide (nt) genomic fragment, which rescues *lin-4 (e912)* phenotype , in a pCR2.1-TOPO vector backbone (Invitrogen). This genomic fragment contains the 94-nt *lin-4* precursor sequence and flanking genomic 5′ (498 nt) and 3′ (103 nt) sequences. Different lin4_m constructs were made using mutation-containing primers to amplify pHZ018 and replacing WT *lin-4* in pHZ018 with the amplified PCR product. The specific mutations made for each construct are illustrated in Figure 1C. The lin4_m constructs (30 ng/μL or 17 ng/μL) were individually injected into PD7143 [*lin-4 (e912)/mIn1* II; *pha-1(e2123ts)* III; *rde-1(ne300)* V] with a selection marker, pC1 [*pha-1(+)*, 120 ng/μL or 133 ng/μL, respectively] (Granato et al. 1994). Transgenic lines derived from these injections were assayed for *lin-4* activity [sublines without the pharyngeal *mIs14(gfp)* are homozygous for the chromosomal *lin-4(e912)* deletion mutation]. We note that transgenic strains with wild-type *lin-4* are subject to differing degrees of partial rescue and mosaicism, with perfect rescue (full phenotypic rescue and egg-laying) observed in only a fraction of transgenic animals (Supplemental Fig. 2ai). Data for WT *lin-4* rescue were taken from our previous work (Zhang et al. 2011).

let-7 *seed mutation (let7_m) constructs*

The wild-type *let-7* construct (CZ234, gift from Robin Trujillo and Chang-Zheng Chen) consists of a 2460-nt genomic fragment in a pCR2.1-TOPO vector backbone (Invitrogen). This genomic fragment contains the *let-7* precursor sequence and flanking genomic 5′ (1747 nt) and 3′ (614 nt) sequences. Different let7_m constructs were made using mutation-containing primers to amplify CZ234 and replacing WT *let-7* in CZ234 with the amplified PCR product. The specific mutations made for each construct are illustrated in Figure 2D. The let7_m constructs (m1-m13, 3–6 ng/μL) were individually injected into PD5584 [*mnDp1(X;V)/+ V; let-7 (mn112) unc-3(e151)* X] with pPD117.01 (*pmec-7::gfp*, 12–50 ng/μL) and CZ233 (10–44 ng/μL). Transgenic lines derived from these injections (uncoordinated animals with green fluorescent touch receptor neurons) were assayed for *let-7* activity (rescue of vulval "bursting"

TABLE 1. Plasmids used in transgene-based experiments

Plasmid	Description
pC1	*pha-1* (+) (Granato et al. 1994)
pHZ018	*lin-4* (+) (Lee et al. 1993)
pHZ195, pHZ197	lin4_m1
pHZ200, pHZ201	lin4_m2
pHZ204, pHZ202	lin4_m3
pHZ161	lin4_m4
pHZ206, pHZ205	lin4_m5
pHZ216	lin4_m6
pHZ209, pHZ210	lin4_m7
pHZ225	lin4_m8
pHZ227	lin4_m9
pHZ212	lin4_m10
pHZ220	lin4_m11
pHZ269-270	lin4_m12
pHZ267-268	lin4_m13
pHZ264-266	lin4_m14
pHZ181	let7_m1
pHZ184	let7_m2
pHZ186	let7_m3
pHZ190	let7_m4
pHZ236-238	let7_m5
pHZ239	let7_m6
pHZ192	let7_m7
pHZ241-244	let7_m8
pHZ233	let7_m9
pHZ253, pHZ254	let7_m10
pHZ255, pHZ256	let7_m11
pHZ257-260	let7_m12
pHZ261	let7_m13
pHZ123	let7_m14
pHZ129, pHZ131	let7_m15
pHZ228	let7_seedmut
pHZ230-232	let7_strandswap
CZ234	*let-7* (+) (Reinhart et al. 2000)
pPD117.01	*pmec-7::gfp* (Chalfie et al. 1994; Fire et al. 1998)
CZ233	pCR2.1-TOPO vector with no insert (Invitrogen)

phenotype). The let7_m constructs (m14-m15, 3–6 ng/μL) were individually injected into PD5585 [*let-7 (n2853)* X] with pPD117.01 (*pmec-7::gfp*, 12–50 ng/μL) and CZ233 (10–44 ng/μL). Transgenic lines derived from these injections (green fluorescent touch receptor neurons) were assayed for *let-7* activity (rescue of vulval "bursting" phenotype) at 25°C.

Endogenous CRISPR lin-4 and let-7 seed mutants (lin4_cm and let7_cm)

Seed mutations were engineered at the genomic *lin-4* and *let-7* loci as previously described with several modifications (Arribere et al. 2014; Zhao et al. 2014). The sensitivity of microRNAs to seed region changes required us to use a strategy to cleanly introduce the desired mutations without any additional base pair changes. This was accomplished by choosing CRISPR/Cas9 target sequences that overlapped the location of the desired mutation. Repairs templated from the provided oligonucleotide would then be sufficiently unique to resist further cleavage by CRISPR/Cas9. Oligonucleotide

TABLE 2. Strains used in transgene-based experiments

Strain	Description
PD1448	*lin-4 (e912)/mIn1 II; pha-1(e2123ts) III; rde-1(ne300) V; ccEx1448[pC1; pHZ018]*
PD1469	*lin-4 (e912) II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1469[pC1; pHZ018]*
PD1491 (Line 208.3.2.2)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1491[pC1; pHZ195]*
PD1492 (Line 209a.1.1.2)	*lin-4 (e912); pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1492[pC1; pHZ197]*
PD1493 (Line 210.1.2.0)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1493[pC1; pHZ201]*
PD1494 (Line 211.2.1.1)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1494[pC1; pHZ200]*
PD1495 (Line 212a.1.1.0)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1495[pC1; pHZ204]*
PD1496 (Line 213.1.2.1)	*lin-4 (e912) II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1496[pC1; pHZ202]*
PD1497 (Line 214.3.1.0)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1497[pC1; pHZ206]*
PD1498 (Line 215.1.1.1)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1498[pC1; pHZ205]*
PD1499 (Line 216.1.1.0)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx1499[pC1; pHZ209]*
PD2904 (Line 217.1.1.1)	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V; ccEx2904[pC1; pHZ210]*
PD2907 (Line 203z.2.1.1)	*let-7 (mn112) unc-3(e151) X; ccEx2907[pmec-7:gfp; pHZ181]*
PD2908 (Line 205z.1.1.0)	*let-7 (mn112) unc-3(e151) X; ccEx2908[pmec-7:gfp; pHZ184]*
PD2909 (Line 225.1.1.0)	*let-7 (mn112) unc-3(e151) X; ccEx2909[pmec-7:gfp; pHZ186]*
PD2910 (Line 220.3.2.1)	*let-7 (mn112) unc-3(e151) X; ccEx2910[pmec-7:gfp; pHZ190]*
PD2911 (Line 221.2.1.1)	*let-7 (mn112) unc-3(e151) X; ccEx2911[pmec-7:gfp; pHZ190]*
PD2912 (Line 227.1.1.0)	*let-7 (mn112) unc-3(e151) X; ccEx2912[pmec-7:gfp; pHZ192]*
PD2940 (Line 232.1.1.0)	*mnDp1(X;V)/+ V; let-7 (mn112) unc-3(e151) X; ccEx2940[pmec-7:gfp; pHZ230]*
PD2942 (Line 233.1.2.0)	*mnDp1(X;V)/+ V; let-7 (mn112) unc-3(e151) X; ccEx2942[pmec-7:gfp; pHZ232]*
PD5584	*mnDp1(X;V)/+ V; let-7 (mn112) unc-3(e151) X*
PD5585	*let-7 (n2853) X*
PD7143	*lin-4 (e912)/mIn1 II; pha-1 (e2123ts) III; rde-1 (ne300) V*
PD7183, PD7187	*let-7 (n2853) X; ccEx7183/7[pmec-7:gfp; pHZ123]*
PD7184	*let-7 (n2853) X; ccEx7184[pmec-7:gfp; pHZ129]*
PD7185	*let-7 (n2853) X; ccEx7185/7[pmec-7:gfp; pHZ131]*

To balance *lin-4 (e912)*-bearing chromosomes, we used a derivative of the *mIn1* inversion chromosome that carries an integrated *gfp* marker [*mIs14(gfp)*] in addition to a mutant (recessive) copy of *dpy-10(e128)* (Edgley and Riddle 2001).

repair templates and hybridization pairs (Table 3) for cloning into the pRB1017 backbone, as well as the resulting gRNA vectors (Table 4) are listed above. Repair template oligonucleotides were ordered as Ultramers from IDT. Details of strains used in the CRISPR/Cas9 experiments can be found in Table 5.

lin-4 mutants (lin4_cm)

pPD207.594 (gRNA vector) was designed to specifically target *lin-4 (ma161)*, without cutting the corresponding wild-type *lin-4* sequence, and pPD207.627 (gRNA vector) was designed to target the *lin-4** strand (for compensatory mutation). An injection mix comprised of the following was injected into PD7173 *lin-4 (ma161)II/mIn1* heterozygotes: 50 ng/μL pDD162 (*Cas9* expression vector, Dickinson et al. 2013), 25 ng/μL pJA50 (*unc-58* gRNA vector, Arribere et al. 2014), 0.5 uM AF-JA-76 (*unc-58* repair oligonucleotide, Arribere et al. 2014), 25 ng/μL pPD207.594 (*ma161* gRNA vector), 15 ng/μL pPD207.627 (*lin-4** gRNA vector) and 0.5 μM of one of the following *lin-4* mutation-containing repair template oligonucleotides (AF-KLA-193, AF-KLA-194, AF-KLA-195, AF-KLA-196). F1 progeny carrying *mIn1(gfp)* balancer and exhibiting the appropriate homologous recombination marker "shaker" phenotype were singled 3–4 d after injection. Nonshaker GFP-expressing F2 progeny were screened for desired mutations by single worm PCR. Following establishment of lines carrying the *lin-4* mutation balanced by the *mIn1(gfp)* balancer but lacking mutations at coselected *unc-58* locus, individual homozygous *lin-4* mutant worms

were picked at L3/L4 stage and monitored over several days for egg-laying, protruding vulva, and bursting (Supplemental Fig. 2aii). We noted that occasionally some of the monitored worms failed to thrive; they usually shriveled and died before becoming adults. These instances were omitted from our assays for rescue activities and were also observed in lin4_m transgenic worms.

Outcrosses with wild type (VC2010) were performed to eliminate *unc-58* mutation in PD1095 (lin4_cm7). Resultant males were successively crossed to *mInl* II homozygous hermaphrodites (from PD7173) twice and *lin-4 (cc10952)/mIn1* II lines (PD1263-5) were identified from following single progeny.

let-7 mutants (let7_cm)

To create *let-7* mutants, an injection mix comprised of the following was injected into VC2010: 50 ng/μL pDD162 (*Cas9* expression vector, Dickinson et al. 2013), 25 ng/μL pJA58 (*dpy-10* gRNA vector, Arribere et al. 2014), 0.5 μM AF-ZF-827 (*dpy-10* repair oligonucleotide, Arribere et al. 2014), 50 ng/μL pPD207.646 (*let-7* gRNA vector), 50 ng/μL pPD207.658 (*let-7** gRNA vector), and 0.5 μM of one of the following *let-7* mutation-containing repair template oligonucleotides (AF-KLA-291, AF-KLA-266). F1 progeny exhibiting the appropriate homologous recombination marker "roller" were singled 3–4 d after injection and genotyped for desired *let-7* mutations. Single F2s (from identified *let-7* mutant F1s) were used to establish homozygous lines. The *let-7* genotypes of these selected F2s were also confirmed by PCR. let7_cm14/15 strains were further

TABLE 3. Oligonucleotide repair templates used to make CRISPR/Cas9 endogenous mutants

Oligo name	Sequence	Description
AF-JA-76	Arribere et al. (2014)	*unc-58 (e665)*
AF-KLA-193	caatttctagagtttggttggtttatgagtttatgcttccggcctgttTcctgagacctcaagtgtgagtgtactattgatgcttcacacctgggctctccgg Ataccaggacggtttgagcagatctttttttctgttttcacggggtttttt	lin4_cm2
AF-KLA-194	tttctagagtttggttggtttatgagtttatgcttccggcctgttccctCagacctcaagtgtgagtgtactattgatgcttcacacctgggctctGcgggta ccaggacggtttgagcagatctttttttctgttttcacggggtt	lin4_cm6
AF-KLA-195	ttctagagtttggttggtttatgagtttatgcttccggcctgttccctgGgacctcaagtgtgagtgtactattgatgcttcacacctgggctcCccgggta ccaggacggtttgagcagatctttttttctgttttcacggggtt	lin4_cm7
AF-KLA-196	ttctagagtttggttggtttatgagtttatgcttccggcctgttccctgTgacctcaagtgtgagtgtactattgatgcttcacacctgggctcAccgggta ccaggacggtttgagcagatctttttttctgttttcacggggtt	lin4_cm11
AF-KLA-266	gaaagttgtgagagcaagacgacgcagcttcgaagagttctgtctccggtaagAtagaaaattgcatagttcaccggtggtaatattcca	let7_cm16
AF-KLA-291	tcaggcaagcaggcgattggtggacggtctacactgtggatccggtgagAtagtaggttgtatagtttggaatattaccaccggtgaactatgcaattt tctaTcttaccggagacagaactcttcgaagctgcgtcgtcttgctctcacaactt	let7_cm14/15
AF-ZF-827	Arribere et al. 2014	*dpy-10 (cn64)*

Mutated residues shown in uppercase.

outcrossed to eliminate *dpy-10* mutation. Specifically, these strains were crossed with wild-type males (with GFP-marked X chromosome, generated from a cross between PD4810 [*plmn-1::lmn-1:: gfp,* Liu et al. 2000] and wild-type VC2010). Singled progeny were selfed and let7_cm14/15 strain homozygous lines (non-GFP) were obtained after confirming *let-7* and *dpy-10* genotypes by PCR. The endogenous *let-7* mutants were monitored for development at 16C, 20C, 23C, and 25C.

Preparation of small RNA libraries for Illumina sequencing

Small RNA data from the N2 and *lin-4 (e912)* strains were taken from our previous work (Zhang et al. 2011). The *lin-4* transgenic strains were derived by segregating lines of transgene-carrying animals homozygous for the chromosomal *lin-4 (e912)* mutation from the *lin-4 (e912)/mIn1* transgenic lines created from the injec-

tion of lin4_m constructs. The *let-7* transgenic strains were derived by isolating *let-7 (mn112) unc-3(e151)* animals (uncoordinated animals expressing p*mec-7::gfp*) from the *mnDp1/+; let-7 (mn112) unc-3(e151)* transgenic lines created from the injection of let7_m constructs. All strains used in the making of small RNA libraries were harvested as synchronized L4 populations grown at 25°C and frozen as pellets in liquid N_2. Synchronization was obtained following treatment with sodium hypochlorite to isolate embryos as previously described (Brenner 1974; Stiernagle 2006).

Small RNAs were isolated from these frozen pellets using the mirVana Isolation Kit (Ambion) and used for library preparation following the 5′-monophosphate method that enriches for microRNAs, as previously described (Lau et al. 2001). The exception was that modified 5′-adaptor oligonucleotides, which included 4-nt barcodes, were utilized to enable pooling of several libraries for Illumina sequencing (Parameswaran et al. 2007).

TABLE 4. Plasmids used in CRISPR/Cas9-based experiments

Plasmid	Description	Annealed oligo sequences	Reference
pPD207.594	*lin-4 (ma161)* gRNA vector	TCTTGacttgaggtctca*A*ggaac, AAACgttcc*T*tgagacctcaagtC	This paper
pPD207.627	*lin-4** gRNA vector	TCTTGcctgggctctccgggtacc, AAACggtacccggagagcccaggC	This paper
pPD207.646	*let-7* gRNA vector	TCTTGctacactgtggatccggtg, AAACcaccggatccacagtgtagC	This paper
pPD207.658	*let-7** gRNA vector	TCTTGatgcaatttctaccttac, AAACgtaaggtagaaaattgcatC	This paper
pRB1017	gRNA backbone vector	–	Arribere et al. (2014)
pDD162	*Cas9* expression vector	–	Dickinson et al. (2013)
pJA50	*unc-58* gRNA vector	–	Arribere et al. (2014)
pJA58	*dpy-10* gRNA vector	–	Arribere et al. (2014)

For gRNA vectors, annealed oligo sequences in lowercase are identical to targeted genomic sequences. TCTT and AAAC create overhangs for cloning into the *BsaI* site in pRB1017. The additional G (C on reverse complement strand) is needed for expression from the U6 promoter (Arribere et al. 2014). Bases within the targeted genomic region that vary from wild type are shown in italics.

Sequence analysis of small RNA libraries

Thirty-six nucleotide reads were generated from the small RNA libraries using the Illumina Genome Analyzer system. The captured small RNA sequences, after removal of linker and adaptor sequences, were aligned using BLAT against a reference list of known *C. elegans* mature microRNA sequences downloaded from miRBase (http://www. mirbase.org). The number of captured small RNA sequences that perfectly matched *C. elegans* annotated microRNAs and lin4_m/ let7_m microRNAs was obtained for each library.

Calculation of minimum free energy hybridization (mfe)

Using RNAhydrid (Rehmsmeier et al. 2004), mfe were calculated between entire 3′ UTR sequences of target genes and microRNA sequences (both full-length microRNA and microRNA seed sequence). We listed the

TABLE 5. Strains used in CRISPR/Cas9-based experiments

Strain	Genotype	Description
PD1086	*lin-4 (cc1086)/mln1* II	lin4_cm11
PD1088	*lin-4 (cc1088)/mln1* II	lin4_cm2
PD1093	*lin-4 (cc1093)/mln1* II	lin4_cm6
PD1095	*lin-4 (cc10952)/mln1* II; *unc-58(cc10951)* X	lin4_cm7 (parental line)
PD1254/5	*let-7 (cc12511)* X	let7_cm14
PD1256	*let-7 (cc12521)* X	let7_cm14
PD1259/ 60/62	*let-7(cc12531)* X	let7_cm15
PD1263/4/ 5	*lin-4 (cc10952)/mln1* II	lin4_cm7 (outcrossed lines)
PD1266	*let-7(cc1266)* X	let7_cm16
PD1267	*let-7(cc1267)* X	let7_cm16
PD4810	*ccIs4810[plmn-1::lmn-1:: gfp]* X	Liu et al. 2000
PD7173	*lin-4 (ma161)/mln1* II	–
VC2010	Wild type	Thompson et al. 2013

three most favorable mfe for each microRNA sequence, as well as the average mfe calculated from the top three mfe across all tested target 3′ UTRs.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

REFERENCES

Abbott AL, Alvarez-Saavedra E, Miska EA, Lau NC, Bartel DP, Horvitz HR, Ambros V. 2005. The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans*. *Dev Cell* **9:** 403–414.

Abrahante JE, Daul AL, Li M, Volk ML, Tennessen JM, Miller EA, Rougvie AE. 2003. The *Caenorhabditis elegans* hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. *Dev Cell* **4:** 625–637.

Alvarez-Saavedra E, Horvitz HR. 2010. Many families of *C. elegans* microRNAs are not essential for development or viability. *Curr Biol* **20:** 367–373.

Ambros V. 1989. A hierarchy of regulatory genes controls a larva-to-adult developmental switch in *C. elegans*. *Cell* **57:** 49–57.

Ambros V, Horvitz HR. 1984. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226:** 409–416.

Arribere JA, Bell RT, Fu BXH, Artiles KL, Hartman PS, Fire AZ. 2014. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics* **198:** 837–846.

Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature* **455:** 64–71.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116:** 281–297.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136:** 215–233.

Beitzinger M, Peters L, Zhu JY, Kremmer E, Meister G. 2007. Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biol* **4:** 76–84.

Brennecke J, Stark A, Russell RB, Cohen SM. 2005. Principles of microRNA-target recognition. *PLoS Biol* **3:** e85.

Brenner S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics* **77:** 71–94.

Bushati N, Cohen SM. 2007. microRNA functions. *Annu Rev Cell Dev Biol* **23:** 175–205.

Cevec M, Thibaudeau C, Plavec J. 2010. NMR structure of the let-7 miRNA interacting with the site LCS1 of lin-41 mRNA from *Caenorhabditis elegans*. *Nucleic Acids Res* **38:** 7814–7821.

Chalfie M, Horvitz HR, Sulston JE. 1981. Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell* **24:** 59–69.

Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. 1994. Green fluorescent protein as a marker for gene expression. *Science* **263:** 802–805.

Chan SP, Ramaswamy G, Choi EY, Slack FJ. 2008. Identification of specific let-7 microRNA binding complexes in *Caenorhabditis elegans*. *RNA* **14:** 2104–2114.

Cox GN, Kusch M, DeNevi K, Edgar RS. 1981. Temporal regulation of cuticle synthesis during development of *Caenorhabditis elegans*. *Dev Biol* **84:** 277–285.

Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. 2013. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat Methods* **10:** 1028–1034.

Didiano D, Hobert O. 2006. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat Struct Mol Biol* **13:** 849–851.

Didiano D, Hobert O. 2008. Molecular architecture of a miRNA-regulated 3′ UTR. *RNA* **14:** 1297–1317.

Doench JG, Sharp PA. 2004. Specificity of microRNA target selection in translational repression. *Genes Dev* **18:** 504–511.

Ecsedi M, Rausch M, Großhans H. 2015. The let-7 microRNA directs vulval development through a single target. *Dev Cell* **32:** 335–344.

Edgley ML, Riddle DL. 2001. LG II balancer chromosomes in *Caenorhabditis elegans*: mT1(II;III) and the mIn1 set of dominantly and recessively marked inversions. *Mol Genet Genomics* **266:** 385–395.

Esquela-Kerscher A, Johnson SM, Bai L, Saito K, Partridge J, Reinert KL, Slack FJ. 2005. Post-embryonic expression of *C. elegans* microRNAs belonging to the lin-4 and let-7 families in the hypodermis and the reproductive system. *Dev Dyn* **234:** 868–877.

Fire A, Ahnn J, Kelly W, Harfe B, Kostas S, Hsieh J, Hsu M, Xu S. 1998. GFP applications in *C. elegans*. In *GFP strategies and applications* (ed. Chalfie M, Kain S), pp. 153–168. Wiley, New York.

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19:** 92–105.

Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. 2011. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat Struct Mol Biol* **18:** 1139–1146.

Granato M, Schnabel H, Schnabel R. 1994. pha-1, a selectable marker for gene transfer in *C. elegans*. *Nucleic Acids Res* **22:** 1762–1763.

Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27:** 91–105.

Grosshans H, Johnson T, Reinert KL, Gerstein M, Slack FJ. 2005. The temporal patterning microRNA let-7 regulates several transcription

factors at the larval to adult transition in *C. elegans*. *Dev Cell* **8:** 321–330.

Ha I, Wightman B, Ruvkun G. 1996. A bulged lin-4/lin-14 RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. *Genes Dev* **10:** 3041–3050.

Hayes GD, Frand AR, Ruvkun G. 2006. The mir-84 and let-7 paralogous microRNA genes of *Caenorhabditis elegans* direct the cessation of molting via the conserved nuclear hormone receptors NHR-23 and NHR-25. *Development* **133:** 4631–4641.

Jan CH, Friedman RC, Ruby JG, Bartel DP. 2011. Formation, regulation and evolution of *Caenorhabditis elegans* 3′UTRs. *Nature* **469:** 97–101.

Jovanovic M, Reiter L, Clark A, Weiss M, Picotti P, Rehrauer H, Frei A, Neukomm LJ, Kaufman E, Wollscheid B, et al. 2012. RIP-chip-SRM–a new combinatorial large-scale approach identifies a set of translationally regulated bantam/miR-58 targets in *C. elegans*. *Genome Res* **22:** 1360–1371.

Kelly WG, Xu S, Montgomery MK, Fire A. 1997. Distinct requirements for somatic and germline expression of a generally expressed *Caernorhabditis elegans* gene. *Genetics* **146:** 227–238.

Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A. 2004. A combined computational-experimental approach predicts human microRNA targets. *Genes Dev* **18:** 1165–1178.

Kloosterman WP, Wienholds E, Ketting RF, Plasterk RH. 2004. Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Res* **32:** 6284–6291.

Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M. 2005. Silencing of microRNAs in vivo with "antagomirs". *Nature* **438:** 685–689.

Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* **294:** 853–858.

Lau NC, Lim LP, Weinstein EG, Bartel DP. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294:** 858–862.

Lee RC, Ambros V. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294:** 862–864.

Lee RC, Feinbaum RL, Ambros V. 1993. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75:** 843–854.

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. 2003. Prediction of mammalian microRNA targets. *Cell* **115:** 787–798.

Lewis BP, Burge CB, Bartel DP. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120:** 15–20.

Li M, Jones-Rhoades MW, Lau NC, Bartel DP, Rougvie AE. 2005. Regulatory mutations of mir-48, a *C. elegans* let-7 family MicroRNA, cause developmental timing defects. *Dev Cell* **9:** 415–422.

Liu J, Rolef Ben-Shahar T, Riemer D, Treinin M, Spann P, Weber K, Fire A, Gruenbaum Y. 2000. Essential roles for *Caenorhabditis elegans* lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. *Mol Biol Cell* **11:** 3937–3947.

Meneely PM, Herman RK. 1979. Lethals, steriles and deficiencies in a region of the X chromosome of *Caenorhabditis elegans*. *Genetics* **92:** 99–115.

Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR. 2007. Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet* **3:** e215.

Moss EG, Lee RC, Ambros V. 1997. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA. *Cell* **88:** 637–646.

Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. 2007. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res* **35:** e130.

Rajewsky N. 2006. microRNA target predictions in animals. *Nat Genet* **38:** S8–S13.

Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. 2004. Fast and effective prediction of microRNA/target duplexes. *RNA* **10:** 1507–1517.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403:** 901–906.

Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature* **455:** 58–63.

Sethupathy P, Megraw M, Hatzigeorgiou AG. 2006. A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat Methods* **3:** 881–886.

Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP. 2010. Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* **38:** 789–802.

Stark A, Brennecke J, Russell RB, Cohen SM. 2003. Identification of *Drosophila* MicroRNA targets. *PLoS Biol* **1:** E60.

Stern-Ginossar N, Elefant N, Zimmermann A, Wolf DG, Saleh N, Biton M, Horwitz E, Prokocimer Z, Prichard M, Hahn G, et al. 2007. Host immune system gene targeting by a viral miRNA. *Science* **317:** 376–381.

Stiernagle T. 2006. Maintenance of *C. elegans*. In *WormBook* (ed. The *C. elegans* Research Community). doi: 10.1895/wormbook.1.101.1.

Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, Au V, Chaudhry I, Fernando L, Hutter H, et al. 2013. The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res* **23:** 1749–1762.

Vadla B, Kemper K, Alaimo J, Heine C, Moss EG. 2012. lin-28 controls the succession of cell fate choices via two distinct activities. *PLoS Genet* **8:** e1002588.

Vella MC, Choi EY, Lin SY, Reinert K, Slack FJ. 2004a. The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. *Genes Dev* **18:** 132–137.

Vella MC, Reinert K, Slack FJ. 2004b. Architecture of a validated microRNA::target interaction. *Chem Biol* **11:** 1619–1623.

Yekta S, Shih IH, Bartel DP. 2004. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304:** 594–596.

Zhang H, Fire AZ. 2010. Cell autonomous specification of temporal identity by *Caenorhabditis elegans* microRNA lin-4. *Dev Biol* **344:** 603–610.

Zhang L, Hammell M, Kudlow BA, Ambros V, Han M. 2009. Systematic analysis of dynamic miRNA-target interactions during *C. elegans* development. *Development* **136:** 3043–3055.

Zhang H, Maniar JM, Fire AZ. 2011. "Inc-miRs": functional intron-interrupted miRNA genes. *Genes Dev* **25:** 1589–1594.

Zhao P, Zhang Z, Ke H, Yue Y, Xue D. 2014. Oligonucleotide-based targeted gene editing in *C. elegans* via the CRISPR/Cas9 system. *Cell Res* **24:** 247–250.

Zisoulis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, Pasquinelli AE, Yeo GW. 2010. Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat Struct Mol Biol* **17:** 173–179.