nature structural & molecular biology

# An alternative mode of microRNA target recognition

Sung Wook Chi[1–4], Gregory J Hannon[2] & Robert B Darnell[1]

**MicroRNAs (miRNAs) regulate mRNA targets through perfect pairing with their seed region (positions 2–7). Recently, a precise genome-wide map of miRNA interaction sites in mouse brain was generated by high-throughput sequencing and analysis of clusters of ~50-nucleotide mRNA tags cross-linked to Argonaute (Ago HITS-CLIP). By analyzing Ago HITS-CLIP 'orphan clusters'—Ago binding regions from HITS-CLIP that cannot be explained by canonical seed matches—we have now identified an alternative binding mode used by miRNAs. Specifically, G-bulge sites (positions 5–6) are often bound and regulated by miR-124 in brain. More generally, bulged sites comprise ≥15% of all Ago-miRNA interactions in mouse brain and are evolutionarily conserved. We call position 6 the 'pivot' nucleotide and suggest a model in which a transitional 'nucleation bulge' leads to functional bulge mRNA-miRNA interactions, expanding the number of potential miRNA regulatory sites.**



Phenotypic diversity generated through the regulation of RNA transcripts has been proposed as an explanation for the discrepancy between organismal complexity and the relatively limited number of primary coding transcripts[1,2]. Regulation by miRNAs is one such mechanism, as each miRNA directly binds hundreds of mRNAs to induce post-transcriptional repression, leading to diverse cellular phenotypes[3,4]. The best characterized features determining miRNA-target recognition are short, six-nucleotide (nt) seed sites, which perfectly complement the 5′ end of the miRNA (positions 2–7)[5]. So-called 'seed pairing rules' are widely used to predict functional miRNA target sites, often in combination with evolutionary conservation[6], secondary structure[7] or neighboring context information[8].

Seed rules have been informative in identifying targets through miRNA overexpression or knockdown studies, especially in combination with microarray[9] or proteomic approaches[10,11]. However, such strategies suffer from both false-positive (~40–66%)[10,12] and false-negative predictions (~50–70%)[10,11,13] and cannot identify noncanonical target sites. Several biological studies have functionally validated the hypothesis that perfectly matched miRNA seeds are neither necessary nor sufficient for all functional miRNA-target interactions[14–17]. For instance, genetically verified lin-4 (ref. 14), let-7 (ref. 15) or lys-6 (ref. 17) targets in *Caenorhabditis elegans* as well as some mouse miRNA targets[16] contain only imperfect binding sites, with bulges or G·U wobble pairs in the seed region. Although these studies strongly suggest the existence of noncanonical target sites, these sites have not been recognized as general features of miRNA-mRNA interactions, partly due to the difficulty in determining how frequently such atypical sites are used *in vivo* and what general rules could be used to predict them. False-negative predictions from seed rules could be interpreted as target transcripts with noncanonical target sites, especially for false-negative predictions derived from coimmunoprecipitation of the RNA–Ago complex (~21–50%)[13,18,19], although uncertainty regarding the specificity and resolution of these experiments has made it difficult to interpret and identify such noncanonical sites.

Recently, a precise, genome-wide map of miRNA-binding sites in mouse brain was decoded by applying a biochemical method that involves direct recovery of CLIP[20] containing Ago, followed by HITS of the isolated RNA interaction sites[21,22] and bioinformatic analysis of the sequences based on seed matches[23]. Ago HITS-CLIP has been used to map Ago-mRNA binding footprints (~45–62 nt) with high specificity (~93% specificity, ~13–27% false positives and ~15–25% false negatives)[23] and has also been applied to *C. elegans*[24] and cultured cells[25,26], allowing the identification of *in vivo* miRNA binding sites on a genome-wide scale. However, not all identified Ago binding sites follow classical seed rules. In fact, 27% of Ago-mRNA clusters (normalized overlapping tags from Ago cross-linked mRNAs) in mouse brain are orphans with no predicted seed matches among the top Ago-bound miRNAs (~88%, from the top 20 miRNA families). These results suggest that a substantial number of Ago–miRNA binding sites *in vivo* might fail to follow the rule that mRNA recognition is dictated by canonical seed matches[23].

To search for such noncanonical seed pairing, we analyzed Ago HITS-CLIP orphan clusters. mRNAs harboring G-bulge sites were often bound by miR-124 in mouse brain, and they conformed to a rule in which formation of a transitional nucleation bulge is determined by the annealing of five consecutive nucleotides in positions 2–6 (a 'pivot pairing' rule). Applying this rule globally, we found that bulged sites are common, comprising >15% (>1,441 sites) of all Ago-miRNA interactions in mouse brain, thereby expanding the number of potential regulatory sites for miRNAs and providing insight into the biochemical mechanisms by which miRNA-Ago complexes bind their targets.

[1]Laboratory of Neuro-Oncology, The Rockefeller University, Howard Hughes Medical Institute, New York, New York, USA. [2]Cold Spring Harbor Laboratory, Watson School of Biological Sciences, Howard Hughes Medical Institute, Cold Spring Harbor, New York, USA. [3]Graduate School, Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul, Korea. [4]Samsung Research Institute for Future Medicine, Samsung Medical Center, Seoul, Korea. Correspondence should be addressed to S.W.C. (swchi@skku.edu) and R.B.D. (darnelr@rockefeller.edu).

Received 26 April 2011; accepted 19 December 2011; published online 12 February 2012; doi:10.1038/nsmb.2230

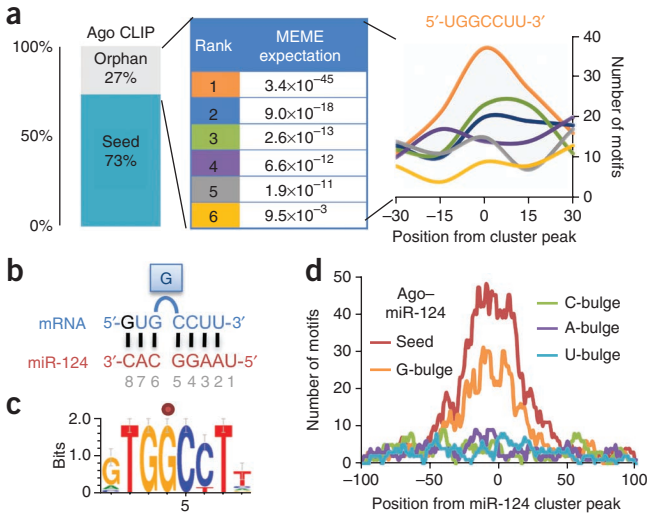


**Figure 1** Identification of G-bulge sites pairing to miR-124 by Ago HITS-CLIP analysis. (**a**) Over-represented motifs in orphan clusters, in which Ago footprint regions have no predicted seed matches among the top 20 Ago–miRNAs families (left panel, see also **Supplementary Fig. 1a**). Six significantly enriched motifs (*E*-value < 0.01) were identified in the orphan Ago footprint regions by MEME analysis (**Supplementary Fig. 1b**), and their *E*-values (MEME expectation) are indicated in the middle panel. Right, distribution of the motifs relative to the peaks of Ago mRNA clusters, with the same colors as in the middle panel. The 5-UGGCCUU-3′ (orange line) motif is the most significantly enriched motif near the peaks ($k = 2.2$ versus $k = 1.8$ in the uniform distribution). (**b**) The 5′-UGGCCUU-3′ (blue) is a G-bulge (in positions 5–6) match to miR-124 (positions 2–8, red). The bulge nucleotide (between positions 5 and 6) is highlighted. (**c**) MEME analysis of the G-bulge motif in 2,392 orphan *de novo* Ago miR-124 clusters, after miR-124 transfection into HeLa cells[23]. The relative height of the individual bases represents the frequency; error bars indicate an approximate, Bayesian 95% confidence interval. (**d**) The position of seed (red) and four possible bulge matches (G, orange; C, green; A, purple; U, blue) to miR-124 are plotted relative to the peak of 3,083 *de novo* Ago miR-124 clusters (normalized, biological complexity ≥ 2). Of note, *de novo* miR-124 clusters could be generated as the consequence of miR-124 overexpression that reduced or precluded Ago binding to sites occupied in untransfected cells, as previously observed[23,28].

## RESULTS

### Identification of G-bulge sites that pair to miR-124

To uncover new rules for miRNA binding, we first conducted an unbiased search for all sequence motifs 6- to 8-nt long that were enriched within orphan clusters detected in mouse brain Ago HITS-CLIP, by using MEME[27] (**Fig. 1a** and **Supplementary Fig. 1b**). Six motifs were identified (*E*-value < 0.01), and their locations relative to cluster peaks[23] were further examined. The 5′-UGGCCUU-3′ sequence was identified as the most significant motif that was enriched (**Fig. 1a**). This motif is a match for miR-124, but only if a G-bulge in the mRNA is allowed (**Fig. 1b**). miR-124 is a brain-specific miRNA previously shown to be the eighth most frequent miRNA associated with Ago, and its canonical seed sites are the most enriched in Ago mRNA clusters that have been identified in mouse brain[23]. Of note, among all 11,463 Ago-mRNA clusters[23], 684 clusters are miR-124 G-bulge clusters (~6% of total clusters, and these are enriched relative to the fraction of miR-124 (~3% among all brain miRNAs bound to Ago) and equal to ~46% of the miR-124 seed clusters (1,480 clusters)). Further analysis revealed that the same G-bulge motif was present in *de novo* Ago miR-124 clusters, those clusters that appear after miR-124 transfection of HeLa cells (**Fig. 1c**), which is consistent with this motif acting as a bona fide miR-124–dependent Ago binding site. G-bulge sites were enriched in *de novo* Ago miR-124 footprints, compared to seed sites (418 G-bulge and 691 seed clusters were identified as 6-mers from positions 1–8 in a 62-nt sequence). Among all possible bulge types, only G-bulge sites were significantly enriched in *de novo* Ago miR-124 clusters (**Fig. 1d**, kurtosis ($k$) = 4.02 in G-bulges versus $k = 1.82$ in C-bulges, $k = 2.23$ in A-bulges and $k = 2.25$ in U-bulges; the distribution of G-bulges is the same as in seed sites ($k = 4.38$)), and this was also found to be the case in several additional analyses evaluating all possible bulge and wobble sites in Ago clusters (**Supplementary Table 1** and **Supplementary Fig. 1**). Considered together, these data provide direct evidence that miR-124 may act on brain transcripts harboring noncanonical bulge sites and that such interactions are specific for G-bulges relative to other nucleotide bulges.

### Validation of G-bulge sites in Ago–miR-124 clusters

To measure the extent of G-bulge binding to miR-124 *in vivo*, we analyzed a set of Ago–miR-124 bulge clusters detected from both brain and miR-124 transfected HeLa cells (**Supplementary Fig. 2a**); such sites are identified by a markedly stringent test, in that they denote native mouse brain Ago-mRNA footprints that are conserved in HeLa cells, referred to as conserved *de novo* Ago clusters. From this dataset, we conservatively estimated (**Fig. 2a**) that one fourth of all Ago–miR-124–mRNA interactions are mediated through binding to bulge sites, with three-quarters of these attributable to G-bulge sites (~18% of total). We used luciferase reporter assays to validate G-bulge sites in *Mink1* and *epb41* (**Fig. 2b** and **Supplementary Fig. 2b,c**). Generation of point mutants demonstrated that in these contexts, G-bulge sites were able to mediate miR-124–dependent repression (**Fig. 2c,d**) as efficiently as canonical seed sites. We then extended these validation analyses to a large number of mRNAs (**Fig. 2e** and **Supplementary Fig. 3a**) that show miR-124–dependent changes in transcript levels after miR-124 transfection ($n = 2{,}694$, $P < 0.05$)[9]. In analysis of the cumulative distribution of miR-124–dependent changes, transcripts with G-bulge sites identified in *de novo* Ago–miR-124 HeLa clusters were downregulated ($P < 0.05$, relative to the distribution of total miR-124–dependent transcripts, Kolmogorov-Smirnov test (KS-test)), although less so than the transcripts with seed sites in the clusters ($P < 0.01$, KS-test). Notably, transcripts with G-bulge sites were also downregulated relative to the total number of Ago–miR-124–bound transcripts (**Supplementary Fig. 3a**), showing that the observed repression was a G-bulge specific effect. To compare the extent of repression mediated by perfect and bulged seed matches for miR-124, the downregulated transcripts were examined further, and this analysis indicated that G-bulge sites were present in an abundance that was comparable to that of seed sites, whereas possible binding sites with other types of bulges were not as abundant (**Fig. 2f**). The degree of repression mediated by G-bulges versus seed sites substantially overlapped that seen for canonical sites (**Supplementary Fig. 3b**). In addition, transcripts with G-bulge sites were also downregulated at the protein level[10], although this was only evident in the smaller number of transcripts that had larger changes (**Supplementary Fig. 3c**). We also found evidence that nucleotide changes in G-bulge sequences could abrogate interactions with Ago *in vivo* (**Supplementary Fig. 3d**) by examining species-specific nucleotide changes between two Ago HITS-CLIP datasets (HeLa versus mouse brain, $P < 0.01$, Fisher's exact test,

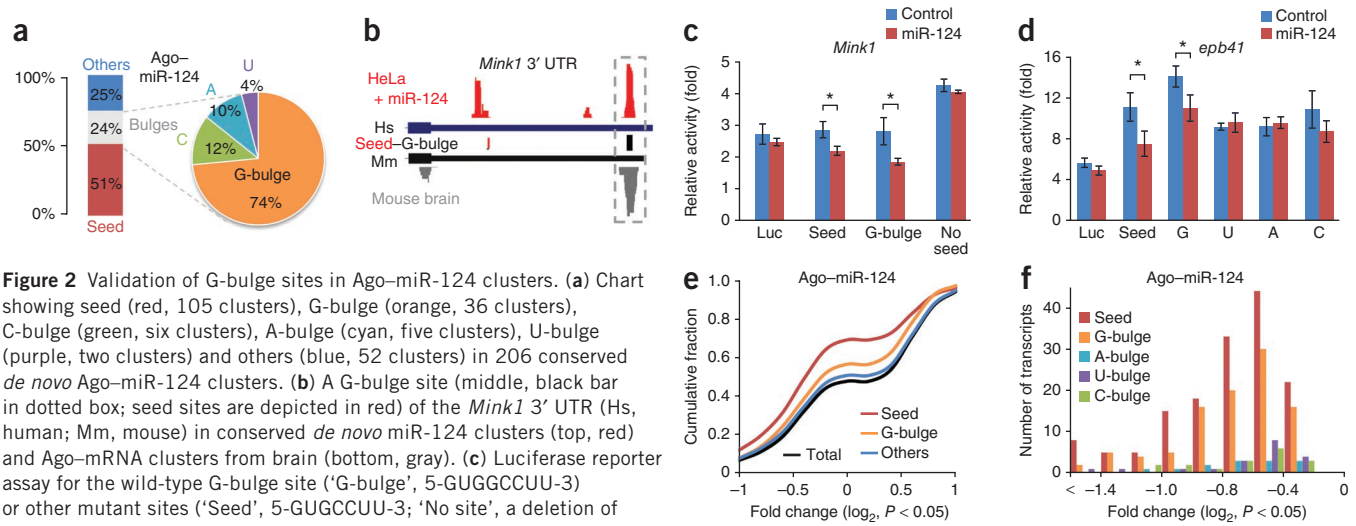


**Figure 2** Validation of G-bulge sites in Ago–miR-124 clusters. (**a**) Chart showing seed (red, 105 clusters), G-bulge (orange, 36 clusters), C-bulge (green, six clusters), A-bulge (cyan, five clusters), U-bulge (purple, two clusters) and others (blue, 52 clusters) in 206 conserved *de novo* Ago–miR-124 clusters. (**b**) A G-bulge site (middle, black bar in dotted box; seed sites are depicted in red) of the *Mink1* 3′ UTR (Hs, human; Mm, mouse) in conserved *de novo* miR-124 clusters (top, red) and Ago–mRNA clusters from brain (bottom, gray). (**c**) Luciferase reporter assay for the wild-type G-bulge site ('G-bulge', 5-GUGGCCUU-3) or other mutant sites ('Seed', 5-GUGCCUU-3; 'No site', a deletion of 5-GUGGCCUU-3; 'Luc', luciferase vector with no insert) of *Mink1* in the presence of a miR-124 (red bar) or a control miRNA mimic (blue bar). Relative activity, average *Renilla* luciferase activity normalized to firefly luciferase in three replicates; error bars, s.d. Asterisks denote instances of $P < 0.01$ (*t*-test). (**d**) The same luciferase reporter assay as in **c** but done for a G-bulge site in *epb41* with more single-point mutations in bulged position. Asterisks denote instances of $P < 0.05$ (*t*-test). (**e**) Transcripts with G-bulge sites in *de novo* Ago–miR-124 clusters (orange line) showed miR-124–dependent suppression relative to previous analysis of all regulated transcripts in miR-124 transfected HeLa cells[9]. (**f**) Numbers of miR-124–dependent transcripts containing seed (150 transcripts), G-bulge (98 transcripts) and other bulge sequences (U,19 transcripts; A,13 transcripts; and C,17 transcripts) in *de novo* Ago–miR-124 clusters are plotted in all ranges of repression.

**Supplementary Fig. 3e**). Overall, these data lead us to conclude that G-bulge sites bound by miR-124 are functional.

### An alternative mode of miRNA target recognition

We explored the mechanisms that might account for the preference for G-bulges in miR-124 orphan clusters. The free energy calculation for G-, C-, U- or A-bulges in positions 5–6 was the same (**Fig. 3a** and **Supplementary Fig. 4a**) and hence did not explain the specificity for G. However, when we considered an intermediate in the miR-124–mRNA duplex in which the G-bulge nucleotide was temporarily used to bind to miR-124, five consecutive nucleotides were available for annealing (positions 2–6) compared to only four consecutive nucleotides (positions 2–5) for all other possible bulges, and this intermediate yielded a free energy that was substantially lower than that seen with only four consecutive annealing nucleotides (**Fig. 3a** and **Supplementary Fig. 4b**). This intermediate (termed the transitional nucleation) may occur *in vivo*, such that stability of the transition state would be mainly determined by pairing in miRNA position 6.

To test the validity of such a pivot-pairing rule, we initially examined whether bulge sites with this pattern (resulting in a nucleation bulge, in which the bulge sequence was competent to pair to the miRNA pivot nucleotide in a transitional nucleation intermediate, **Fig. 3a**) were enriched in Ago footprint regions. In an analysis of the top 30 Ago-bound miRNAs (~92% of all miRNAs bound to Ago), nucleation bulges were slightly but significantly enriched near Ago-mRNA cluster peaks ($k = 1.92$, within −50 and +50 nt, $P < 0.05$, relative to background uniform distribution ($k = 1.80$), KS-test. **Fig. 3b**), as is also seen with canonical seed sites ($k = 2.33$, $P < 0.01$, KS-test)[23]. Such an enrichment was not seen with bulge sites harboring only 4-nt transitional interactions ($k = 1.81$, $P = 0.95$, KS-test, which is the same as the background uniform distribution ($k = 1.80$)), termed 'non-nucleation bulges' (in which the position 5–6 bulge sequence was identical to the pivot nucleotide and therefore unable to pair to the pivot, **Fig. 3a** and **Supplementary Fig. 4b**).

### Nucleation bulges are conserved, functional and widespread

Evolutionary conservation of miRNA binding sites can provide strong evidence for their biological significance. Comparison of the conservation of canonical seed sites and nucleation bulges revealed that within the context of otherwise poorly conserved 3′ UTRs (median conservation rate ~5.4%, 7-mers), nucleation bulges for all known miRNAs were significantly conserved over the background (~6.3 versus 5.4%, $P < 0.01$, KS-test) as well as seed sites (~6.6 versus 5.4%, $P < 0.01$, KS-test) and were distributed to the same degree as conservation rates observed in seed sites (~6.3 versus 6.6%, respectively, **Fig. 3c**), whereas non-nucleation bulges were not conserved and were similar to background distribution (~5.1 versus 5.4%, respectively). Among all Ago footprints in mouse brain (average conservation rate ~25%, 7-mers), nucleation bulges for the top 20 Ago-bound miRNAs (~88% of all miRNAs bound to Ago) were also evolutionarily conserved over background rates (~32% versus 25%, $P = 0.06$, *t*-test, **Fig. 3d**), but less so than canonical seed sites (~51% versus 25%, $P = 0.02$, *t*-test); non-nucleation bulges were even less conserved than the background rate (~22% versus 25%, $P = 4.0 \times 10^{-5}$, *t*-test). When non-nucleation bulges were used as a control, nucleation bulges were also significantly conserved (~32% versus 22%, $P = 0.01$, *t*-test) but less so than in canonical seed sites (~51% versus 22%, $P = 1.9 \times 10^{-7}$, *t*-test). In addition, the 3′ UTR and coding sequence were analyzed separately, with similar results (**Supplementary Fig. 4c**).

We further investigated the functional significance of the nucleation bulge by compiling brain-expressed transcripts from a meta-analysis of microarray experiments conducted with seven different miRNAs transfected into cell lines[28] (excluding the previous analysis of miR-124). Transcripts harboring nucleation bulges in Ago clusters showed significant miRNA-dependent repression ($P = 0.03$, KS-test, **Fig. 3e**); although this was to a lower degree than miRNA-dependent repression of transcripts with canonical seed sites ($P = 1.05 \times 10^{-24}$, KS-test), transcripts with non-nucleation bulges showed no repression ($P = 0.95$, KS-test). Repression mediated by sites harboring a nucleation bulge showed a broad range of fold changes, overlapping

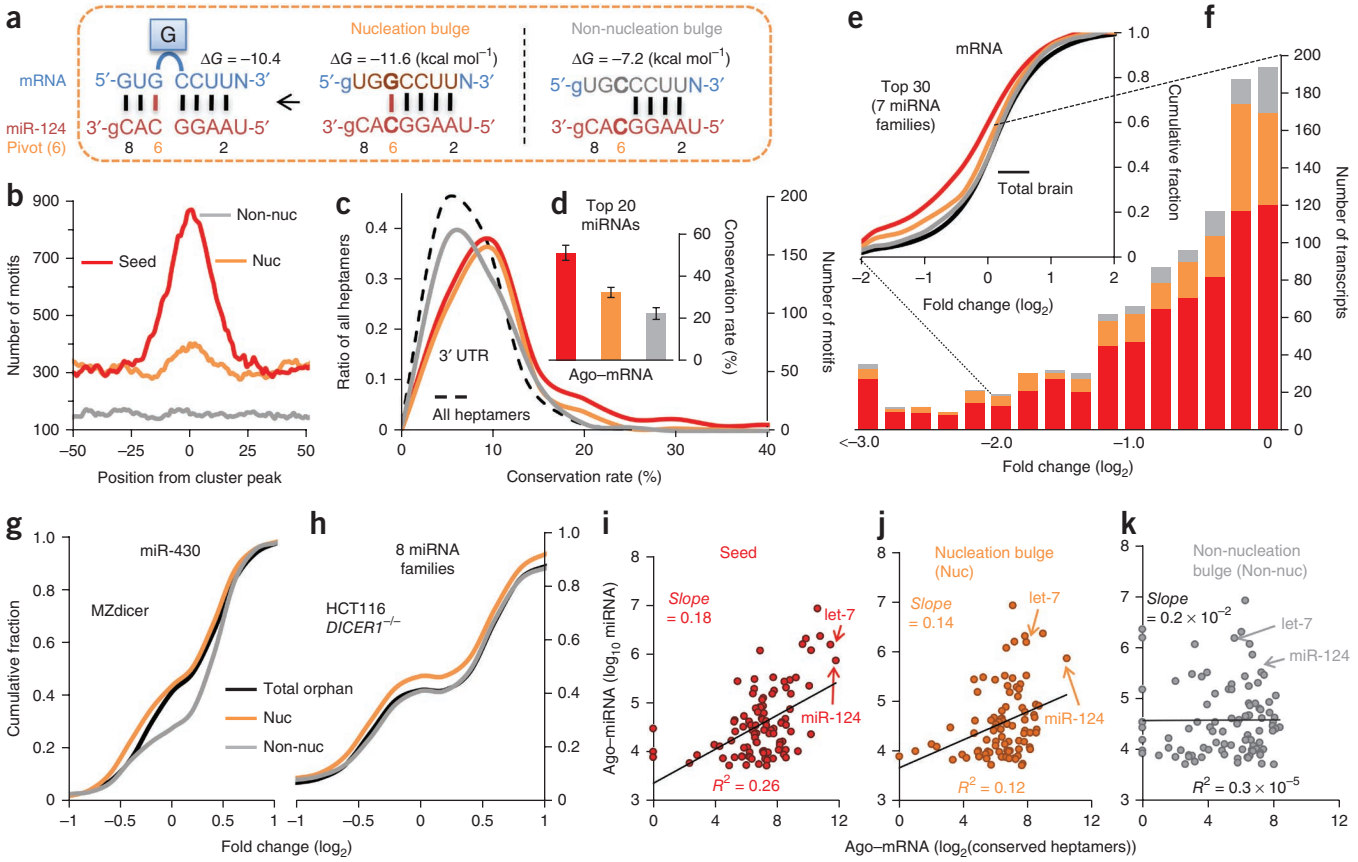


**Figure 3** Nucleation bulges are widely used and evolutionarily conserved as functional miRNA target sites. (**a**) Minimum free energy (Δ*G*) calculation for miR-124 pairing to a G-bulge site (left) and a transition nucleation model, showing a nucleation bulge (middle) with pivot pairing (bold text) and a non-nucleation bulge (right) with no pivot pairing (bold text). (**b**) The position of seed (Seed), nucleation bulge (Nuc) and non-nucleation bulge sites (Non-nuc) from the top 30 Ago-bound brain miRNAs plotted relative to the peak of 11,463 Ago footprint regions (biological complexity ≥ 2). (**c**) Distribution of conservation rates for three different sites (Seed, Nuc and Non-nuc) of mouse miRNAs and all heptamers (dotted line) in 3′ UTRs. (**d**) Average conservation rates of the three different sites (as in **c**) for the top 20 Ago-bound miRNAs were calculated in the Ago footprint regions; error bars indicate standard error. (**e**) Meta-analysis of brain-expressed transcripts harboring the three different sites (as in **c**) in compiled microarray data[28], using seven different miRNA transfections. **c**–**e** are colored as in **b**. (**f**) Numbers of the seven different miRNA-dependent transcripts containing the three different sites analyzed in **e** plotted in all ranges of repression. (**g**) Cumulative distribution of miR-430 orphan transcripts. (**h**) Meta-analysis of orphan transcripts for eight miRNA families. (**i**,**j**) Linear regression analysis comparing the number of conserved miRNA seed matches (7-mers to positions 2–8) (**i**) or nucleation bulges (7-mers) (**j**) with the frequency of the top 100 miRNAs experimentally determined by mouse brain Ago HITS-CLIP. (**k**) Non-nucleation bulges analyzed by linear regression (7-mers).

with and similar to the degree of repression through canonical seed-match sites (−0.79 versus −0.97 median $\log_2$ fold changes; **Fig. 3f** and **Supplementary Fig. 4d**).

We also expanded our analysis to a second organism by using recently available Ago HITS-CLIP and microarray data from the wild-type versus the *alg-1* (gk214) mutant *C. elegans*[24]. In this dataset, the degree of transcript repression from nucleation bulge and canonical seed sites also overlapped (**Supplementary Fig. 4e**–**g**). The same enrichment of functional nucleation bulges (**Supplementary Fig. 5a**–**c**) was also seen in *de novo* Ago–miR-124 or Ago–miR-7 clusters identified by photoactivatable ribonucleoside-enhanced CLIP (PAR-CLIP)[25]. However, nucleation G-bulge sites were more specifically enriched in *de novo* Ago–miR-124 clusters identified by HITS-CLIP than by PAR-CLIP (**Supplementary Fig. 5d**).

We tested whether the pivot-pairing rule could explain 'orphan transcripts' evident in previously published datasets in which miRNAs had been exogenously added to Dicer-null cells. These orphan transcripts were defined as those showing miRNA-dependent changes, but in none of these could 6-mer seed matches for the added miRNA be found (similarly to false-positive predictions from the seed rule). In a cumulative distribution analysis for miR-430–injected Dicer-null zebrafish (MZdicer, **Fig. 3g**)[29] or in a meta-analysis of data generated for eight different miRNAs in Dicer-null cells (HCT116 *DICER1*$^{-/-}$, **Fig. 3h**)[30], orphan transcripts with predicted nucleation bulges were significantly downregulated ($P < 0.05$ and $P = 0.005$, KS-test) relative to total orphan transcripts. Of note, total orphan transcripts (**Fig. 3g**,**h**) showed an increase in fold change, probably caused by the subtraction of transcripts with seed matches, which show higher fold decreases.

We examined the correlation between the number of conserved bulge sites predicted by the pivot rule to be present in 11,463 Ago footprints and the number of miRNAs associated with such bulge sites, using linear regression analysis. Nucleation bulge sites for the top 100 Ago-bound miRNAs showed positive correlation (**Fig. 3i**), as did canonical seed sites (**Fig. 3j**), although the correlation was not as high as for canonical seed sites. However, non-nucleation bulge sites showed no correlation (**Fig. 3k**). Notably, we also confirmed that 5-mer seed matches (positions 2–6), which could confound the result

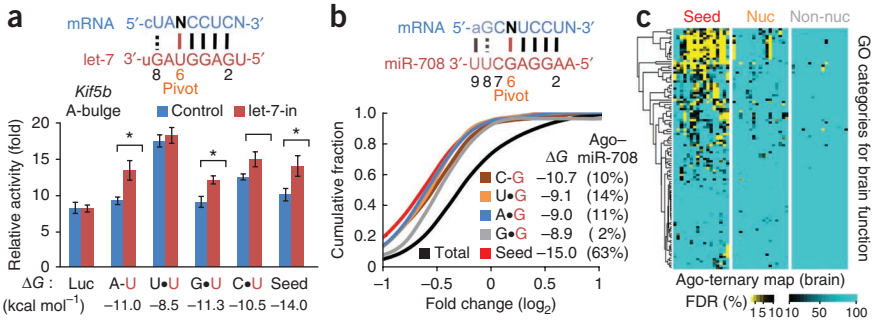


**Figure 4** Functional nucleation bulges in let-7 and miR-708, and gene ontology (GO) analysis. (**a**) Transitional nucleation model of let-7 (top) and luciferase reporter assay (bottom) for the A-bulge site in *Kif5b* (**Supplementary Fig. 6**). Luciferase assays were done as described in **Figure 2d**, except reporter activities were compared between let-7-inhibitor transfected (let-7-in) and control transfected HeLa cells (Control), to inhibit high endogenous levels of HeLa let-7. Asterisks denote instances of $P < 0.01$ (*t*-test). Bracket without asterisk indicates $P < 0.05$ (*t*-test). (**b**) A transition nucleation state model of miR-708 (top) and free energies of transitional nucleation ($\Delta G$, bottom). The frequencies of seed and different bulge sites identified in *de novo* clusters from miR-708 transfected HeLa cells are also indicated in the lower panel as percentages; composition of seed and bulge motifs in the Ago footprint region (64 nt): seed (red, 864 clusters), C-bulge (brown, 143 clusters), U-bulge (orange, 185 clusters), A-bulge (cyan, 153 clusters) and G-bulge (gray, 28 clusters). Cumulative distribution of transcripts with canonical seeds or each type of bulge (colored as indicated) in conserved *de novo* Ago miR-708 clusters are shown with all transcripts (black) in the lower panel. (**c**) Heat maps derived from gene ontology analysis of target transcripts for each of the top 20 miRNAs based on Watson-Crick pivot pairing versus seed rules (Seed, seed sites; Nuc, nucleation bulge; Non-nuc, non-nucleation bulge) show a false discovery rate. The tree indicates the hierarchical clustering of gene ontology subcategories for brain function.

from bulge sites, are not functional, as we observed that those could not interact with Ago–miRNA (**Supplementary Fig. 5e**), nor could they mediate miRNA-dependent repression *in vivo* (**Supplementary Fig. 5f**). Considered as a whole, our analyses demonstrate that nucleation bulges are biologically significant, as they are evolutionarily conserved, functional and widespread.

### Application and validation of pivot pairing rule

To define and apply the pivot-pairing rule with more precision, we searched for nucleation bulges in orphan clusters predicted to have let-7–bulge mRNA interactions, as let-7 is a well-studied miRNA and is highly expressed (ranked no. 5 in Ago–miRNA binding) with canonical seed sites enriched in brain (second only to miR-124 in Ago-mRNA clusters)[23]. Notably, let-7 has a pivot nucleotide (U) capable of wobble pairing, and in theory, this might contribute transitional nucleation energy to either A or G residues. We identified a let-7 nucleation A-bulge in a *Kif5b* orphan cluster (**Supplementary Fig. 6a**) and found that it mediated a level of repression comparable to a canonical seed site in luciferase reporter assays (**Fig. 4a**). We noted that the extent of repression by let-7 was roughly proportional to the absolute value of free energy from transitional nucleation ($|\Delta G|$); a mutated G-bulge site with similar $|\Delta G|$ showed let-7–dependent repression comparable to that seen with the native A-bulge, probably due to G·U wobble pivot pairing. Even a mutated C-bulge site showed some repression, possibly through noncanonical Watson-Crick pivot pairing with extended extra G·U pairing at position 8, whereas a U-bulge site (a non-nucleation bulge) was not functional.

To extend this observation, we also examined miR-708, which has a G pivot nucleotide and also has the potential for extended transitional nucleation pairing from the pivot onward (positions 6–9), including the possibility of G·U pairing in three of four nucleotides and potentially more promiscuous bulge partners (**Fig. 4b**, upper). Therefore, stable nucleation states of miR-708 are predicted by five consecutive base pairs at the 5′ end of miR-708 (with a G pivot nucleotide at position 6, as indicated) as well as potential additional G·U and/or A·U pairing in positions 8–9, predicting that in addition to C-bulges, substitution of other bulge nucleotides might result in similar transitional nucleation stability (inset in **Fig. 4b**, lower). We addressed miR-708 bulge interactions by analyzing Ago HITS-CLIP and microarray data, following miR-708 transfection of HeLa cells (**Supplementary Fig. 6b,c**). From among 3,433 *de novo* Ago–miR-708 clusters with predicted target sites, 63% matched the canonical miR-708 seed sequence, whereas 37% had bulge sites. All four possible bulge sites were detected, in rough proportion to the predicted free energy of binding to the G-pivot nucleotide in the nucleation state (C: 10%, U: 14%, A: 11% and G: 2%, **Fig. 4b**), although we recognize that factors other than free energy predictions are likely to contribute to the efficiency of nucleation (see discussion below). The 3,433 *de novo* Ago–miR-708 clusters with predicted target sites were used for this analysis, as a relatively small number of transcripts with Ago–miR-708 clusters in the brain were also expressed in HeLa cells (140, versus 423 observed for miR-124). Although the numbers were small, we also observed a similar frequency of seed (72%) and bulge predictions (C: 6%, U: 15%, A: 6% and G: 0%) in the 47 conserved *de novo* Ago miR-708 clusters (overlapping set of *de novo* miR-708 HeLa cluster and p13 mouse clusters) as in the *de novo* cluster analysis alone. Notably, transcripts harboring these *de novo* clusters were significantly downregulated after miR-708 transfection (all $P \leq 0.01$ relative to total transcripts, KS-test), and again, G-bulge sites were the least effective in miR-708–dependent repression ($P = 0.01$, KS-test), correlating with the lowest $|\Delta G|$ for G·G interactions. These experiments demonstrate that miR-708 transfection induces Ago-mRNA binding clusters that include bulge sites enabling pivot nucleotide interactions.

Finally, we applied the pivot-pairing rule to the previous Ago ternary map established for the 20 most abundant Ago–miRNAs in the brain. This analysis identified 1,441 clusters decoded as nucleation bulge sites (2,162 clusters when G·U wobble pivot pairing was considered, ~15–22% of Ago–miRNA-mRNA interactions). Using gene ontology analysis, we examined the functions encoded by these additional targets and found that these targets were also enriched in brain-function gene ontology categories, less so than Ago-mRNA targets with canonical seed sites but more than control, non-nucleation bulge targets (**Fig. 4c** and **Supplementary Fig. 7a**), suggesting that neuronal functions are also regulated through interactions between nucleation bulges and Ago–miRNAs. These results demonstrate that the application of the pivot-pairing rule and identification of miRNA bulge interactions can expand the Ago-ternary map and our understanding of the biological actions mediated by miRNAs.

## DISCUSSION

Previous work demonstrated that most miRNA targets are mediated by seed pairing[6,9–11]. Although early biological studies demonstrated several instances in which bulged sequences are functional[14–17], they have not been recognized as general features

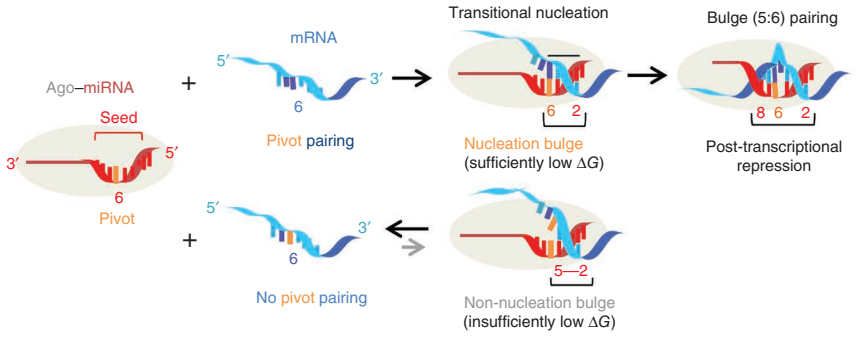


**Figure 5** Pivot pairing and transitional nucleation models. Nucleation bulges enabling pivot pairing (upper) induce transitional nucleation (five consecutive pairings in positions 2–6) and initiate formation of the miRNA-mRNA duplex by stabilizing the thermodynamics (for example, –11.6 kcal mol$^{-1}$ for miR-124). This transition state is followed by the formation of a bulge (positions 5–6) and propagation of base-pairing distally, leading ultimately to post-transcriptional repression by Ago–miRNA. Non-nucleation bulges without pivot pairing only make four consecutive matches and are probably not stable enough for nucleation to initiate formation of the miRNA–mRNA duplex (for example, –7.2 kcal mol$^{-1}$ for miR-124, lower). Oval shaded areas represent Ago protein.

of miRNA-mRNA interactions. Recently, Ago HITS-CLIP has been applied to the mouse brain[23], to *C. elegans*[24] and to cultured cells[25,26], allowing the identification of *in vivo* miRNA binding sites on a genome-wide scale. The ability to map Ago-mRNA binding footprints with high resolution (~45-nt footprints, with ~93% specificity)[23] and the identification of Ago-mRNA orphan clusters motivated us to analyze the genome-wide use of atypical miRNA sites. We observed that Ago binds to a large number of bulged sites *in vivo*, expanding on observations made of individual bulged seed sites[14–17]. Notably, from these *in vivo* miRNA-mRNA interactions, we demonstrate that nucleation bulges occur at a specific position in the mRNA seed site (positions 5–6) when that nucleotide is able to bind sufficiently and robustly with the miRNA pivot nucleotide (position 6). On average, the degree of Ago interaction and repression seen with nucleation bulges is somewhat less than what is seen with canonical seed sites, but is statistically significant in all experiments and analyses and is similar to that seen in a number of suppression studies of canonical seed sites ($P < 0.05$)[31–33]. The pivot-pairing rule provides a qualitative change in our understanding and assessment of miRNA-mRNA regulation, and it is a mechanism that is conserved across a range of species, from *C. elegans* to mammals.

We propose a transitional nucleation model in which the pivot-bulge interaction serves as a general means of enabling a transitional nucleation state by stabilizing nucleation base-pairing (positions 2–6), allowing subsequent bulge formation and propagation of the seed interaction (**Fig. 5**). This model is supported by structural studies of Ago[34–37]. In the structure of the Ago–miRNA binary complex, the same residues (positions 2–6) identified here are exposed, enabling the initiation of nucleation for mRNA pairing (other nucleotides in seed regions were embedded (position 1) or partially buried (positions 7–8))[36]. In the Ago ternary complex, bulges at positions 4–5 and 5–6 in the mRNA, but not in the miRNA, can function *in vitro*[35], and nucleation and subsequent propagation is believed to be necessary for cleavage of perfectly matched miRNA–target duplexes[37]. Notably, our nucleation model was presaged by a similar hypothesis in which four nucleotides within the seed (positions 1–6) were used to calculate the initiation potential of miRNA-mRNA interactions, in efforts to improve the prediction of target sites by considering secondary structure[7]. Although the nucleation potential calculated from Δ*G*s when only four nucleotides were used was not able to explain the rules of miRNA binding in Ago orphan clusters, these prior studies provide conceptual support for our findings[7,18]. Here we define the rules for such pairing and find that they frequently occur, as detected by Ago HITS-CLIP in brain (≥15% of Ago–miRNA-mRNA interactions among the top 20 Ago-bound miRNAs) or miRNA transfected HeLa cells (~7–18% or one-third of seed sites estimated from brain-conserved *de novo* Ago–miR-708 or Ago–miR-124 clusters, respectively), demonstrating that they are widely used for miRNA-mRNA interactions *in vivo*.

Our results also provide evidence that non-Watson-Crick pivot pairings can support bulge formation if they can contribute sufficient Δ*G* to stabilize the transitional nucleation state. However, non-Watson-Crick base-pairings are incompletely understood[38], such that accurate calculation of Δ*G* to precisely predict nucleation bulges for all miRNAs is challenging. Moreover, such calculations are likely to serve only as approximate guides to more complex biology. For example, formation of a tethered complex between a Piwi protein from *Archaeoglobus fulgidus*, such as Ago, which contains a MID-PIWI domain, and a short RNA had a higher seed-to-target binding affinity than did the short RNA in isolation (up to ~300-fold enhancement)[39]. Thus, the role of Ago in stabilizing miRNA pivot pairing might contribute more substantially to nucleation than is estimated on the basis of free energy calculations alone. For this reason, for the target predictions done in this study, we limited our analysis to Watson-Crick pivot pairings. Although this enabled us to decode new functional sites comprising ~15–22% of Ago–miRNA-mRNA interactions, it may also be of interest to examine non-Watson-Crick pivot pairing.

We noted that bulge sites include a potential 5-nt seed sequence (positions 2–6). Therefore, in our all analyses, we used 7-mer motifs to search for nucleation bulges, which differentiate longer bulge sites, from 5-nt seeds. We also confirmed that sites harboring only 5-mer seed matches (positions 2–6) are not functional (**Supplementary Fig. 5e**,**f**). It is possible that some 6-mer seed matches to positions 1–6 might function as 5-mer sites, as there is some indirect[6] and *in vitro* evidence[39] that the first position of a miRNA cannot form a base pair. However, in our previous Ago HITS-CLIP study[23], we were able to address this question *in vivo* and observed that all 6-mers matching to miRNA positions 1–8 were enriched in Ago-mRNA footprints. Further study would be needed to clarify if and when the first position participates in functional miRNA-mRNA interactions.

Although bulge sites account for many of the orphan clusters, unexplained orphans (~25% of the conserved *de novo* Ago miR-124 clusters, **Fig. 2a**) still remain, and these show evidence of being functional (**Fig. 2e**). Although we have not found a consensus motif or free energy change associated with miR-124 binding in these remaining orphans (**Supplementary Fig. 7b**), they may be useful for uncovering additional rules or expanding the study of noncanonical miRNA binding sites, such as functional 'seedless' elements[40] or centered pairing sites[41]. These remaining orphans could include only Ago-mRNA interactions, such as those orphans containing G-rich motifs that were recently discovered in Dicer-null mouse embryo stem cells by conducting Ago HITS-CLIP[26] analysis. Notably, our initial analysis of orphan clusters also identified some as harboring G-rich motifs (**Supplementary Fig. 1b**).

In summary, Ago HITS-CLIP analysis enabled us to identify a new class of miRNA target site nucleation bulges and an alternative mode of miRNA target recognition by a pivot-pairing rule. Based on these findings, we propose a transitional nucleation model in which a transitional nucleation state determines the binding of miRNAs to nucleation bulge mRNAs. The identification of functional noncanonical miRNA-mRNA interactions is essential to understanding the mechanisms of miRNA target recognition, discovering new miRNA targets and applying RNA interference analysis for experimental and clinical purposes, where understanding the specificity of target recognition is especially important.

## METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

*Note: Supplementary information is available on the Nature Structural & Molecular Biology website.*

### ACKNOWLEDGMENTS

We thank the members of the Darnell and Hannon laboratories for helpful discussions. This work was supported in part by grants from the US National Institutes of Health (R.B.D. and G.J.H.) and a grant from the Korean Health Technology R&D Project, Ministry of Health & Welfare (A111989, to S.W.C.). R.B.D. and G.J.H. are investigators of the Howard Hughes Medical Institute.

### AUTHOR CONTRIBUTIONS

S.W.C. carried out the experiments and computational analyses. S.W.C. and R.B.D. analyzed the data. S.W.C., G.J.H. and R.B.D. designed the research and wrote the paper.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Sharp, P.A. The centrality of RNA. *Cell* **136**, 577–580 (2009).
2. Licatalosi, D.D. & Darnell, R.B. RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.* **11**, 75–87 (2010).
3. Ambros, V. The functions of animal microRNAs. *Nature* **431**, 350–355 (2004).
4. He, L. & Hannon, G.J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5**, 522–531 (2004).
5. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
6. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
7. Long, D. *et al.* Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* **14**, 287–294 (2007).
8. Grimson, A. *et al.* MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27**, 91–105 (2007).
9. Lim, L.P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–773 (2005).
10. Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455**, 64–71 (2008).
11. Selbach, M. *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63 (2008).
12. Mourelatos, Z. Small RNAs: the seeds of silence. *Nature* **455**, 44–45 (2008).
13. Easow, G., Teleman, A.A. & Cohen, S.M. Isolation of microRNA targets by miRNP immunopurification. *RNA* **13**, 1198–1204 (2007).
14. Ha, I., Wightman, B. & Ruvkun, G. A bulged lin-4/lin-14 RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. *Genes Dev.* **10**, 3041–3050 (1996).
15. Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K. & Slack, F.J. The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. *Genes Dev.* **18**, 132–137 (2004).
16. Tay, Y., Zhang, J., Thomson, A.M., Lim, B. & Rigoutsos, I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **455**, 1124–1128 (2008).
17. Didiano, D. & Hobert, O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat. Struct. Mol. Biol.* **13**, 849–851 (2006).
18. Hammell, M. *et al.* mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nat. Methods* **5**, 813–819 (2008).
19. Karginov, F.V. *et al.* A biochemical approach to identifying microRNA targets. *Proc. Natl. Acad. Sci. USA* **104**, 19291–19296 (2007).
20. Ule, J. *et al.* CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**, 1212–1215 (2003).
21. Licatalosi, D.D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469 (2008).
22. Darnell, R.B. HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip. Rev. RNA* **1**, 266–286 (2010).
23. Chi, S.W., Zang, J.B., Mele, A. & Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).
24. Zisoulis, D.G. *et al.* Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **17**, 173–179 (2010).
25. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
26. Leung, A.K. *et al.* Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nat. Struct. Mol. Biol.* **18**, 237–244 (2011).
27. Bailey, T.L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).
28. Khan, A.A. *et al.* Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nat. Biotechnol.* **27**, 549–555 (2009).
29. Giraldez, A.J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75–79 (2006).
30. Linsley, P.S. *et al.* Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. *Mol. Cell. Biol.* **27**, 2240–2252 (2007).
31. Kawahara, Y. *et al.* Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**, 1137–1140 (2007).
32. Zhao, Y. *et al.* Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1–2. *Cell* **129**, 303–317 (2007).
33. Gehrke, S., Imai, Y., Sokol, N. & Lu, B. Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression. *Nature* **466**, 637–641 (2010).
34. Song, J.J., Smith, S.K., Hannon, G.J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434–1437 (2004).
35. Wang, Y. *et al.* Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921–926 (2008).
36. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D.J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–213 (2008).
37. Wang, Y. *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754–761 (2009).
38. Tinoco, I. Jr. & Bustamante, C. How RNA folds. *J. Mol. Biol.* **293**, 271–281 (1999).
39. Parker, J.S., Parizotto, E.A., Wang, M., Roe, S.M. & Barford, D. Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol. Cell* **33**, 204–214 (2009).
40. Lal, A. *et al.* miR-24 Inhibits cell proliferation by targeting *E2F2*, *MYC*, and other cell-cycle genes via binding to 'seedless' 3′UTR microRNA recognition elements. *Mol. Cell* **35**, 610–625 (2009).
41. Shin, C. *et al.* Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell* **38**, 789–802 (2010).

## ONLINE METHODS

**Bioinformatics.** In general, for the bioinformatics analysis, we used the University of California, Santa Cruz genome browser (http://genome.ucsc.edu/), Galaxy (http://galaxy.psu.edu/) and Python scripts, as described previously[23]. The minimum free energy ($\Delta G$) of miRNA–mRNA duplexes were calculated using RNAhybrid[42]. To calculate $\Delta G$ of the intermediate duplex (transitional nucleation), a helix constraint (–f, 2–6) was used. All statistical tests were done using Scipy (http://www.scipy.org/), and other analyses were done with the Python scripts, using Biopython (http://biopython.org/). The motif analysis was done using MEME (http://meme.sdsc.edu/) and WebLogo (http://weblogo.threeplusone.com/). Detailed methods for the bioinformatics analysis are described in **Supplementary Methods**.

**Transfection of miRNA mimics and inhibitors.** Transfections of miRNA mimics and inhibitors were carried out as described previously[23]. Briefly, HeLa cells were transfected using Lipofectamine 2000 (Invitrogen) in 100-mm$^2$ plates with 75 nM RNA duplexes, miR-124, miR-708 and negative control miRNA from miRIDIAN miRNA mimic (Dharmacon) or 75 nM miRNA inhibitors to let-7 and negative control inhibitor from miRIDIAN miRNA hairpin inhibitors (Dharmacon).

**Luciferase reporter assay.** Luciferase reporter assays were done according to the manufacturer's protocol for dual-luciferase assays (Promega). A psiCheck-2 plasmid (Promega) was used for the luciferase reporter and cotransfected with miR-124, miR-708, control miRNA mimic, let-7 inhibitor or control miRNA inhibitor (Dharmacon) using Lipofectamine 2000 (Invitrogen). Primers for *Mink1* (forward, 5′-CCGCTCGAGAGCAGCAAGTAACCCTTCTCCTCCCTCCCCCCTCCCCC CTCCTCAATGTAG-3′; seed-reverse, 5′-ATTTGCGGCCGCTTAACAAACAGGATATCCAAGGC ACTACATTGAGGAGGGGGGAGGGG-3′; G-bulge-reverse, 5′-ATTTGCGGCCGCTTAACAAACA GGATATCCAAGGCCACTACATTGAGGAGGGGGGAGGGG-3′; and no seed-reverse, 5′-ATTTGCG GCCGCTTAACAAACAGGATATCCCTACATTGAGGAGGGGGGAGGGG-3′), for *epb41* (forward, 5′-CCGCTCGAGACATGGAAGTTGCTTCAGATATCTGATACTGTGAATGTTTGAA CATATCCG-3′; G-reverse, 5′-ATTTGCGGCCGCAAGGGTAGCTGGAGAGGTGAAGGCCACGG ATATGTTCAAACATTCACA-3′; A-reverse, 5′-ATTTGCGGCCGCAAGGGTAGCTGGAGAGGTGA AGGTCACGGATATGTTCAAACATTCACA-3′; U-reverse, 5′-ATTTGCGGCCGCAAGGGTAGCTG GAGAGGTGAAGGACACGGATATGTTCAAACATTCACA-3′; and G-reverse, 5′-ATTTGCGGCCGCAAGGGTAGCTGGAGAGGTGAAGGGCACGGATATGTTCAAACATTCACA-3′) and for *Kif5b* (forward, 5′-CCGCTCGAGTTGAAAAGTAATTGAAGTTTGAAGAGGACATAAAATCAGTCTTT CACTAAC-3′; A-reverse, 5′-ATTTGCGGCCGCCTACAATCCCAAGGAATAGAGGTTAGTTAGTGAAAGACTGATTTTATG-3′; U-reverse, 5′-ATTTGCGGCCGCCTACAATCC CAAGGAATAGAGGA TAGTTAGTGAAAGACTGATTTTATG-3′; G-reverse, 5′-ATTTGCGGCCGCCTACAATCCCAAGG AATAGAGGCTAGTTAGTGAAAGACTGATTTTATG-3′; C-reverse, 5′-ATTTGCGGCCGCCTACA ATCCCAAGGAATAGAGGGTAGTTAGTGAAAGACTGATTTTATG-3′; and seed-reverse, 5′-ATTTGC GGCCGCCTACAATCCCAAGGAATAGAGGTAGTTAGTGAAAGACTGATTTTATG-3′) were used to generate fragments (104 nt) containing the bulge and various mutation sites, and were cloned into the psiCheck-2 plasmid, downstream of *Renilla* luciferase. Twenty-four hours after transfection, relative activity (*Renilla* luciferase activity normalized to firefly luciferase) was measured by dual-luciferase assays (Promega) according to the manufacturer's protocol.

**Microarrays.** RNA from miR-708 or control miRNA transfected HeLa was extracted using Trizol and the RNAeasy kit, and mRNA was amplified and labeled by the method provided by Affymetrix. Human exon 1.0 ST arrays were used, and the data were analyzed by using Affymetrix Power Tools, as described previously[23].

**Ago HITS-CLIP.** Ago HITS-CLIP was carried out as described[23] using monoclonal Ago antibodies, 2A8 (ref. 43) and 7G1-1* (an old batch of 7G1-1 from the University of Iowa Developmental Studies Hybridoma Bank, a mixture of clones of mouse-specific anti-FMRP antibody and anti-Ago antibody, as confirmed previously)[23]. Because it is specific to only mouse FMRP[44], 7G1-1* can be used as an anti-Ago antibody in any human cell line, such as HeLa, without the peptide-blocking FMRP epitope. In brief, miR-708- or control miRNA–transfected HeLa were UV-irradiated to covalently cross-link RNA–protein complexes. After lysing the cells, extracts were partially digested with RNase A to reduce the modal size of cross-linked RNA bound to Ago to ~50 nt. After immunoprecipitating Ago complex with 2A8 or 7G1-1*, Ago–miRNAs (~110 kDa) and Ago-mRNAs (~130 kDa) were purified by SDS-PAGE, followed by nitrocellulose transfer and further purification of RNAs by proteinase K treatment. After generating cDNA libraries by PCR, sequences were analyzed with an Illumina Genome Analyzer. Degenerate barcodes (4-nt tag followed by G) were introduced into the 5′ fusion linker to increase the complexity in unique tags and to avoid artifacts from PCR contamination.


42. Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507–1517 (2004).
43. Nelson, P.T. *et al.* A novel monoclonal antibody against human Argonaute proteins reveals unexpected characteristics of miRNAs in human blood cells. *RNA* **13**, 1787–1792 (2007).
44. Brown, V. *et al.* Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107**, 477–487 (2001).