

Systematic analysis of dynamic miRNA-target interactions during *C. elegans* development

Liang Zhang^{1,*}, Molly Hammell², Brian A. Kudlow¹, Victor Ambros² and Min Han^{1,†}

Although microRNA (miRNA)-mediated functions have been implicated in many aspects of animal development, the majority of miRNA::mRNA regulatory interactions remain to be characterized experimentally. We used an AIN/GW182 protein immunoprecipitation approach to systematically analyze miRNA::mRNA interactions during *C. elegans* development. We characterized the composition of miRNAs in functional miRNA-induced silencing complexes (miRISCs) at each developmental stage and identified three sets of miRNAs with distinct stage-specificity of function. We then identified thousands of miRNA targets in each developmental stage, including a significant portion that is subject to differential miRNA regulation during development. By identifying thousands of miRNA family-mRNA pairs with temporally correlated patterns of AIN-2 association, we gained valuable information on the principles of physiological miRNA::target recognition and predicted 1589 high-confidence miRNA family::mRNA interactions. Our data support the idea that miRNAs preferentially target genes involved in signaling processes and avoid genes with housekeeping functions, and that miRNAs orchestrate temporal developmental programs by coordinately targeting or avoiding genes involved in particular biological functions.

KEY WORDS: GW182, AIN-1, AIN-2, miRISCs, Target recognition, Seed matches, Protein glycosylations

INTRODUCTION

MicroRNAs (miRNAs) are 18- to 25-nucleotide non-coding RNAs that are involved in the post-transcriptional regulation of gene expression, mainly through imperfect base pairing between the miRNA and the 3' UTR of the target mRNA (Ambros, 2004). miRNAs were first identified as developmental regulators in *Caenorhabditis elegans* (Lee et al., 1993; Reinhart et al., 2000; Wightman et al., 1993). They were later recognized as evolutionarily conserved regulatory RNAs present in a wide range of animal species (Pasquinelli et al., 2000). Subsequent studies in many animal systems have provided mounting evidence for the important roles of miRNAs in broad aspects of animal development (Bushati and Cohen, 2007; Stefani and Slack, 2008). For example, in *C. elegans*, ablation of key components of the miRNA pathway leads to dramatic and pleiotropic developmental defects (Grishok et al., 2001; Knight and Bass, 2001; Zhang et al., 2007). Studies of individual *C. elegans* miRNAs have also identified specific roles in regulating development of the hypodermis, vulvae, muscle and nervous system (e.g. Abbott et al., 2005; Chang et al., 2004; Grosshans et al., 2005; Johnson et al., 2005; Johnston and Hobert, 2003; Li et al., 2005; Moss et al., 1997; Simon et al., 2008; Yoo and Greenwald, 2005). However, the number of experimentally identified miRNA targets is still small relative to the vast scale of predicted miRNA::mRNA interactions, and a systematic analysis of the global dynamics of miRNA activity throughout development is still lacking.

Many computational miRNA target prediction algorithms have been developed based on the common features of known miRNA::target interactions and these have greatly facilitated the search for miRNA targets (Bartel, 2009). However, there are several limitations associated with these prediction methods. For example,

most do not account for physiological factors that could influence the existence or outcome of the predicted miRNA::mRNA interactions, including the stoichiometry of miRNAs and targets, or the involvement of regulatory proteins (Bhattacharyya et al., 2006; Didiano and Hobert, 2006; Doench and Sharp, 2004; Mishima et al., 2006). Another limitation is that the predictions depend largely on the available 3'UTR annotations, which may be incomplete, inaccurate and significantly variable among different databases (Ritchie et al., 2009). It is also notable that results from different computational prediction algorithms are highly divergent, and their accuracy under physiological conditions remains to be determined (Rajewsky, 2006; Ritchie et al., 2009).

To obtain a global view of miRNA::target interactions during *C. elegans* development, we applied a biochemical method to systematically identify miRISC-associated miRNAs and mRNAs at different developmental stages. The GW182 family proteins (AIN-1/2 in *C. elegans*) function specifically in miRNA-mediated post-transcriptional regulation of gene expression and are associated with miRNA-specific RNA-induced silencing complexes (RISCs) (Behm-Ansmant et al., 2006; Ding et al., 2005; Ding and Grosshans, 2009; Eulalio et al., 2008; Jakymiw et al., 2005; Liu et al., 2005; Pauley et al., 2006; Rehwinkel et al., 2005; Schneider et al., 2006; Zhang et al., 2007). Biochemical analysis of AIN-1/2-containing complexes revealed that these proteins are associated with miRNA-specific Argonaute proteins but not with small interfering RNA (siRNA)-related Argonaute proteins or common P-body components involved in mRNA degradation in *C. elegans* (Zhang et al., 2007), indicating that the AIN-1/2 containing messenger ribonucleoprotein particles (mRNPs) are dedicated miRNA-effector complexes that are distinct from other P-body mRNPs, siRNA-specific RISCs and miRNA microprocessors.

Based on the specific association of AIN-1/2 with miRNA-effector complexes, we previously established an immunoprecipitation (IP) of AIN-1/2 approach to systematically identify endogenous miRNA targets in *C. elegans* (Zhang et al., 2007), and used those data to develop an improved miRNA target prediction algorithm (Hammell et al., 2008). In this paper, we employ AIN-1/2 IP for the systematic analysis of the dynamics of

¹Howard Hughes Medical Institute and Department of MCDB, University of Colorado, Boulder, CO 80309, USA. ²Program in Molecular Medicine, University of Massachusetts School of Medicine, Worcester, MA 01605, USA.

*Present address: Laboratory of Mammalian Cell Biology and Development, The Rockefeller University, New York, NY 10065, USA

†Author for correspondence (mhan@colorado.edu)

miRNA::target interactions during development. Our findings suggest that microRNAs orchestrate broad temporal programs of gene expression during embryonic and larval development by coordinately targeting specific sets of genes involved in cell signaling and other developmentally important cellular functions.

MATERIALS AND METHODS

C. elegans methods

Worm strains were maintained and manipulated as described by Brenner (Brenner, 1974). All experiments were carried out at 20°C. The single copy integrated AIN-2::GFP transgenic strain was generated previously (Zhang et al., 2007). Large-scale culturing of worms for AIN-2::GFP IP was done on 10 cm NGM plates with concentrated HB101 bacterial lawns. Well-fed animals were harvested and bleached to isolate eggs (Schafer, 2006). The isolated eggs were further purified by spin-floating in 30% sucrose. A portion of the purified eggs was directly flash frozen in liquid nitrogen to be used later as the 'egg' sample. Other purified eggs were hatched and synchronized in S-Basal solution overnight at 20°C. The synchronized L1 worms were mixed with concentrated HB101 bacteria and spotted to 10-cm NGM plates. L1, L2, L3 and L4 animals were collected after 4, 21, 30 and 42 hours on food, respectively. Synchronization was confirmed by examining harvested animals under a Nomarski microscope. Worm samples with >80% of the animals at the correct stage were used for subsequent experiments. The harvested larvae were directly flash frozen in liquid nitrogen for later use. Each replicate of each stage was cultured and harvested independently. Four replicates were collected for each of the egg, L1 and L3 stages. Three replicates were collected for L2 and L4 stages. In total, we performed 18 experiments from the five developmental stages.

Immunoprecipitation (IP)

AIN-2::GFP IP experiments were carried out with a slightly modified procedure from that which was previously described (Zhang et al., 2007). The frozen samples of harvested eggs and larvae were homogenized by grinding with a pestle and mortar in the presence of liquid nitrogen. The homogenized samples were solubilized on ice in three volumes of cold lysis buffer [25 mM HEPES-NaOH (pH 7.4), 150 mM NaCl, 0.2 mM DTT, 10% glycerol, 1% Triton X-100, Complete Protease Inhibitors (Roche), and a recombinant ribonuclease inhibitor (RNaseOUT by Invitrogen) (Tabara et al., 2002)]. The lysed samples were then clarified by centrifuging at 20,000 *g* at 4°C for 15 minutes. Supernatants were further clarified by incubation at 4°C with protein A agarose beads (Sigma) at a ratio of 25 μ l beads per 1 ml worm lysate for 2 hours with gentle rotation. The samples were then centrifuged again at 20,000 *g* at 4°C for 5 minutes. The supernatants were kept for subsequent experiments. The mAb 3E6 anti-GFP monoclonal antibody (Invitrogen) was used for all AIN-2::GFP IP experiments according to the vendor's protocol. For IP, the anti-GFP antibody was added to worm lysate at 1-2 μ g/ml and incubated for 1 hour at 4°C with gentle rotation. Then, 50 μ l lysate was taken out and extracted by Trizol (Invitrogen) according to the vendor's protocol to be used later as the total RNA sample. Protein A agarose beads (Sigma) were added to the remaining lysate at 25 μ l/ml and the sample was incubated at 4°C for another 2 hours with gentle rotation. The beads were then briefly spun down and resuspended in cold Washing Buffer 1 [50 mM Tris (pH 7.5), 150 mM NaCl, 0.1% NP-40, 1 mM EDTA, 100 units/ μ l RNaseOUT (Invitrogen)]. After 10 minutes of incubation at 4°C with gentle rotation, the beads were briefly spun down again and washed again in cold Washing Buffer 1. The beads were then similarly washed sequentially once in cold Washing Buffer 2 (50 mM Tris pH 7.5, 500 mM NaCl, 0.1% NP-40, 100 units/ μ l RNaseOUT) and once in cold Washing Buffer 3 (50 mM Tris pH 7.5, 0.1% NP-40, 100 units/ μ l RNaseOUT). At the end, the washed beads were divided equally into two tubes and the immunoprecipitated RNAs were extracted separately with Trizol (Invitrogen), according to the manufacturer's protocol. RNA from one tube was used for microarray analysis. RNA from the other tube was used for miRNA cloning and subsequent pyrosequencing analysis.

Microarray data generation and analysis

As described in our previous study (Zhang et al., 2007), the RNA samples from the Trizol extractions were shipped to the Washington University Microarray Center (St Louis, MO, USA) for microarray analysis. The

samples were amplified once using the MessageAmp aRNA Kit (Ambion). The IP and Total aRNAs from the same experiment were labeled with cy3 or cy5 using the MicroMax ASAP Kit (Perkin Elmer). The differentially labeled IP and Total aRNAs from the same experiment were hybridized to the same Long Oligomer-based Spotted Microarray for *C. elegans* made by the Genome Sequencing Center at Washington University (http://genome.wustl.edu/services/microarray/caenorhabditis_elegans). The hybridizations were performed according to standard protocol. Dye label flipping was included for different replicates from the same stage to control potential dye label bias. Each microarray had an IP RNA signal channel and the corresponding total RNA signal channel from different dye labeling. Eighteen microarrays experiments were done in total.

The value of (median signal – median local background) from each channel of each spot was used as the signal value for this spot (each representing a *C. elegans* transcript) in this channel for data analysis. Every spot that had a signal value <25 in either channel was flagged as 'unreliable' owing to high variations at low signal range. The signal value of each channel of each spot on each microarray was then divided by the average signal value of the corresponding channel of all *C. elegans* transcripts on the corresponding microarray. Then, the ratio between the signals from the IP channel and the Total channel of each spot that was not flagged as 'unreliable' was calculated. The IP/Total ratios were then log₂ transformed. This value was identified as the 'log₂ transformed fold of enrichment', which directly reflects the enrichment of the transcript in the IP RNA sample compared with the total RNA sample. This log₂ (IP/Total) value was used for all the subsequent data analysis unless otherwise mentioned. Flagged 'unreliable' spots were treated as missing values in this experiment during subsequent analysis.

For each transcript represented on the microarray, we first used all 18 data points from all of the 18 microarrays to calculate its average in-stage standard deviation as $SD = \sqrt{(\sum_{i=1}^K (n_i - 1) SD_i^2) / (N - K)}$, degree of freedom (df) = N - K, where K is the number of developmental stages in which the given transcript had at least one non-missing log₂(IP/Total) value; n_i is the number of non-missing values in stage i among the K stages; SD_i is the standard deviation of the log₂(IP/Total) values from all the replicates of stage i ; N is the total number of non-missing values for this transcript at all stages. A transcript must have at least one stage with at least two non-missing values to be testable. All of the non-missing values of a transcript at each stage were averaged to generate the stage-average values (M_i for stage i). The standard error of M_i was calculated as $SE_i = SD / \sqrt{n_i}$ (df = N - K), where SD is the average in-stage standard deviation calculated above and n_i is the number of non-missing values at stage i , as explained above. Based on the SE_i , a one-tailed Student's t -test was used to calculate the P -value of enrichment in stage i (Test if $M_i > 0$). The T statistic was constructed as $T = M_i / SE_i$ (df = N - K). $M_i > 0$ and enrichment $P < 0.001$ were used as the threshold of enrichment for each stage. We also used the same method to reprocess our previous three GFP IP microarray data from mixed populations (Zhang et al., 2007). We detected zero enriched transcripts at $P < 0.01$ significance level (see Table S2 in the supplementary material). Therefore, we concluded that potential non-specific binding between GFP and miRISC was minimal and thus could be ignored.

The heat map displayed in Fig. 2A was generated by a self-organization map followed by K-mean clustering using the standard cluster software (Eisen et al., 1998). Stage-averaged log₂(IP/Total) values were used. Only transcripts that were testable in at least one stage were included.

The microarray data have been deposited with the Gene Expression Omnibus with the Accession number GSE16208.

FatiScan analysis

For FatiScan analysis, stage-averaged log₂(IP/Total) values of each stage were used. Only the testable genes of each stage were analyzed, and the data of each stage were separately analyzed by FatiScan. The FatiScan algorithm is a segmentation test that checks for asymmetrical distributions of biological labels associated with genes that are ranked in a list. In this study, the genes in each stage were ranked in descending order by their stage-averaged log₂(IP/Total) values. Thirty partitions were tried for each stage. P -values were generated by two-tailed Fisher's exact test followed by false discovery rate (FDR) adjustment (Al-Shahrour et al., 2005a; Al-Shahrour et

al., 2006; Al-Shahrouh et al., 2005b). For each biological label, the smallest *P*-value from all partitions was used as the *P*-value for this label. Labels that were depleted from high ranks or enriched in low ranks were regarded as labels that were avoided by miRNAs. Labels that were enriched in high ranks or depleted from low ranks were regarded as labels that were preferred by miRNAs. To generate the heat map in Fig. 4B, absolute values of $[\log_{10}(P)]$ were used for miRNA preferred labels and $-1 \times \text{Absolute Value}[\log_{10}(P)]$ was used for miRNA avoided labels. The heat map and tree structure were generated by self-organization map followed by complete linkage hierarchical clustering using the standard cluster software (Eisen et al., 1998).

Reporter analysis of the 3'UTR of potential miRNA targets

The *rpl-28* promoter was used to drive ubiquitous expression. The ppd129.57 plasmid (a gift from A. Fire, Stanford University, CA, USA), which contains *rpl-28 promoter:4NLS::gfp:let-858 3'UTR*, was used as the vector for all the reporter constructs. The 3'UTRs of the candidate genes were PCR amplified from genomic DNA and cloned into ppd129.57 to replace the *let-858 3'UTR* and make the GFP reporters.

To make the control mCherry reporters, the *4NLS::gfp*-coding region was replaced by a *his-24::mCherry* translational fusion sequence. For reporter analysis, the candidate GFP reporter, mCherry control, *myo-2::gfp* marker and PBSSK vector DNA were mixed and injected at a concentration of 4, 4, 5, and 95 ng/μl, respectively, into the gonad of wild-type animals following standard protocols. To study the impact of the overexpression of *mir-35* family miRNAs on *egl-1* expression, the coding region of the *mir-35* family cluster was PCR amplified from genomic DNA and cloned into ppd129.57 to replace the *gfp*-coding region. The resulting plasmid (5ng/μl) was added into the standard injection mixture of the *egl-1* GFP reporter as described above. Stable transgenic lines were isolated according to standard procedures. Two independent lines were analyzed.

The primer pairs used to amplify the 3'UTR of candidate genes were:

egl-1 3'UTR, 5'-ATATGCGGCCGCGTGATCAAAATCTCCAACCTT-TTCTC-3' and 5'-ATATGCGGCCGCTCCCGCTTCTGAGAATA-TTG-3'.

hbl-1 3'UTR, 5'-ATATGCGGCCGCGTCTCTGTTAAGGAAACA-CCTCC-3' and 5'-ATATGCGGCCGCGCAGGGATTGCAATTTGAA-AC-3'.

sqv-3 3'UTR, 5'-ATTTGCGGCCGCTCCATAATATTATATCATTAT-TTTATTG-3' and 5'-ATTTGCGGCCGCGAAAAAATAGTACACAC-TCCCTCTAG-3'.

twk-12 3'UTR, 5'-ATTTGCGGCCGCTTATCACACTTTTAAATA-TTGCTACTC-3' and 5'-ATTTGCGGCCGCTTATCTGTTTGTCCAC-TTACCG-3'.

The primer pairs for amplifying the miR-35 cluster were: 5'-TATGAC-CGGTCCAAGTCAGCAGTATAAAAGGAATG-3' and 5'-TAGTCAAT-TGACGAGGATAAAGTGCAACGAG-3'.

miRNA cloning and 454 pyrosequencing/data analysis

miRNA cloning procedures were slightly modified from those in the manufacturer's protocol of the miRCat microRNA Cloning Kit [Integrated DNA technologies (IDT)]. Different RNA species in the RNA samples from Trizol extractions were separated based on molecular mass by running on 12.5% Sequagels (National Diagnostics). The 18 IP RNA samples were each purified separately. Equal amounts of total RNA samples from different replicates of the same stage were combined before gel purification, resulting in five Total RNA samples corresponding to each of the five developmental stages. The positions of miRNAs on the gels were indicated by the miSPIKEtm internal control RNA. The miRNA bands were excised from gels and extracted using the spin column protocol described in the miRCat kit manual. The purified miRNAs were subsequently ligated with the miRNA cloning 3'RNA linker-1 from IDT. The ligation products were gel purified and ligated with the 5' M.R.S. miRNA Cloning Linker from IDT. The final ligation products were reverse transcribed according to the miRCat kit manual. The resulting cDNAs were PCR amplified, with reactions stopped in the linear phase. The products were purified by running on 8% native PAGE gels and excising bands of the correct size. The PCR-gel purification steps were repeated for one more round to further reduce non-

specific bands. The purified miRNA clones were then PCR amplified (PCR stopped in linear phase) with primers containing 454 sequencing adaptors and a four-nucleotide index code sequence unique to each sample. After labeling each sample with a unique index code, the products were purified by running on an agarose gel. In total, there were 23 samples (18 IP samples and 5 Total samples). All the samples were quantified with a spectrometer and then equally mixed at equal concentrations into one combined sample for pyrosequencing in one 454 Production run on an FLX 70x75 PTP plate. The pyrosequencing run was done by the Genomics Services in the Interdisciplinary Center for Biotechnology Research at the University of Florida (Gainesville, FL, USA). The pyrosequencing run was divided into two separated regions that served as two sequencing technical replicates to control for sequencing variations.

AIN-2-association pattern-matching analysis

Because members of the same miRNA family tend to recognize the same binding sites on target mRNAs and these members also display similar patterns of AIN-2 association, we chose to consider each miRNA family as a single unit to reduce the complexity of the correlation analysis. In addition, because low abundance miRNAs tend to generate highly variable data, we only considered the relatively abundant miRNA families, which made up more than 0.1% of all miRNAs in AIN-2 miRISC in at least one developmental stage (as we obtained roughly 10,000 miRNA reads for each sample in our pyrosequencing, 0.1% is roughly 10 reads).

The correlation *P*-values of each miRNA family/miRNA pair were determined by two-tailed Student's *t*-test of the Pearson coefficient between AIN-2 IP concentration of the the miRNA family and AIN-2 IP enrichment of the mRNA in our 18 independent experiments. *P*-values of negative correlations were manually assigned negative values to distinguish them from the *P*-values of positive correlations in Table S4 in the supplementary material. Only the non-missing values in the 18 microarray experiments were used for the correlation calculation, and only the miRNA family::mRNA pairs with sufficient non-missing values for the correlation *P*-value calculation were considered as 'Testable'.

Calculations of the enrichment for miRNA::target features

The 'minimum' miRNA target prediction was generated as described previously (Hammell et al., 2008). Briefly, the 'minimum' prediction applied liberal filters for: (1) the hybrid interaction energy between miRNA and miRNA-binding sites on mRNA, $\Delta G_{\text{hybrid}} (\Delta G_{\text{hybrid}} < -15 \text{ kcal/mol}$ for imperfectly matched seeds, whereas a threshold of $\Delta G_{\text{hybrid}} < -10 \text{ kcal/mol}$ was allowed for perfect seed matches; to include seed-only sites); and (2) the binding site seed configuration, which was set to allow up to three G:U wobble pairs in the seed region or a single seed bulge on the mRNA side of the duplex, as described previously. No conservation filter was used in creating the minimal prediction catalog.

The enrichments calculated for seed matching and total interaction energy in Fig. 5A-D were performed as follows. All minimally predicted miRNA-binding sites in *C. elegans* 3'UTRs were divided into two groups: the pattern-matched miRNA::mRNA pairs and the testable non-pattern-matched miRNA::mRNA pairs. Potential miRNA-binding sites for any member of the same seed family were grouped together as a 'miRNA family site'. We removed any overlapping miRNA sites in a 3'UTR for miRNAs of the same family by randomly choosing one representative site for each miRNA family. As miRNA families share seed sequences, this should not affect the seed enrichment calculations.

For the seed enrichment calculations of Fig. 5A, all non-overlapping, non-redundant miRNA-binding sites were further divided into seed configuration bins, where each testable miRNA-family::mRNA pair was counted once for each seed type. This allowed us to calculate the fold enrichment for having at least one miRNA-binding site of each seed configuration relative to all testable pairs in the pattern-matched and non-pattern-matched categories. The pattern-matched miRNA::mRNA pairs were much more likely to have a corresponding binding site than were the non-pattern-matched pairs, which is why all bins show some enrichment in Fig. 5A. For the seed and energy enrichment calculations of Fig. 5B-D, only conserved, non-overlapping miRNA-binding sites were used in the enrichment calculations. This allowed us to calculate the enrichment for particular subsets of seed types

(Fig. 5B) and interaction energies (Fig. 5C,D) relative to all conserved predicted binding sites. The pattern-matched miRNA::mRNA pairs were more likely to have strong seed matches, with highly favorable interaction energies, suggesting that these parameters correlate with functional binding sites. Although the total interaction energy is a continuous variable, unlike the discrete seed configuration types, we divided the interaction energy into 2 kcal/mol bins for the purposes of calculating enrichments in Fig. 5C,D.

RESULTS

Distinct sets of miRNAs are deployed in miRISCs at different stages in development

To analyze miRNAs in AIN-2-associated miRISCs during development, we combined AIN-2::GFP IP with miRNA cloning and pyrosequencing to generate the miRNA composition profile in the AIN-2-containing miRISC at each developmental stage, and then compared this composition with the miRNA expression profile of whole cell lysates at each corresponding stage. Overall, 86 miRNAs were detected in at least one of the AIN-2 IP samples. Sixty-seven miRNAs were detected in at least one of the total RNA samples, and all of these 67 miRNAs were also detected in the AIN-2 IP samples (Fig. 1A), which is consistent with our previous finding that AIN-2 is broadly involved in all miRNA functions during development (Zhang et al., 2007). The expression patterns of the *lin-*

4 and *let-7* family miRNAs revealed by our analysis of total RNA samples were consistent with their published expression patterns based on northern blots (Abbott et al., 2005; Lim et al., 2003) (Fig. 1B), indicating the validity of the miRNA cloning and sequencing procedure used in this study. Although the relative abundances of miRNAs in the AIN-2 IP generally correlated with their levels in the total lysate at the corresponding developmental stage, there were several instances in which this correlation was weak (see Fig. S1 in the supplementary material).

Through unsupervised hierarchical clustering analysis, four major cluster groups were identified for the 86 miRNAs based on the pattern of their association within the AIN-2-containing miRISCs at each developmental stage, as measured by their concentration in the AIN-2 IP samples for each stage (Fig. 1A). Except for cluster group 4, which is composed of scattered low-abundance miRNAs, each of the other three cluster groups represents a clear temporal pattern of association with the AIN-2 miRISCs (Fig. 1A). This clustering indicates that distinct sets of miRNAs are active at different developmental stages, supporting the idea that the switch between different development programs is accompanied by significant changes in the underlying miRNA::mRNA regulation network (Ason et al., 2006). Interestingly, the most dramatic changes in the miRNA composition of AIN-2 miRISC were observed between egg

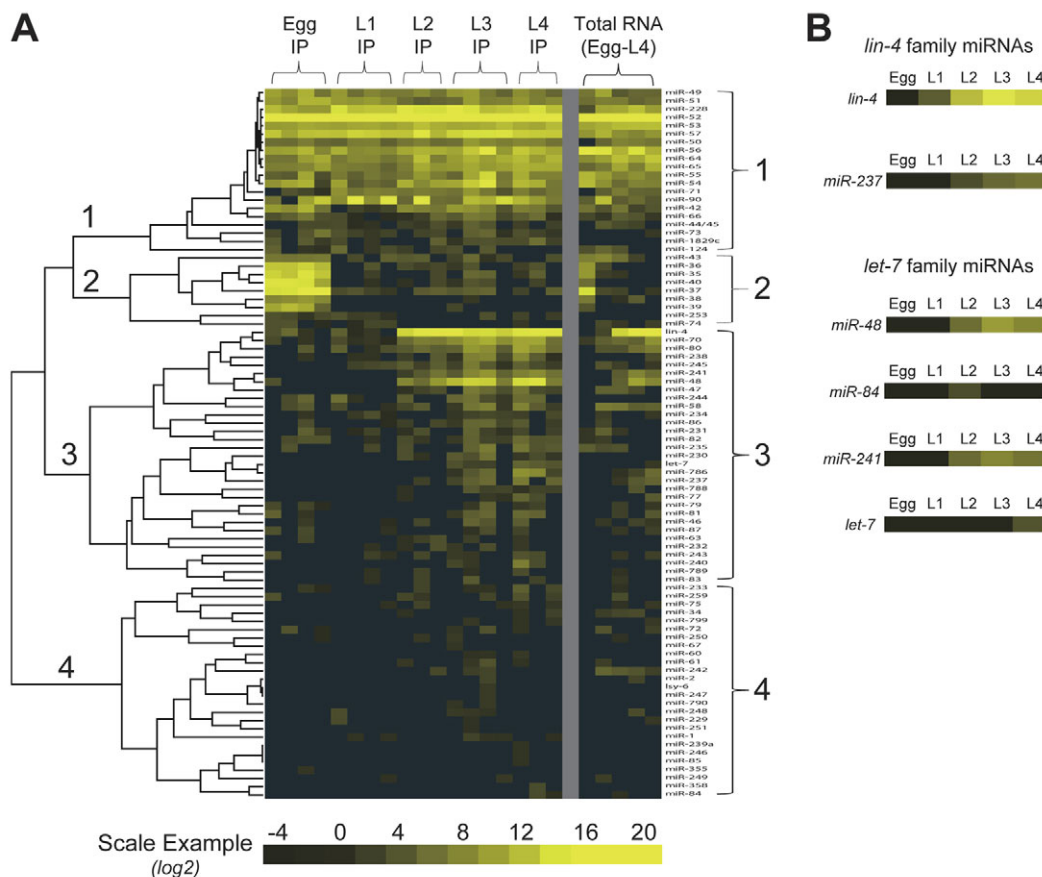


Fig. 1. Dynamic association of miRNAs with AIN-2-miRISCs during development. (A) Heat map illustration of log₂-transformed miRNA concentrations measured by pyrosequencing in 18 AIN-2::GFP IP and five total RNA samples from worms at five indicated developmental stages. Each row represents a miRNA and each column represents an individual sample. Only the 86 detected miRNAs were included. Different columns for the same stage are different replicates for that stage. The heat map and tree structure were generated by the self-organization map approach followed by complete linkage hierarchical clustering using the cluster software. The miRNAs can be grouped into four major clusters: all-stage miRNAs (1), early-stage miRNAs (2), late-stage miRNAs (3) and miRNAs with sporadic low level expressions (4). (B) The expression pattern of *lin-4* and *let-7* family miRNAs determined by pyrosequencing. The results were taken directly from the total RNA samples displayed in A.

Table 1. Number of AIN-2-associated miRNA targets in each stage

Stage	Egg	L1	L2	L3	L4	Any stage [‡]
Targets*	999	1399	1084	1153	991	2094
Percentage [†]	7.2	10.8	8.7	8.6	7.2	13.9
Stage-specific targets [§]	225	265	73	102	72	–
% stage-specific targets [¶]	22.5	18.9	6.7	8.8	7.3	–

*The number of transcripts that are enriched in AIN-2::GFP IP ($P < 0.001$) in each stage (see Materials and methods).

[†]The percentage of the enriched transcripts among all testable transcripts in the corresponding stage (see Materials and methods).

[‡]The transcripts that are enriched in any one of the five stages.

[§]The number of transcripts that are specifically enriched in each stage but not in any other stages.

[¶]The percentage of stage specifically enriched transcripts among all the enriched transcripts in each stage.

and L1 stages, and between L1 and L2 stages, which is consistent with the dynamic changes in the mRNA targets associated with miRISCs during these transitions (see below).

The 5' miRNA seed region is a key determinant for miRNA/mRNA target recognition (Lewis et al., 2005; Xie et al., 2005). Based on shared seed sequences, *C. elegans* miRNAs have been grouped into discrete families, some of which share the seed sequence with miRNAs from other organisms (Miska et al., 2007; Ruby et al., 2006). Our results indicate that members of the same miRNA family tend to have similar temporal patterns of miRISC association despite the divergent chromosome locations of the genes of the family members in most cases (see Table S1 in the supplementary material); this is consistent with their binding to shared targets.

More than 2000 mRNAs associate with AIN-2/GW182-containing miRISC at one or more developmental stage

To generate a global view of the dynamics of miRNA-mediated regulation of gene expression during *C. elegans* development, we analyzed the mRNAs in the AIN-2::GFP IP results from five developmental stages. The magnitude of the combined interaction of miRNAs with a given target mRNA was assessed by measuring the fold enrichment of that mRNA in AIN-2 IP samples, relative to the abundance of the mRNA in the corresponding total lysate (Zhang et al., 2007). Because this enrichment in the IP sample versus total lysate directly reflects the miRISC-associated fraction of a given mRNA, high enrichment indicates the likelihood of strong miRNA-mediated regulation of the mRNA, whereas low or negative enrichment indicates the likelihood of weak or absent miRNA regulation of the mRNA. It is also possible that poor enrichment could reflect interactions that occur only in a rare subset of cells at any given stage of development.

Our results revealed that at each developmental stage, approximately 1000 transcripts were significantly enriched ($P < 0.001$) in the AIN-2-containing miRISCs, indicating that these transcripts are miRNA targets at the corresponding stage. Overall, 2094 out of the 15013 testable transcripts (13.9%) were enriched in at least one stage (Table 1; see also Table S2 in the supplementary material for the full list of the 2094 genes and their AIN-2 IP enrichment at each developmental stage).

Many miRNA::target interactions display dynamic temporal patterns during development

Analysis of the AIN-2::GFP IP data from the stage-synchronized worms generated an overview of the dynamic developmental pattern of miRNA-mediated regulation of the 2094 miRNA targets (Fig. 2). Many of these miRNA targets appeared to be subject to stable and continuous regulation over all five stages, which is consistent with the observation that most miRNAs have relatively steady expression

patterns during *C. elegans* development (Fig. 1) (Lim et al., 2003). For other mRNAs, complex temporal patterns of differential miRNA::target interaction were clearly evident (Fig. 2), indicating that the miRNA/mRNA regulation network is highly dynamic during animal development. Interestingly, the egg and L1 stages appear to have miRNA and mRNA target profiles that are distinct from those in the other developmental stages (Fig. 2A). For example, these two stages have a dramatically higher percentage of stage-specific miRNA targets compared with the L2-L4 stages (Table 1). This, together with the significant changes in the miRNA composition of the AIN-2 miRISC occurring between the egg and L1 stages, and between the L1 and L2 stages (Fig. 1A), suggests that the embryo to larva and the L1 to L2 transitions are accompanied by substantial changes in miRNA-mediated post-transcriptional regulation of gene expression.

Several prominent examples of dynamic miRNA::target regulatory interaction during development are shown in Fig. 2B, including *hbl-1*, the *C. elegans hunchback* homolog; *egl-1*, a proapoptotic gene; *twk-12*, which encodes a potassium channel subunit; and *sqv-3*, which encodes a conserved $\beta(1,4)$ galactosyltransferase. Additional interesting examples of mRNAs with dynamic AIN-2-miRISC association patterns include *imb-1*, which encodes an importin beta family protein; *wrt-2*, which encodes a hedgehog-like protein; *ptr-5*, which encodes a Patched-related protein; and *drsh-1*, which encodes the Droscha pri-miRNA processing enzyme (Fig. 2D,E). Importantly, the dynamic AIN-2-miRISC enrichment patterns we observed here are unlikely to be a simple consequence of the changes in mRNA expression levels, because most of the mRNA expression profiles of the above genes are not correlated with their enrichment in AIN-2 IP (Fig. 2B,C). Furthermore, the developmental pattern of AIN-2 IP enrichment and the total RNA expression level of all of the testable transcripts were, on average, negatively correlated (Median Pearson's correlation coefficient = -0.46 ; Fig. 2F), which suggests that many of the miRNA::target interactions identified in this study could cause a decrease in the cellular level of the target mRNA, either directly, through post-transcriptional mRNA downregulation, or indirectly, through effects on transcription.

Overall, 589 of the 2094 miRNA targets (28%) displayed a significantly different degree of AIN-2 IP enrichment at different stages (ANOVA Test, $P < 0.01$; see Table S2 in the supplementary material), indicating that these transcripts are subject to stage-dependent miRNA regulation. One of the stage-dependent miRNA targets is *hbl-1*. Previous studies have shown that *hbl-1* is a direct target of the *let-7* family miRNAs and that *let-7/hbl-1* regulation is essential for the developmental transition of lateral hypodermal cells from the L2 to L3 stage (Abbott et al., 2005; Abrahante et al., 2003; Grosshans et al., 2005; Lin et al., 2003). Reflecting the known regulation of *let-7* family miRNAs on *hbl-1* at the L2/L3 transition, the enrichment of *hbl-1* in our AIN-2 IP microarray data displayed

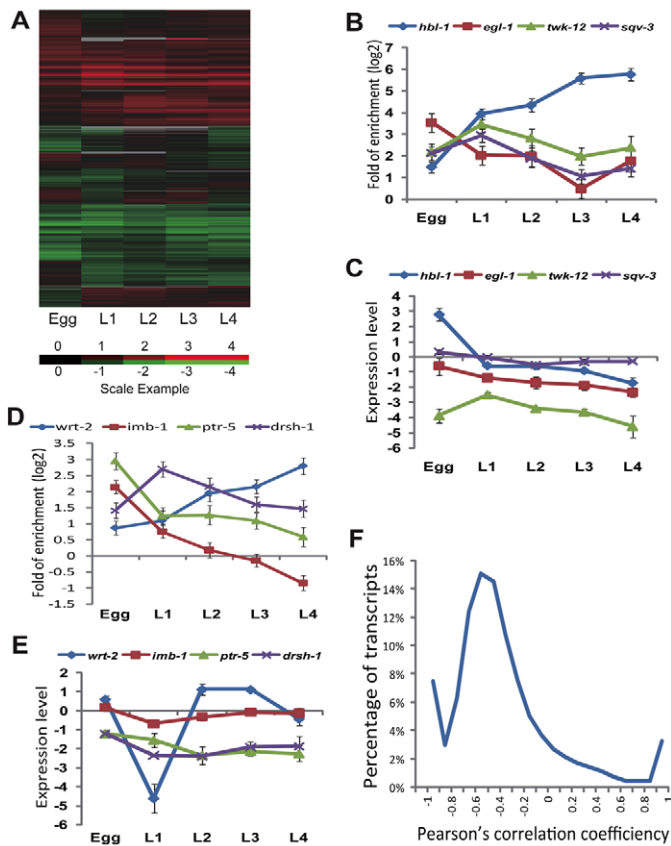


Fig. 2. Dynamic pattern of fold of enrichments of miRNA targets in AIN-2 IP during development. (A) Heat map illustration of enrichment of the 15,013 testable transcripts in the AIN-2::GFP IP from worms at the five developmental stages. Log₂-transformed stage-averaged fold of enrichment was displayed. Each row represents a transcript. Each column represents a stage as indicated. Red indicates enrichment in IP, green indicates depletion from IP, and gray indicates missing values. (B,D) Eight examples of dynamic miRNA::target interactions during development. The vertical axis represents log₂-transformed stage-averaged fold of enrichment. (C,E) The mRNA expression pattern of the example genes in B,D during development. The vertical axis represent log₂ transformed stage-averaged mRNA expression level, measured by dividing the microarray signal of the mRNA in each total lysate sample by the average signal of all mRNAs in the same sample. Error bars in B-E indicate s.e.m. (F) Distribution of the Pearson's correlation coefficient between the expression level and the fold of enrichment in AIN-2 IP of all the testable transcripts during development. The vertical axis represents the percentage of transcripts within each 0.1 interval of the correlation coefficient value. The median Pearson's correlation coefficient value of all the testable transcripts is -0.46.

a significant increase from the L2 to the L3 stage ($P=0.003$; Fig. 2B). Interestingly, substantial enrichment of *hbl-1* was also observed at L1 and L2, and to a lesser extent even at the egg stage (Fig. 2B). At the L1 and embryonic stages, the *let-7* family miRNAs were essentially not present (Abbott et al., 2005), which suggests that *hbl-1* is subject to regulation by other miRNAs at these stages before the emergence of the *let-7* family miRNAs. Further support for this finding came from analysis of the expression of reporter transgenes. A GFP reporter containing the *hbl-1* 3'UTR, but not an mCherry reporter containing a control 3'UTR on the same extrachromosomal

array, displayed a strongly attenuated expression in all tissues during the egg, L1 and L2 stages, and this expression was reduced further to virtually undetectable levels in the L3 and L4 stages (Fig. 3B,C; Table 2).

By using this GFP reporter assay, we also examined the developmental miRNA regulation dynamics of two other regulatory genes whose dynamic miRISC association is shown in Fig. 2B: the pro-apoptotic gene *egl-1* (Conrad and Horvitz, 1998) and the two-pore domain potassium channel subunit gene *twk-12*. *egl-1* was strongly enriched in the AIN-2 miRISC in embryos, but this enrichment was significantly attenuated when the animals progressed into larval development. This pattern of enrichment in the AIN-2::GFP complex was temporally correlated with the repression of an *egl-1* 3'UTR reporter in eggs and de-repression of the same reporter in later larval stages (Fig. 3D; Table 2). By contrast, the *twk-12* gene appeared to be subject to relatively stable miRNA regulation across development, with a modest peak at the L1 stage. Consistent with this observation, we found that a *twk-12* 3'UTR reporter showed stable repression in each stage that we examined (Fig. 3E; Table 2). These observations indicate that the temporal pattern of mRNA association with the AIN-2 miRISC is correlated with the pattern of repression and de-repression of the 3'UTR of the mRNA.

miRNAs preferentially regulate genes involved in cell signaling and avoid housekeeping genes

The distribution of the magnitude of miRISC-association (as reflected by the fold enrichment in the AIN-2 IP) for all the transcripts on our microarray is approximately normal with a slight bias toward the positive side (Zhang et al., 2007) (Fig. 4A). However, clearly biased distributions of miRISC-association intensities were observed for genes that belong to individual functional groups. For example, at the L2 stage, ribosomal genes displayed a very strong tendency towards negative enrichment in the AIN-2 IP, whereas receptor (Gene Ontology term: receptor activity) genes tended to show higher than average positive enrichment (Fig. 4A; FDR $P<0.01$ in both cases). These data suggest that, at the L2 stage, ribosomal genes are strongly excluded from miRNA regulation, while receptors genes are preferentially regulated by miRNAs.

To systematically assess the impact of miRNAs on genes associated with different biological functions during development, we used the FatiScan algorithm (Al-Shahrour et al., 2005a) in the Babelomics 2.0 suite (Al-Shahrour et al., 2005b) to analyze potential asymmetrical distributions of functional groups of genes with respect to their levels of enrichment in the AIN-2 IP at each developmental stage. This analysis revealed that genes in eight KEGG annotated pathways, three Biological Process Gene Ontology Terms, and as many as 40 Molecular Function Gene Ontology Terms displayed significantly asymmetrical distributions in the levels of miRNA::target interactions within the group (FDR $P<0.01$) in at least one developmental stage (Fig. 4B; see also Table S3 in the supplementary material). Housekeeping genes, such as ribosomal genes, oxidative phosphorylation genes, and basic translational factors, appeared to be associated with no or very weak miRNA regulation at all stages (Fig. 4B), indicating that miRNAs tend to avoid constitutively expressed genes that are required for essential biological functions. This result is in remarkable agreement with a previous bioinformatic analysis that predicted these genes as *Drosophila* miRNA 'anti-targets', which generally have short 3'UTRs and low densities of conserved miRNA recognition sites (Stark et al., 2005). By contrast, AIN-2-associated mRNAs were

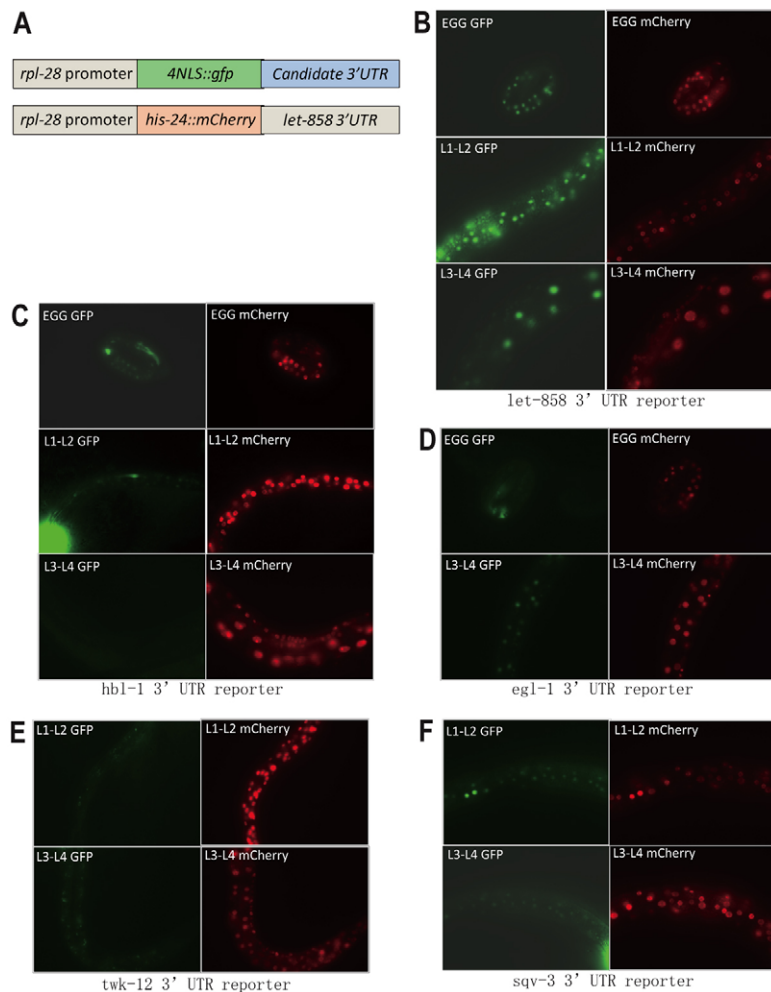


Fig. 3. Reporter analysis of dynamics of miRNA regulation on the expression of four genes during development. (A) Cartoon illustration of the dual-color reporter analysis strategy. As the control for transgene expression, *rpl-28* driving *his-24::mCherry:let-858 3'UTR* was co-expressed with the GFP reporters in each animal from the same extra-chromosomal array.

(B-F) Fluorescence images of living animals expressing GFP reporters of four indicated 3'UTRs. The stage of each image is indicated. All GFP reporters contain a 4× SV40 nuclear localization signal (NLS) and were driven by the ubiquitous *rpl-28* promoter. The mCherry images were taken immediately after the GFP images of the same animals using exactly the same scope settings.

highly enriched for genes that are involved in cell-cell signaling processes, such as receptors, kinases and ion channels (Fig. 4B), indicating that the signal transduction processes are subject to intensive miRNA regulation in vivo.

Genes involved in protein glycosylation pathways are subject to direct miRNA regulation

Protein glycosylation is important for cell-cell signaling and defects in glycosylation pathways have been linked to a wide range of human developmental disorders (Freeze and Aebi, 2005). Although genes involved in protein glycosylation pathways have not been previously identified as direct miRNA targets, we found that this gene category was significantly enriched for genes with strong miRNA regulation (Fig. 4B, the first five terms at the top of the heat map). To support this finding, we analyzed the activity of a 3'UTR GFP reporter for *sqv-3*, which encodes a conserved $\beta(1,4)$ galactosyltransferase that has been characterized for its function in vulval morphogenesis (Bulik et al., 2000; Herman and Horvitz, 1999; Okajima et al., 1999). Consistent with the observation that the *sqv-3* transcript displayed a significant enrichment in AIN-2 IP in all five developmental stages (Fig. 2B), the 3'UTR of *sqv-3* conferred a strong negative regulation on the reporter gene expression in all four larval stages (Fig. 3F, Table 2). Therefore, our data suggest that at least some conserved genes involved in protein glycosylation pathways, which play important roles in cell signaling, are subject to direct miRNA regulation during animal development.

Distinct biological functions are preferentially targeted by miRNAs at different developmental stages

Stage-specific changes in miRNA regulation were clearly evident for many gene classes (Fig. 4B). For example, genes classified as being involved in embryonic development were significantly avoided by egg miRNAs. The statistical significance of this preference was diminished in the subsequent L1, L2 and L3 larval stages (Fig. 4B), suggesting that a substantial fraction of genes involved in embryonic development need to be coordinately repressed by miRNAs after the embryo/larva transition.

Other striking examples of stage-dependent miRNA regulation include: the nucleotide binding protein genes, which were specifically avoided by miRNAs in the embryo; the sugar binding protein genes, which were specifically avoided in the L1 stage (Fig. 4B); and various ion channel genes, which appeared to be more strongly targeted by miRNAs before the onset and also at the end of larval development (Fig. 4B). Although the biological significances of these patterns remain to be elucidated, these data, together with the above example of embryonic development genes, suggest that miRNAs can coordinately target or avoid certain categories of genes in a temporally regulated manner, thereby imposing broad patterns of regulation upon the signaling networks that mediate development.

Table 2. Analysis of 3'UTR reporter expression using transgenic animals

3'UTR (line number)	Reference tissue	GFP intensity	Egg	L1,L2	L3,L4
<i>let-858</i> (A)	All	++	97.0%±3.0%	97.2%±2.8%	100%±0.0%
		+	3.0%±3.0%	2.8%±2.8%	0.0%±0.0%
		-	0.0%±0.0%	0.0%±0.0%	0.0%±0.0%
		<i>n</i>	46	56	77
<i>let-858</i> (B)	All	++	73.2%±1.0%	92.3%±3.9%	96.2%±0.3%
		+	26.8%±1.0%	7.7%±3.9%	3.8%±0.3%
		-	0.0%±0.0%	0.0%±0.0%	0.0%±0.0%
		<i>n</i>	136	67	80
<i>hbl-1</i> (A)	All	++	0.0%±0.0%	0.0%±0.0%	0.0%±0.0%
		+	94.1%±3.4%	72.7%±3.9%	5.8%±3.3%
		-	5.9%±3.4%	27.3%±3.9%	94.2%±3.3%
		<i>n</i>	46	70	88
<i>hbl-1</i> (B)	All	++	0.0%±0.0%	0.0%±0.0%	0.0%±0.0%
		+	64.5%±1.5%	74.2%±0.8%	6.3%±6.3%
		-	35.5%±1.5%	25.8%±0.8%	93.8%±6.3%
		<i>n</i>	79	62	80
<i>egl-1</i> (A)	Intestine	++	0.0%±0.0%		26.6%±3.1%
		+	21.3%±1.5%		71.6%±2.2%
		-	78.7%±1.5%		1.9%±1.9%
		<i>n</i>	99		89
<i>egl-1</i> (B)	Intestine	++	1.7%±1.7%		3.1%±1.5%
		+	25.6%±0.6%		93.0%±2.4%
		-	72.7%±1.5%		3.9%±3.9%
		<i>n</i>	70		85
<i>sqv-3</i> (A)	All	++		1.7%±1.7%	6.0%±1.2%
		+		76.1%±21.4%	94.0%±1.2%
		-		22.2%±22.2%	0.0%±0.0%
		<i>n</i>		67	66
<i>sqv-3</i> (B)	All	++		2.5%±1.4%	10.8%±0.9%
		+		81.7%±15.9%	89.2%±0.9%
		-		15.8%±15.8%	0.0%±0.0%
		<i>n</i>		70	74
<i>twk-12</i> (A)	All	++		0.0%±0.0%	0.0%±0.0%
		+		2.8%±2.8%	1.1%±1.1%
		-		97.2%±2.8%	98.9%±1.1%
		<i>n</i>		55	79
<i>twk-12</i> (B)	All	++		0.0%±0.0%	0.0%±0.0%
		+		1.8%±1.8%	1.2%±1.2%
		-		98.2%±1.8%	98.8%±1.2%
		<i>n</i>		64	62
<i>egl-1</i> 3'UTR+ <i>mir-35</i> (A)	Intestine	++			0.0%±0.0%
		+			10.4%±3.1%
		-			89.6%±3.1%
		<i>n</i>			76
<i>egl-1</i> 3'UTR+ <i>mir-35</i> (B)	Intestine	++			0.0%±0.0%
		+			18.1%±0.8%
		-			81.9%±0.8%
		<i>n</i>			120

Each of the *rpl-28* promoter:4NLS::gfp::candidate 3'UTRs constructs was co-injected with the reference *rpl-28* promoter:*his-24*::mCherry::*let-858* 3'UTRs construct to create transgenic animals. Two independent transgenic lines were analyzed for each candidate 3'UTR (line A and line B). The GFP expression level was measured by comparing the fluorescence intensity of GFP with that of mCherry. The percentage of animals with each GFP expression level in each transgenic line at indicated development stages was counted and listed in the table.

++, GFP fluorescence is stronger than or comparable with the mCherry fluorescence.

+, GFP fluorescence is visible but much weaker than the mCherry fluorescence.

-, GFP fluorescence is mostly invisible despite the presence of strong mCherry fluorescence.

n indicates the total number of counted animals.

Reference tissue indicates the tissue that the counting was based on.

AIN-2-associated pattern-matched miRNA family::mRNA pairs are significantly enriched for known and predicted miRNA targets in *C. elegans*

Previously, we have demonstrated that the presence of miRNA targets in the AIN-2 miRISC is dependent on the presence of regulatory miRNAs (Zhang et al., 2007). Therefore, a tight correlation between the temporal enrichment patterns of an miRNA and an mRNA in AIN-2 complexes under different physiological conditions could be used as an indication of an

in vivo miRNA::mRNA regulatory interaction. Because members of the same miRNA family tend to recognize the same binding sites on target mRNAs, and because these members also display similar patterns of AIN-2 association, we chose to consider each miRNA family as a single unit to reduce the complexity of the correlation analysis. In addition, we excluded low abundance miRNAs, which tend to generate highly variable data. This left us 16 miRNA families (see Table S4 supplementary material).

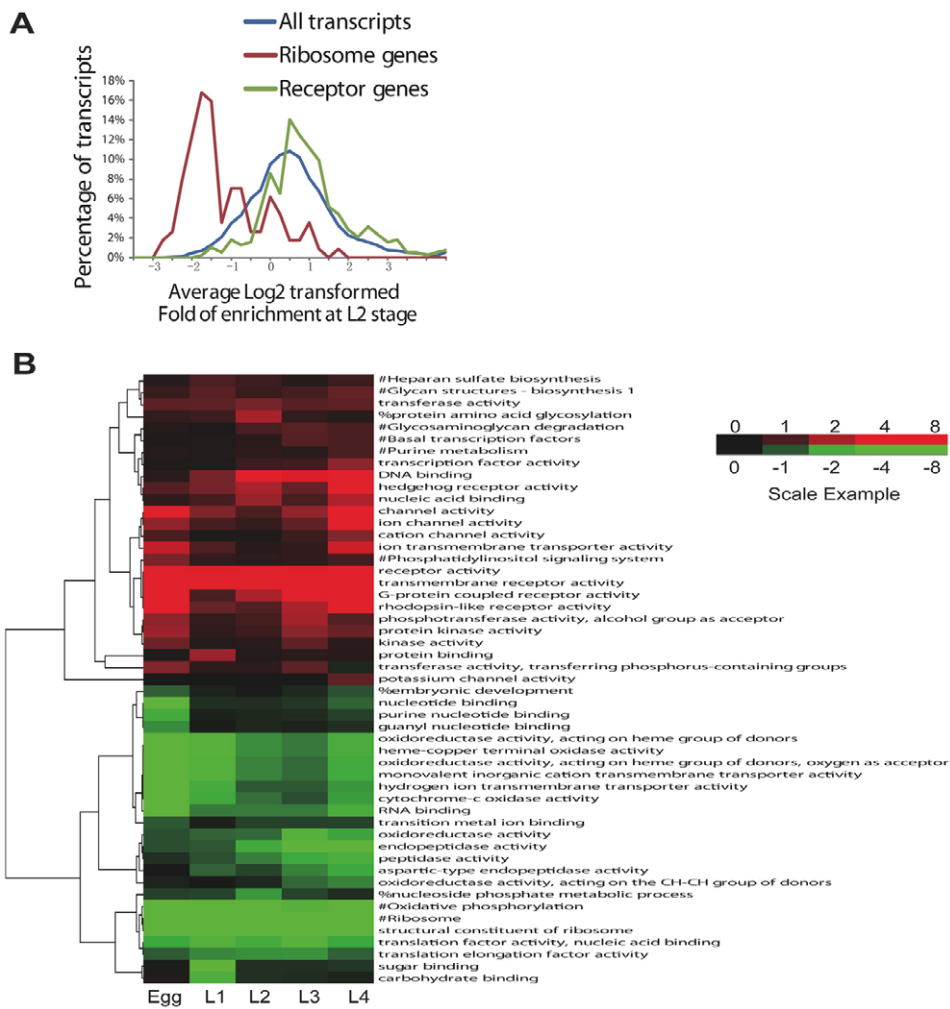


Fig. 4. miRNA regulation prominently prefers genes involved in cell signaling and avoids genes of housekeeping functions during development.

(A) Distribution of miRNA regulation intensity on selected gene classes at the L2 stage. The vertical axis represents the percentage of transcripts within each 0.25 interval of log₂-transformed stage-averaged fold enrichment. Receptor genes represent genes annotated with GO term: receptor activity. Ribosome genes represent genes annotated with GO term: structural constituent of ribosome. (B) Heat map and hierarchical clustering display of absolute value of log₁₀-transformed *P*-values for the association between biological labels and the relative miRNA regulation intensity based on Fatiscan analysis (see Table S3 in the supplementary material). Each column represents a stage as indicated. Each row represents a biological label as indicated. '#' indicates KEGG pathway annotations (Kanehisa, 1997; Kanehisa et al., 2006); '%' and all the other labels indicate GO: Biological Process annotations and GO: Molecular Function annotations, respectively (Ashburner et al., 2000). Only labels that were significantly preferred (red) or avoided (green; FDR adjusted, *P*<0.01) in at least one stage were displayed.

To identify tightly correlated miRNA family::mRNA pairs that are likely to represent bona fide, *in vivo* regulatory interactions, we considered the following criteria. (1) In the AIN-2 IP results from different developmental stages, the pattern of association of a miRNA family must have significant positive correlation with the enrichment of its target mRNAs (Positive Correlation, $0 < P < 0.05$). (2) A *bona fide* target of a miRNA family should display at least weak enrichment ($P < 0.05$) in the AIN-2 IP in at least one developmental stage where its correlated miRNA represented more than 0.1% of all miRNAs in AIN-2 IP. (3) If a miRNA family constituted more than 0.1% of all miRNA reads in AIN-2 IP at a developmental stage, its target mRNA should not display significant depletion (Negative enrichment, $P < 0.05$) in the AIN-2 IP of that stage.

Overall, 6,642 of the 240,099 testable (2.8%) miRNA family::mRNA pairs satisfied the above criteria and were considered as AIN-2-association pattern-matched (abbreviated as 'pattern matched' thereafter) miRNA family::mRNA pairs (see Table S4 in the supplementary material). Further analysis indicated that both experimentally identified and computationally predicted miRNA::mRNA interactions were highly enriched in the pattern-matched miRNA family::mRNA pairs (see Fig. S3 in the supplementary material), thereby lending support to the notion that IP pattern matching is an effective way to detect functional miRNA::target interactions.

AIN-2 association pattern-matched miRNA family::mRNA pairs are enriched for two classes of miRNA::target seed matches

The miRNA-binding sites in the 3'UTRs of target mRNAs often have perfect Watson-Crick complementarity to the 5' end of the miRNA (miRNA seed) (Lai, 2002; Lewis et al., 2003; Stark et al., 2003). In our 6,642 pattern-matched miRNA family::mRNA pairs, we detected a dramatic enrichment of potential miRNA-binding sites with perfect 7-nucleotide complementary matching at the 5' end position 2-8 of the corresponding miRNA (Fig. 5A, Perfect 7-mer and Position 1 8-mer), supporting the notion that perfect 7-nucleotide seed matching is a highly favored feature of *in vivo* miRNA::target recognition (Bartel, 2009). At the same time, potential miRNA-binding sites with imperfect complementarity to the 5' end of the miRNA containing up to three G:U wobbles also displayed considerable enrichment in our pattern-matched pairs (more than 2-fold, Fig. 5A), which is consistent with the observation that imperfect base pairing to the miRNA 5'-end seed region can mediate miRNA regulation (Brennecke et al., 2005). The substantial enrichment shown for 7-mer seed containing binding sites suggests that this subset of pattern-matched miRNA family::mRNA pairs represents the highest confidence target among all of the predicted interactions. However, only 24% of the pattern-matched miRNA family::mRNA pairs with minimal predictable miRNA-binding

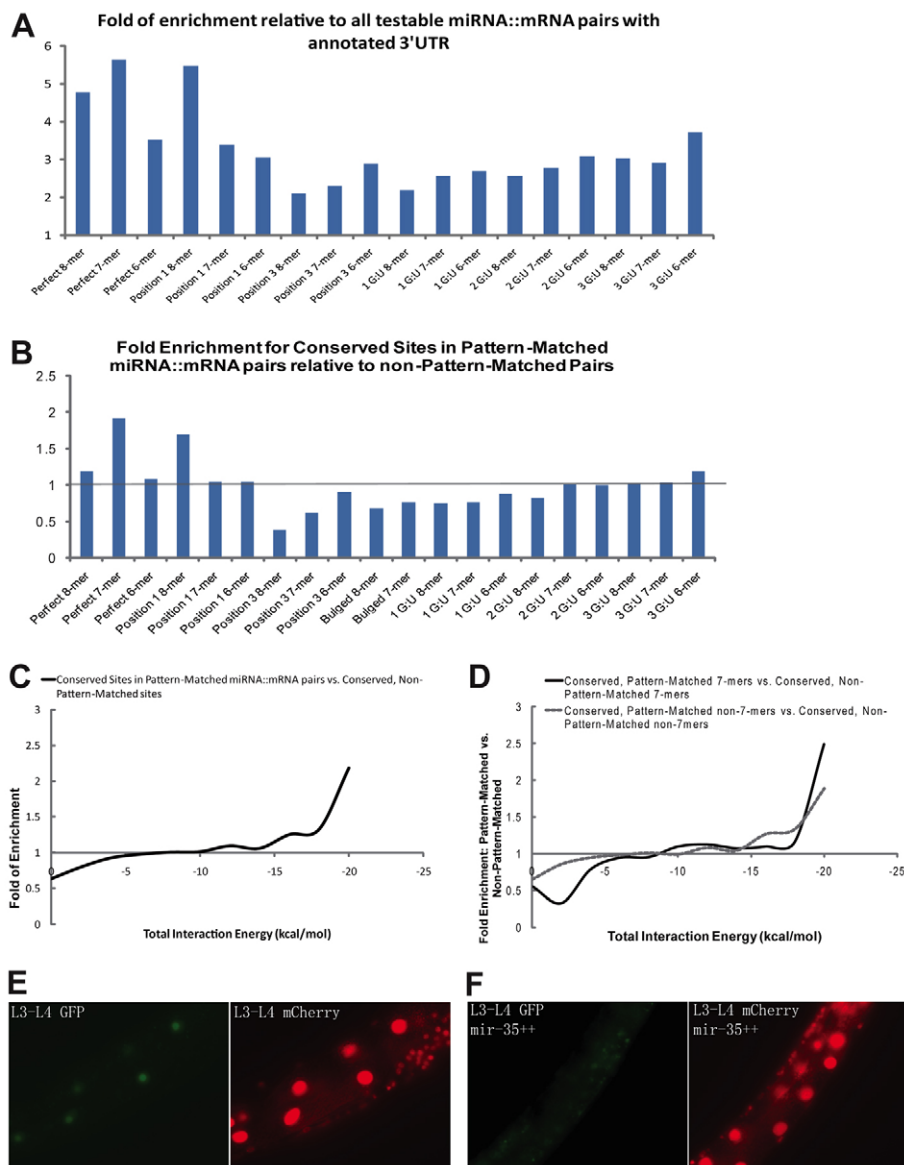


Fig. 5. Matching AIN-2 association patterns of miRNA and mRNA during development greatly facilitates the search for specific miRNA::target interactions.

(A) Fold of enrichment for different types of seed matches in the pattern-matched miRNA family::mRNA pairs containing annotated 3'UTRs compared with all other testable miRNA family::mRNA pairs containing annotated 3'UTRs. Perfect 7-mer refers to a stretch of continuous 7-nucleotide that are a perfect complementary match to the 5' end of the miRNA starting from position 2. Position 1 or position 3 7-mer means the 7-nucleotide perfect match started from position 1 or position 3 instead. 1 G:U 7-mer means there is one G:U wobble pair within the position 2-8 7-mer match. (B) The fold of enrichment for different types of seed matches in conserved potential miRNA family::mRNA-binding sites that are pattern matched versus those that are not pattern matched. (C) The fold of enrichment for different total interaction energy in conserved potential miRNA family::mRNA-binding sites that are pattern matched versus those that are not pattern matched. (D) The fold of enrichment for different total interaction energy in conserved potential miRNA family::mRNA-binding sites with particular seed configurations that are pattern matched versus those that are not pattern matched. 7-mer refers to any miRNA::target helix that has perfect position 2-8 complementarity, which included the position 2-8 7-mers, position 2-9 8-mers, and position 1-8 8-mers from A. 'Non 7-mers' includes all other seed types in A. (E,F) Fluorescence images of live animals expressing the GFP reporter of the *egl-1* 3'UTR in a wild-type (E) or a *mir-35* family overexpression (*mir-35* ++, F) background. The stage of each pictured animal is indicated. See Fig. 3 for information regarding the reporter system.

sites (see below) had 7-mer seed matching (data not shown), suggesting that weaker seed matches also contribute to a large fraction of miRNA interactions across development.

As cross-species conservation has often been used in conjunction with seed matching to predict functional miRNA-binding sites, we further evaluated the enrichment of different seed types in the 6,642 pattern-matched miRNA family::mRNA pairs over non-pattern-matched pairs by considering only conserved potential miRNA-binding sites (Fig. 5B). Among all the conserved potential miRNA family::mRNA-binding sites, only the position 2-8 perfect 7-mer matching (Perfect 7-mer and Position 1 8-mer) displayed a striking enrichment in the pattern-matched results, supporting the idea that conserved 7-mer seed matching is a strong indicator of functional miRNA-binding sites. An additional match at position 1 seems to be insignificant as the Position 1 8-mer matches did not display any further enrichment over Perfect 7-mers (Fig. 5B). Consistent with 7-nucleotide seed matching being highly specific to the 5' end position 2-8 of the miRNA, perfect position 1-7 matching (Position 1 7-mer) displayed no enrichment in the pattern-matched pairs. Furthermore,

additional Watson-Crick base pairing at position 9 and perhaps position 10 after a stretch of perfect matches in the seed region appears to be detrimental for miRNA::target recognition, as matches of both position 3-9 (Position 3 7-mer) and position 3-10 (Position 3 8-mer), but not position 3-8 (Position 3 6-mer), were prominently depleted from the pattern-matched pairs (Fig. 5B). The negative impact of the matching at position 9 on miRNA::target recognition appears to be sufficient to neutralize the positive impact of the conserved position 2-8 seed matching, as the position 2-9 (Perfect 8-mer) match displayed no prominent enrichment (Fig. 5B).

Favorable binding energy is enriched in the AIN-2 association pattern-matched miRNA family::mRNA pairs

Another important factor for miRNA::target mRNA recognition is the total free energy, ΔG_{total} , that is required to alter the local mRNA structure and form the microRNA::target duplex (Kertesz et al., 2007; Long et al., 2007). In our 6,642 pattern-matched pairs, miRNA::mRNA pairs with favorable ΔG_{total} were highly enriched

(Fig. 5C). However, because the 5'-end seed-matching feature observed above can potentially provide highly favorable ΔG_{total} already, we examined whether the favorable energy feature can contribute to enrichment in the pattern-matched results beyond the contribution of the the 5'-end seed-matching feature. Alternatively, miRNA-binding sites with a weak seed match (non-7-mer match) might need support from additional parameters like favorable total energy, in order to form a stable hybrid. Surprisingly, pattern-matched miRNA::mRNA pairs with both strong and weak seed matches showed comparable enrichment for favorable ΔG_{total} (Fig. 5D). Therefore, the enrichment seen for favorable total interaction energy is independent of seed-match status, suggesting that stable interactions are generally enriched among AIN-associated targets.

A total of 1589 potential miRNA family::mRNA regulatory pairs during development were identified through the combination of IP pattern matching and computational target prediction

Some miRNA families displayed similar developmental patterns of AIN-2 miRISC association (Fig. 1A), which made it difficult to effectively distinguish their targets by the pattern-matching approach alone. We thus combined the pattern-matching approach with a 'minimum' computational miRNA target prediction based on liberal filters of hybridization energy and seed matching (see Materials and methods). We reasoned that, besides pattern matching, a true target of a miRNA must have at least a minimally complementary binding site for the miRNA in its 3'UTR, whereas false positive targets would not. Therefore, this combined approach would help to eliminate some false positives by eliminating pairs displaying spurious correlations between a miRNA family and an mRNA.

Overall, 1,589 of the 6,642 (24%) pattern-matched miRNA family::mRNA pairs also carried a potential binding site for the miRNA family on the mRNA (see Table S5 in the supplementary material for a full list of these 1589 pairs and information regarding gene annotation, correlation *P*-value and the AIN-2 IP data of the given miRNA and mRNA in each developmental stage). Among these, all four of the experimentally identified miRNA targets in the correlated pairs were preserved (see Table S5 in the supplementary material). The four shuffled pattern-matching controls (Fig. S3 in the supplementary material) gave an average of 632 (s.e.m.=24) pattern-matched pairs that also carried a potential miRNA-binding site. Therefore, we estimate that the signal to noise ratio of the 1,589 miRNA family::mRNA pairs identified by the combined approach of pattern matching and prediction is around 2.5:1.

Among the 1,589 miRNA family::mRNA regulation pairs identified by the combination approach, *egl-1* was identified as a target of the *mir-35* family of miRNAs (see Table S5 in the supplementary material). The *mir-35* family miRNAs are essential for embryonic development (Miska et al., 2007). They are abundantly present in the AIN-2 miRISCs in embryos but were greatly reduced after the embryo/larva transition (Fig. 1A), which parallels a decrease seen in the intensity of miRNA regulation on an *egl-1* 3'UTR GFP reporter from egg to larval stages as described above (Fig. 3D, Table 2). To provide support for the regulation of the *egl-1* 3'UTR by *mir-35* family miRNAs, we ectopically expressed the *mir-35* family miRNAs in larval stages with the constitutive *rpl-28* promoter. The ectopic expression of *mir-35* family miRNAs induced a marked reduction in the expression of a GFP reporter carrying the *egl-1* 3'UTR in larval stages (Fig. 5E,F; Table 2). These data suggest that *egl-1* is an endogenous target of *mir-35* family miRNAs.

DISCUSSION

Through systematic analysis of miRNAs and mRNAs associated with the AIN-2 miRISC in *C. elegans*, we have identified specific sets of active miRNAs and mRNA targets at individual stages of development. Our experimental data revealed that miRNAs preferentially target genes involved in signaling processes and avoid genes with housekeeping functions during development. The avoidance of miRNA regulation of specific housekeeping genes, such as those required for structural components of the ribosome, oxidative phosphorylation, protein degradation and translation elongation (Fig. 4B), is in remarkable agreement with a previous bioinformatic prediction of these genes as *Drosophila* miRNA 'anti-targets' (Stark et al., 2005). Conversely, the miRNA preferences toward specific signaling molecules, such as hedgehog receptors, ion channels and protein glycosylation enzymes (Fig. 4B), have so far only emerged from our experimental analysis, providing important insights into the regulatory roles of miRNAs in animal development.

The miRNA composition of the AIN-2-associated miRISC displayed a striking stage-dependent pattern, in which most miRNAs clustered into a small number of broad temporal association classes. The progression of development from embryo to a fully developed animal is accompanied by a gradual loss of early stage miRNAs and a gradual accumulation of late stage miRNAs, with the most significant changes observed at the egg to L1, and L1 to L2 transitions. Meanwhile, the miRNA target profile of the AIN-2-associated miRISC also changes accordingly from stage to stage, displaying distinct preferences towards or against specific biological functions at different stages. Taken together, these results suggest that the complement of miRNAs present at a given time during development function together to orchestrate a developmental program by coordinately targeting or avoiding genes involved in certain biological functions.

Further analysis of pattern-matched miRNA family::mRNA pairs resulted in several interesting insights into the features of physiological miRNA::target recognition, which could have important implications for the development of more effective miRNA target prediction algorithms. First, the analysis supports the idea that position 2-8 perfect seed matching, especially when it is conserved across species, is an outstanding feature of physiological miRNA::target interactions and should be weighted much more heavily than any other seed types. Second, matching at position 9 and 10 in addition to a continuous stretch of perfect matches in the 5' end region of the miRNA may be detrimental for in vivo miRNA::target recognition, at least in the case of stable interactions. We speculate that miRNA::mRNA matching near the center of the miRNA is structurally similar to the cleavage site of siRNA::target duplex and is avoided by miRISCs when the cleavage of target mRNA is undesirable. Although it is also possible that miRNA::target matches that lead to target cleavage would be too short-lived to be detected as being enriched in AIN-IP miRISC, the above speculation is consistent with the observation that after the overexpression of miRNAs in HeLa cells, protein production from seed-containing mRNAs with perfect base pairing from nucleotides 9 to 11 and mRNAs lacking seeds was indistinguishable (Selbach et al., 2008). Potential miRNA::target helices with highly favorable total binding energy were dramatically enriched in our pattern-matched results, even in the presence of conserved 7-mer seed matching, suggesting that favorable total binding energy does make a contribution to miRNA::target recognition independently of seed matching. However, the favorable interaction energy did not seem to correlate more strongly with sites bearing a strong or weak seed

match. Therefore, the enrichment seen for favorable total interaction energy cannot be simply explained as a parameter that compensates for weak seed matching as had been previously speculated (Brennecke et al., 2005; Rajewsky, 2006). This result is consistent with the notion that weak seed matches need both compensatory 3' pairing and accessibility of the binding sites in order to attain the same stability that a strong site has by virtue of its well-matched seed region.

Although our results provide important insights into the global dynamics of miRNA-mediated regulatory networks during development, there are some significant limitations in the current data sets. For example, without examining the miRISCs in specific tissues, the miRNA and miRNA target profiles we generated were actually the average miRISC composition of all tissues. Consequently, miRNA:mRNA regulations in certain tissues, particularly in smaller tissues, could have been overlooked. In addition, given the existence of non-overlapping functions and mutually exclusive miRISC presence between AIN-1 and AIN-2, the data from experiments using AIN-2 IP could skew towards certain type of miRISCs. Finally, regulation of an mRNA by multiple different miRNAs at different stages will complicate the correlation between miRNAs and targets, and may elude the detection power of the pattern-matching approach based on the currently limited data points. Further analysis of the miRISCs in specific tissues will resolve the issue of spatially confined miRNA regulations and enhance the overall resolution of the pattern-matching approach.

Acknowledgements

We thank A. Fire for providing several plasmids, and A. Knyazev, W. Wood, T. Blumenthal, D. Xue, R. McIntosh and members of our laboratories for discussions. This work was supported by NIH grants GM47869 to M.H., F32GM087039 to M.C.H., GM34028 and GM066826 to V.A., and by the HHMI. Deposited in PMC for release after 6 months.

Supplementary material

Supplementary material for this article is available at <http://dev.biologists.org/cgi/content/full/136/18/3043/DC1>

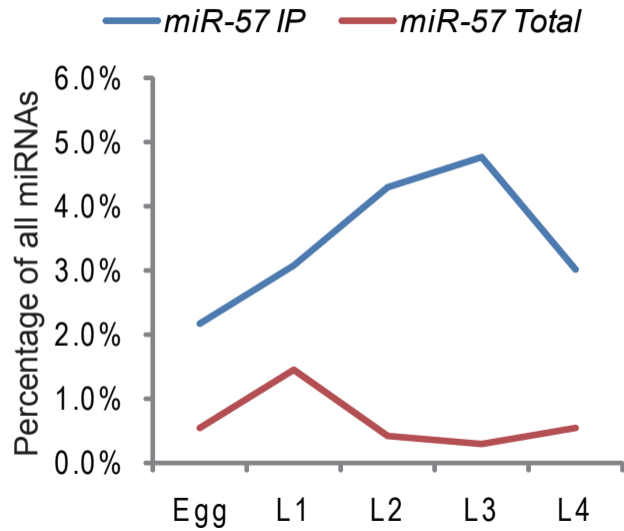
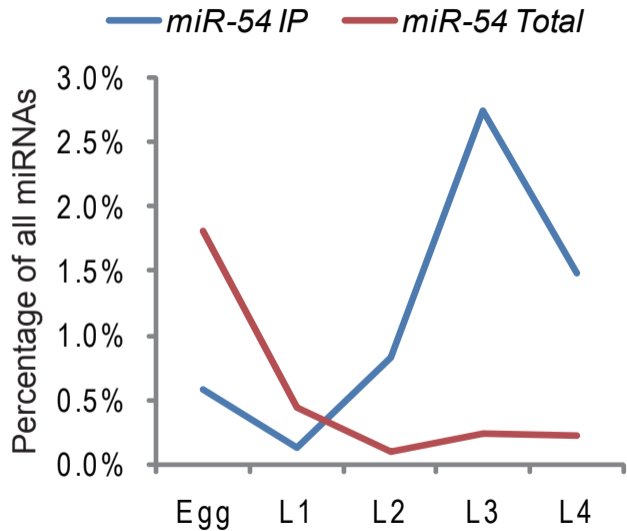
References

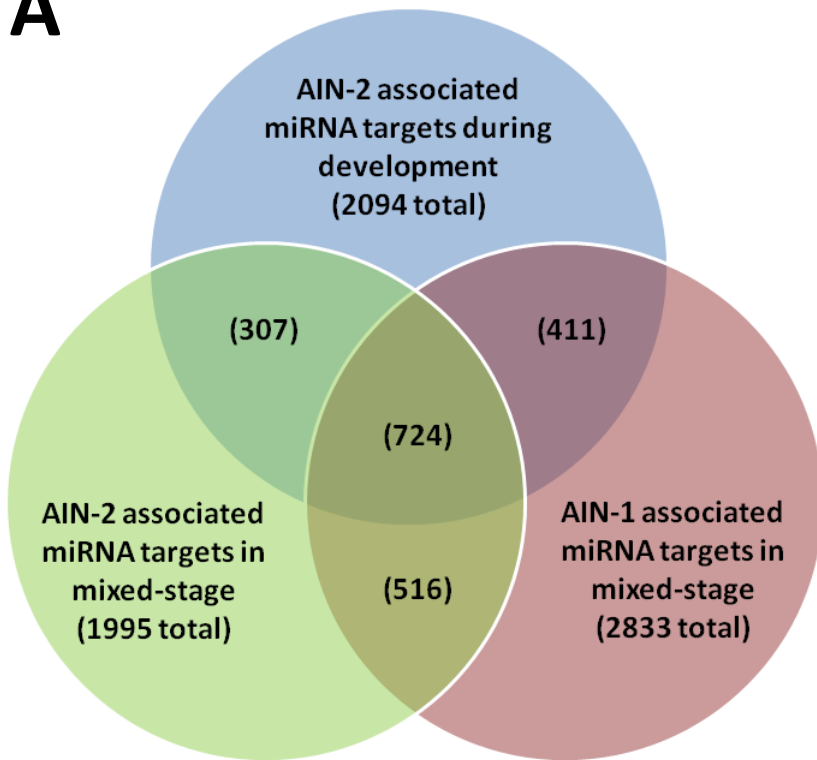
- Abbott, A. L., Alvarez-Saavedra, E., Miska, E. A., Lau, N. C., Bartel, D. P., Horvitz, H. R. and Ambros, V. (2005). The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans*. *Dev. Cell* **9**, 403-414.
- Abrahante, J. E., Daul, A. L., Li, M., Volk, M. L., Tennessen, J. M., Miller, E. A. and Rougvie, A. E. (2003). The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Dev. Cell* **4**, 625-637.
- Al-Shahrour, F., Diaz-Uriarte, R. and Dopazo, J. (2005a). Discovering molecular functions significantly related to phenotypes by combining gene expression data and biological information. *Bioinformatics* **21**, 2988-2993.
- Al-Shahrour, F., Minguez, P., Vaquerizas, J. M., Conde, L. and Dopazo, J. (2005b). BABELOMICS: a suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments. *Nucleic Acids Res.* **33**, W460-W464.
- Al-Shahrour, F., Minguez, P., Tarraga, J., Montaner, D., Alloza, E., Vaquerizas, J. M., Conde, L., Blaschke, C., Vera, J. and Dopazo, J. (2006). BABELOMICS: a systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Res.* **34**, W472-W476.
- Ambros, V. (2004). The functions of animal microRNAs. *Nature* **431**, 350-355.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T. et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25-29.
- Ason, B., Darnell, D. K., Wittbrodt, B., Berezikov, E., Kloosterman, W. P., Wittbrodt, J., Antin, P. B. and Plasterk, R. H. (2006). Differences in vertebrate microRNA expression. *Proc. Natl. Acad. Sci. USA* **103**, 14385-14389.
- Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20**, 1885-1898.
- Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I. and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125**, 1111-1124.
- Brennecke, J., Stark, A., Russell, R. B. and Cohen, S. M. (2005). Principles of microRNA-target recognition. *PLoS Biol.* **3**, e85.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.
- Bulik, D. A., Wei, G., Toyoda, H., Kinoshita-Toyoda, A., Waldrip, W. R., Esko, J. D., Robbins, P. W. and Selleck, S. B. (2000). sqv-3, -7, and -8, a set of genes affecting morphogenesis in *Caenorhabditis elegans*, encode enzymes required for glycosaminoglycan biosynthesis. *Proc. Natl. Acad. Sci. USA* **97**, 10838-10843.
- Bushati, N. and Cohen, S. M. (2007). microRNA functions. *Annu. Rev. Cell Dev. Biol.* **23**, 175-205.
- Chang, S., Johnston, R. J., Jr, Frokjaer-Jensen, C., Lockery, S. and Hobert, O. (2004). MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430**, 785-789.
- Conradt, B. and Horvitz, H. R. (1998). The *C. elegans* protein EGL-1 is required for programmed cell death and interacts with the Bcl-2-like protein CED-9. *Cell* **93**, 519-529.
- Didiano, D. and Hobert, O. (2006). Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat. Struct. Mol. Biol.* **13**, 849-851.
- Ding, L., Spencer, A., Morita, K. and Han, M. (2005). The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in *C. elegans*. *Mol. Cell* **19**, 437-447.
- Ding, X. C. and Grosshans, H. (2009). Repression of *C. elegans* microRNA targets at the initiation level of translation requires GW182 proteins. *EMBO J.* **28**, 213-222.
- Doench, J. G. and Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev.* **18**, 504-511.
- Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863-14868.
- Eulalio, A., Huntzinger, E. and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat. Struct. Mol. Biol.* **15**, 346-353.
- Freeze, H. H. and Aebi, M. (2005). Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr. Opin. Struct. Biol.* **15**, 490-498.
- Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillye, D. L., Fire, A., Ruvkun, G. and Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23-34.
- Grosshans, H., Johnson, T., Reinert, K. L., Gerstein, M. and Slack, F. J. (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.
- Hammell, M., Long, D., Zhang, L., Lee, A., Carmack, C. S., Han, M., Ding, Y. and Ambros, V. (2008). mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nat. Methods* **5**, 813-819.
- Herman, T. and Horvitz, H. R. (1999). Three proteins involved in *Caenorhabditis elegans* vulval invagination are similar to components of a glycosylation pathway. *Proc. Natl. Acad. Sci. USA* **96**, 974-979.
- Jakymiw, A., Lian, S., Eystathioy, T., Li, S., Satoh, M., Hamel, J. C., Fritzler, M. J. and Chan, E. K. (2005). Disruption of GW bodies impairs mammalian RNA interference. *Nat. Cell Biol.* **7**, 1267-1274.
- Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D. and Slack, F. J. (2005). RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635-647.
- Johnston, R. J. and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426**, 845-849.
- Kanehisa, M. (1997). A database for post-genome analysis. *Trends Genet.* **13**, 375-376.
- Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. and Hiraoka, M. (2006). From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* **34**, D354-D357.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U. and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat. Genet.* **39**, 1278-1284.
- Knight, S. W. and Bass, B. L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**, 2269-2271.
- Lai, E. C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.* **30**, 363-364.
- Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854.
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. and Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell* **115**, 787-798.
- Lewis, B. P., Burge, C. B. and Bartel, D. P. (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, M. W., Lau, N. C., Bartel, D. P. and Ruvkun, A. E. (2005). Regulatory mutations of mir-48, a *C. elegans* let-7 family MicroRNA, cause developmental timing defects. *Dev. Cell* **9**, 415-422.
- Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B. and Bartel, D. P. (2003). The microRNAs of *Caenorhabditis elegans*. *Genes Dev.* **17**, 991-1008.
- Lin, S. Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., Gottlieb, E. and Slack, F. J. (2003). The *C. elegans* hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Dev. Cell* **4**, 639-650.
- Liu, J., Rivas, F. V., Wohlschlegel, J., Yates, J. R., 3rd, Parker, R. and Hannon, G. J. (2005). A role for the P-body component GW182 in microRNA function. *Nat. Cell Biol.* **7**, 1261-1266.
- Long, D., Lee, R., Williams, P., Chan, C. Y., Ambros, V. and Ding, Y. (2007). Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* **14**, 287-294.
- Mishima, Y., Giraldez, A. J., Takeda, Y., Fujiwara, T., Sakamoto, H., Schier, A. F. and Inoue, K. (2006). Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. *Curr. Biol.* **16**, 2135-2142.
- Miska, E. A., Alvarez-Saavedra, E., Abbott, A. L., Lau, N. C., Hellman, A. B., McGonagle, S. M., Bartel, D. P., Ambros, V. R. and Horvitz, H. R. (2007). Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet.* **3**, e215.
- Moss, E. G., Lee, R. C. and 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.
- Okajima, T., Yoshida, K., Kondo, T. and Furukawa, K. (1999). Human homolog of *Caenorhabditis elegans* sqv-3 gene is galactosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **274**, 22915-22918.
- Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Muller, P. et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89.
- Pauley, K. M., Eystathiou, T., Jakymiw, A., Hamel, J. C., Fritzler, M. J. and Chan, E. K. (2006). Formation of GW bodies is a consequence of microRNA genesis. *EMBO Rep.* **7**, 904-910.
- Rajewsky, N. (2006). microRNA target predictions in animals. *Nat. Genet.* **38** Suppl, S8-S13.
- Rehwinkel, J., Behm-Ansmant, I., Gatfield, D. and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11**, 1640-1647.
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Ruvkun, A. E., Horvitz, H. R. and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901-906.
- Ritchie, W., Flamant, S. and Rasko, J. E. (2009). Predicting microRNA targets and functions: traps for the unwary. *Nat. Methods* **6**, 397-398.
- Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., Ge, H. and Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**, 1193-1207.
- Schafer, W. R. (2006). Neurophysiological methods in *C. elegans*: an introduction. In *WormBook* (ed. The *C. elegans* Research Community). doi/10.1895/wormbook.1.111.1, <http://www.wormbook.org>.
- Schneider, M. D., Najand, N., Chaker, S., Pare, J. M., Haskins, J., Hughes, S. C., Hobman, T. C., Locke, J. and Simmonds, A. J. (2006). Gawky is a component of cytoplasmic mRNA processing bodies required for early *Drosophila* development. *J. Cell Biol.* **174**, 349-358.
- Selbach, M., Schwanhaussner, B., Thierfelder, N., Fang, Z., Khanin, R. and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58-63.
- Simon, D. J., Madison, J. M., Conery, A. L., Thompson-Peer, K. L., Soskis, M., Ruvkun, G. B., Kaplan, J. M. and Kim, J. K. (2008). The microRNA miR-1 regulates a MEF-2-dependent retrograde signal at neuromuscular junctions. *Cell* **133**, 903-915.
- Stark, A., Brennecke, J., Russell, R. B. and Cohen, S. M. (2003). Identification of *Drosophila* MicroRNA Targets. *PLoS Biol.* **1**, E60.
- Stark, A., Brennecke, J., Bushati, N., Russell, R. B. and Cohen, S. M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* **123**, 1133-1146.
- Stefani, G. and Slack, F. J. (2008). Small non-coding RNAs in animal development. *Nat. Rev. Mol. Cell Biol.* **9**, 219-230.
- Tabara, H., Yigit, E., Siomi, H. and Mello, C. C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DEXH-box helicase to direct RNAi in *C. elegans*. *Cell* **109**, 861-871.
- Wightman, B., Ha, I. and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855-862.
- Xie, X., Lu, J., Kulbokas, E. J., Golub, T. R., Mootha, V., Lindblad-Toh, K., Lander, E. S. and Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338-345.
- Yoo, A. S. and Greenwald, I. (2005). LIN-12/Notch activation leads to microRNA-mediated down-regulation of Vav in *C. elegans*. *Science* **310**, 1330-1333.
- Zhang, L., Ding, L., Cheung, T. H., Dong, M. Q., Chen, J., Sewell, A. K., Liu, X., Yates, J. R., 3rd and Han, M. (2007). Systematic identification of *C. elegans* miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol. Cell* **28**, 598-613.

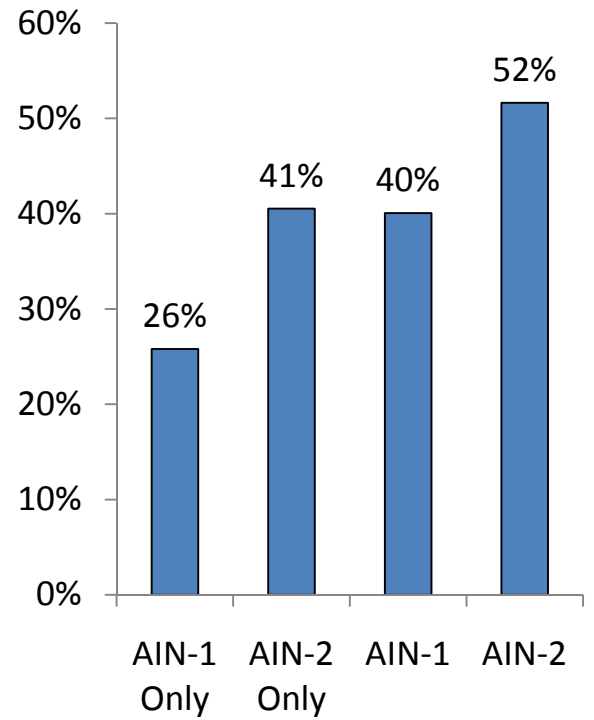
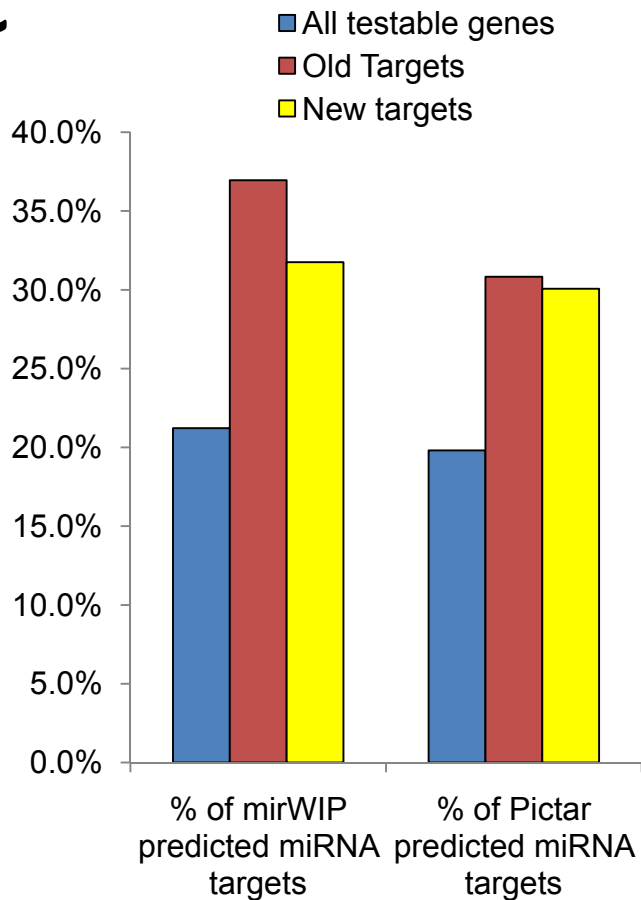
miRNA	Family	Seed	Cluster	Chromosome	Start	End	Strand
<i>lin-4</i>	<i>lin-4</i>	CCCUGA	3	II	5902220	5902313	+
<i>miR-237</i>	<i>lin-4</i>	CCCUGA	3	X	8154027	8154125	+
<i>let-7</i>	<i>let-7</i>	GAGGUA	3	X	14744094	14744192	-
<i>miR-48</i>	<i>let-7</i>	GAGGUA	3	V	14364390	14364487	-
<i>miR-241</i>	<i>let-7</i>	GAGGUA	3	V	14366166	14366261	-
<i>miR-35</i>	<i>miR-35</i>	CACCGG	2	II	11537524	11537620	+
<i>miR-36</i>	<i>miR-35</i>	CACCGG	2	II	11537629	11537725	+
<i>miR-37</i>	<i>miR-35</i>	CACCGG	2	II	11537749	11537846	+
<i>miR-38</i>	<i>miR-35</i>	CACCGG	2	II	11537846	11537943	+
<i>miR-39</i>	<i>miR-35</i>	CACCGG	2	II	11538005	11538091	+
<i>miR-40</i>	<i>miR-35</i>	CACCGG	2	II	11538100	11538191	+
<i>miR-42</i>	<i>miR-35</i>	CACCGG	1	II	11889727	11889821	+
<i>miR-46</i>	<i>miR-46</i>	GUCAUG	3	III	13660395	13660491	+
<i>miR-47</i>	<i>miR-46</i>	GUCAUG	3	X	13921173	13921269	+
<i>miR-49</i>	<i>miR-49</i>	AGCACC	1	X	9989216	9989312	+
<i>miR-83</i>	<i>miR-49</i>	AGCACC	3	IV	7841614	7841711	+
<i>miR-50</i>	<i>miR-50</i>	GAUAUG	1	I	1738636	1738734	+
<i>miR-90</i>	<i>miR-50</i>	GAUAUG	1	III	8873894	8873991	-
<i>miR-51</i>	<i>miR-51</i>	ACCCGU	1	IV	11025983	11026082	-
<i>miR-52</i>	<i>miR-51</i>	ACCCGU	1	IV	14033976	14034075	+
<i>miR-53</i>	<i>miR-51</i>	ACCCGU	1	IV	11027563	11027661	-
<i>miR-54</i>	<i>miR-51</i>	ACCCGU	1	X	13144917	13145015	-
<i>miR-55</i>	<i>miR-51</i>	ACCCGU	1	X	13144748	13144845	-
<i>miR-56</i>	<i>miR-51</i>	ACCCGU	1	X	13144612	13144708	-
<i>miR-57</i>	<i>miR-57</i>	ACCCUG	1	II	7850429	7850527	-
<i>miR-58</i>	<i>miR-58</i>	GAGAUC	3	IV	3233250	3233346	+
<i>miR-80</i>	<i>miR-58</i>	GAGAUC	3	III	8865177	8865274	-
<i>miR-81</i>	<i>miR-58</i>	GAGAUC	3	X	2431067	2431174	+
<i>miR-82</i>	<i>miR-58</i>	GAGAUC	3	X	2435253	2435349	-
<i>miR-63</i>	<i>miR-63</i>	AUGACA	3	X	17598678	17598781	-
<i>miR-64</i>	<i>miR-63</i>	AUGACA	1	III	2172836	2172945	+
<i>miR-65</i>	<i>miR-63</i>	AUGACA	1	III	2172990	2173089	+
<i>miR-66</i>	<i>miR-63</i>	AUGACA	1	III	2173096	2173194	+
<i>miR-73</i>	<i>miR-72</i>	GGCAAG	1	X	2368759	2368857	+
<i>miR-74</i>	<i>miR-72</i>	GGCAAG	2	X	2369037	2369133	+

Table S1. miRNAs with parallel AIN-2-association patterns tend to share the same seed sequence. The table lists all the miRNAs that were detected in the AIN-2 IP and also share the same 6nt seed sequence with at least one other miRNA detected in the AIN-2 IP. Cluster indicates the cluster group of the miRNA in Figure 1A, which represents the temporal miRISC-association pattern of the miRNA during development. The residing chromosome number, the genomic start/end position and the coding DNA strand of each miRNA gene are also listed.

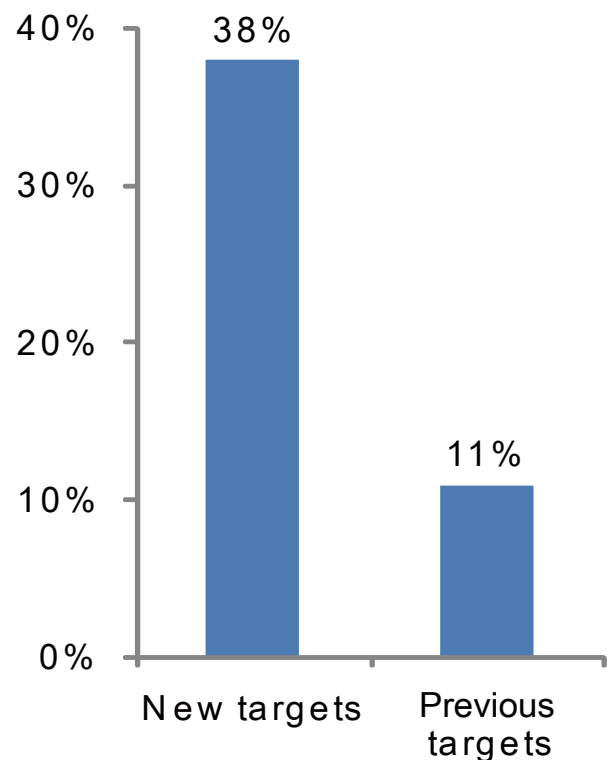


A**B**

Percentage of overlap with this study

**C****D**

Percentage of stage dependent miRNA targets



A

miRNA family/mRNA	Correlation P value
lin-4/lin-14a	0.000423
let-7/hbl-1	0.000481
let-7/pha-4	0.00416
let-7/T14B1.1	0.0272
let-7/daf-12	0.0478 *
lin-4/lin-28a	-0.847
let-7/let-60	-0.493
let-7/lss-4	0.53
let-7/die-1	-0.196

B