HHS Public Access

Author manuscript

*Cell Rep*. Author manuscript; available in PMC 2021 December 14.

Published in final edited form as:

*Cell Rep*. 2021 August 10; 36(6): 109517. doi:10.1016/j.celrep.2021.109517.

# Non-microRNA binding competitively inhibits LIN28 regulation

**Frederick E. Tan**[1,2,3,4], **Shashank Sathe**[1,2,3,4], **Emily C. Wheeler**[1,2,3], **Gene W. Yeo**[1,2,3,5,*]

[1]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA, USA

[2]Stem Cell Program, University of California, San Diego, La Jolla, CA, USA

[3]Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA, USA

[4]These authors contributed equally

[5]Lead contact

## SUMMARY

RNA binding protein (RBP) expression is finite. For RBPs that are vastly outnumbered by their potential target sites, a simple competition for binding can set the magnitude of post-transcriptional control. Here, we show that LIN28, best known for its direct regulation of *let*-7 miRNA biogenesis, is also indirectly regulated by its widespread binding of non-miRNA transcripts. Approximately 99% of LIN28 binding sites are found on non-miRNA transcripts, like protein coding and ribosomal RNAs. These sites are bound specifically and strongly, but they do not appear to mediate direct post-transcriptional regulation. Instead, non-miRNA sites act to sequester LIN28 protein and effectively change its functional availability, thus impeding the regulation of *let*-7 in cells. Together, these data show that the binding properties of the transcriptome broadly influence the ability of an RBP to mediate changes in RNA metabolism and gene expression.

## Graphical abstract



*Correspondence: geneyeo@ucsd.edu.

AUTHOR CONTRIBUTIONS

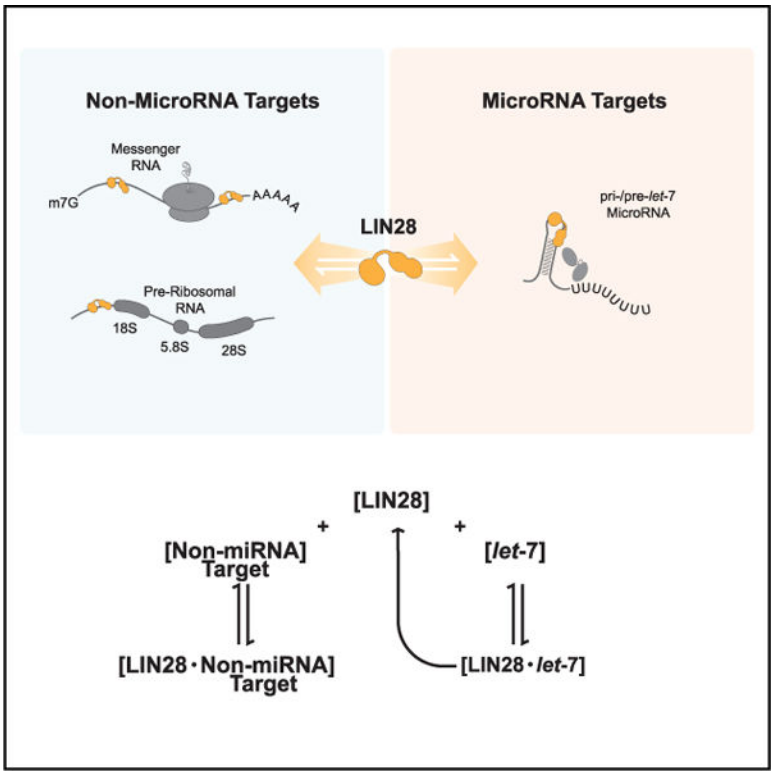
Conceptualization, F.E.T. and G.W.Y.; methodology, F.E.T., S.S., and E.C.W.; formal analysis, F.E.T. and S.S.; investigation, F.E.T.; writing – original draft, F.E.T.; writing – review & editing, F.E.T., S.S., and G.W.Y; supervision, G.W.Y.; funding acquisition, G.W.Y.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109517.

DECLARATION OF INTERESTS

G.W.Y. is co-founder, member of the Board of Directors, on the scientific advisory board, equity holder, and paid consultant for Locanabio and Eclipse Bio-Innovations. G.W.Y. is a visiting professor at the National University of Singapore. G.W.Y.'s interest(s) have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies. The authors declare no other competing financial interests.




## In brief

Tan et al. show that the majority of LIN28 binding sites are found on non-miRNA transcripts like coding and ribosomal RNA. Binding to non-miRNA sites sequesters LIN28 protein and competitively inhibits the regulation of miRNAs. Non-miRNA sites can thus modulate the strength of cellular miRNA activity in cells.

## INTRODUCTION

RNA binding protein (RBP) expression is a key variable in post-transcriptional control (Tan et al., 2019). When target binding sites in the transcriptome are present in excess (Jankowsky and Harris, 2015), changes in RBP expression level influence the competition for RNA binding, which can strongly impact gene expression. Crosslinking and immunoprecipitation (CLIP) is a mainstay method for examining the transcriptome-wide binding behavior of RBPs (Lee and Ule, 2018). However, many CLIP studies are typically performed for only a single expression condition, which can reveal the potential for broad targeting of different RNA types (Nostrand et al., 2018) but provides insufficient insight into which interactions are functionally important or which yield productive post-transcriptional regulation.

Alternatively, given sub-saturating RBP expression levels, CLIP could be performed at different RBP expression levels to capture changes in the steady-state occupancy of target sites. Changes in occupancy would reflect relative differences in binding preference, which may be directly related to the post-transcriptional regulation of a transcript. To examine the

utility of this approach, we performed CLIP on LIN28 to re-examine its dose-dependent regulation of gene expression. LIN28 is a centrally important post-transcriptional regulator of embryonic development, and changes in its expression level are thought to maintain gene expression programs that promote tissue growth and morphogenesis (Shinoda et al., 2013). LIN28 is best known for its ability to regulate miRNA biogenesis (Nam et al., 2011), but it also has the capacity to broadly bind other mammalian transcripts (Cho et al., 2012; Wilbert et al., 2012; Hafner et al., 2013), the purpose of which largely remains unclear.

Our exploration here of the LIN28 regulatory mechanism using enhanced CLIP (eCLIP) and additional biochemical assays reveal that target sites on miRNAs and non-miRNAs are bound by LIN28 with comparable strength. However, the binding of targets on non-miRNAs does not by itself confer large-scale direct post-transcriptional regulation. Rather, sequestration indirectly modulates gene expression by hindering LIN28's ability to regulate miRNAs. In our alternative model, target sites on non-miRNAs act to sequester LIN28 protein from miRNAs and potentiate *let*-7-dependent regulation.

## RESULTS

### Transcriptome-wide LIN28 binding with variable affinity

To examine the characteristics of RNA binding at different LIN28 concentrations, we performed eCLIP using cells stably expressing different levels of LIN28B (Figure 1A; Tables S1 and S2) (Van Nostrand et al., 2016). We showed previously that LIN28B expression in these cell lines (X1 through X4) spanned the physiological range and accounted for a 7-fold difference in LIN28B protein levels (Tan et al., 2019), as well as dose-dependent changes in mature *let*-7 miRNA production (Figure S1). These experiments revealed that the vast majority (~79%) of uniquely mapped CLIP sequencing reads from these cell lines were derived from ribosomal and protein coding RNAs (Figure 1B; Tables S3 and S4), whereas pre-/pri-miRNAs represented only a small minority of the total (<1%). Sequences from other non-coding RNAs, such as tRNAs or long non-coding RNAs (lncRNAs), and repetitive elements typically accounted for ~20% of uniquely mapped reads.

Most LIN28 binding sites (>60%) were detected in all conditions (Figure 1C), but the relative density of CLIP reads at each site changed dynamically as a function of LIN28B level. Because the sum of all RNA sequencing measurements (reads per million) within a library remains constant, a reduction in read density at any one location in the transcriptome must be counterbalanced by an increase at other locations. We expect that these increases or decreases in read density are directly related to the binding preference (relative affinity) of each site for LIN28 protein. As LIN28 levels are increased, equilibrium binding allows higher-affinity sites to approach saturation more quickly, reducing their overall contribution to a CLIP library. By contrast, binding at lower-affinity sites is less preferred by LIN28 but becomes more common when LIN28 is abundant. Lower-affinity sites thus account for a greater fraction of mapped reads at higher LIN28 expression levels. Changes in relative read density therefore reflect the redistribution of read density between sites with differing affinities and are consistent with the limiting availability of LIN28 protein in mammalian cells.

Among the most significant LIN28 binding sites, a small number of prominent regions on ribosomal RNA45S were found to lose read density as LIN28 expression was increased (Figure 1D). Given the relative abundance of ribosomal RNA, reads mapping to these sites accounted for the largest fraction of LIN28 binding in the mammalian transcriptome. Reads mapping to coding RNAs were also substantial, and read density at these target sites was found to change both positively and negatively with LIN28 expression (Figures 1E and 1F). Importantly, changes in read density could not be attributed to differences in RNA abundance, because the composition of the mammalian transcriptome is not affected by LIN28 expression (Tan et al., 2019).

For the conditions profiled in this study (X1 through X4), we observed the redistribution of CLIP reads from ribosomal to coding transcripts, indicating that higher-affinity sites on rRNA are preferred over lower-affinity sites scattered across thousands of coding transcripts (Figure 1B). Moreover, because this redistribution continues even at very high levels of LIN28 expression (X4), we can assume that lower-affinity binding is not completely saturated.

## LIN28 targets an RNA motif and structure

To further examine the significance of this binding behavior, we plotted the regression between read density and LIN28 protein level and computed a slope (μ) for each site, which was either negative (−μ), which we attribute to the loss of reads from a higher-affinity site, or positive (+μ), indicating the gain of reads at a lower-affinity site (Figure 2A). We discovered that the distribution of all slopes, encompassing 11,433 sites and their regressions, was strikingly unimodal and revealed a continuum of binding preferences (Figure 2B).

We explored whether specific molecular features could explain the range of observed affinities, such as the presence of unique RNA sequence motifs or structures. We first focused on the extreme ends of the distribution of slopes, which encompassed the highest (n = 1,592) and lowest (n = 933) affinity sites. LIN28 has two known RNA binding domains, a cold shock domain (CSD) and a 2xCCHC zinc knuckle domain (ZKD), that have been shown to make specific contacts with pri-/pre-*let*-7 miRNAs. We quantified the enrichment of ‘‘NGAU’’ and ‘‘GNNG’’ motifs in LIN28 binding sites, which are bound by the CSD and ZKD, respectively (Loughlin et al., 2011; Ustianenko et al., 2018). Although we did not observe the enrichment of ‘‘NGAU’’ in LIN28 binding sites compared with the sequence background (Figure 2C), more than half of the various combinations matching ‘‘GNNG’’ were enriched (Figure 2D). The level of enrichment for any ‘‘GNNG’’ variant was comparable between the highest and lowest affinity targets, indicating that the presence of ‘‘GNNG’’ may help recruit LIN28 to RNA but is likely not used to discriminate binding strength. Notably, among the more strongly enriched variants was ‘‘GGAG,’’ which is a conserved motif found in mammalian pri-/pre-*let*-7 miRNAs.

Apart from making sequence-specific contacts, the CSD can promote LIN28 binding through non-specific, electrostatic, and hydrophobic interactions (Schröder et al., 1995). Within the context of pri-/pre-miRNAs, the CSD loops single-stranded RNA around its charged surface and stabilizes the head of a miRNA hairpin. RNA structure may thus play a similarly important role at non-miRNA target sites. Indeed, by using the secondary

structure prediction algorithm RNAfold (Gruber et al., 2008), we discovered that higher- and lower-affinity LIN28 binding sites are distinguished by their ability to form stabilizing GC, GU, and AU base pairs (Figure 2E). This base pairing potential could allow LIN28 binding sites to form structures with low free energy compared with the sequence background. Importantly, affinity (μ) was correlated to the potential for forming stable structures (Figure 2F). The exact sequence of a target RNA was not as significant, because a random forest classifier designed to predict affinity based on the sequence composition of binding sites (6-mers) performed poorly, with an accuracy close to random selection (Figures 2F and 2G, left) (Breiman, 2001). RNA structure may also not be accurately represented by isolated 6-mers, because structures can form between sequences separated by long distances within a target site.

## Direct LIN28 binding is uncorrelated to target regulation

These data indicate that LIN28 binds thousands of non-miRNA sites with variable strength. The function of these binding sites, which are found predominantly on ribosomal and coding RNA (Figure 1A), still remains unclear. Previous studies have focused only on mRNAs and how LIN28 influences the production of protein, with many groups proposing a direct role for non-miRNA binding in modulating translation (Cho et al., 2012; Jin et al., 2011; Wilbert et al., 2012). However, in previous work, we showed that LIN28 regulation is strongly associated with the potential for miRNA regulation, with changes in gene expression correlated to the per gene enrichment of *let*-7 and non-*let*-7 miRNA target sites (Tan et al., 2019). To what extent does direct LIN28 binding of mRNA augment this post-transcriptional regulation?

If direct LIN28 binding confers additional post-transcriptional regulation, the degree of LIN28 occupancy on a transcript should exhibit some correlation to the observed change in gene expression. As revealed by polysome sequencing, however, we did not observe any specific relationship between overall CLIP read density and a transcript's enrichment in polysome fractions (Figure 3A). Importantly, the same changes in gene expression are observed, independent of any direct LIN28 binding (Figure 3B). These findings indicate that changes in protein translation do not appear to be directly influenced by LIN28 binding, and that changes in gene expression are still dominated by miRNA-mediated effects.

## Non-miRNA binding sites sequester LIN28 protein

LIN28 does not need to directly bind a transcript to impart post-transcriptional regulation. Non-miRNA binding sites can still influence gene expression indirectly by sequestering LIN28 protein from miRNAs. The vast number of non-miRNA sites on coding and ribosomal transcripts could act as binding decoys and could prevent LIN28 from regulating pre-/pri-miRNA biogenesis. For this to be true, non-miRNAs must be bound by LIN28 as strongly as miRNAs, which would allow them to be competitive inhibitors. Sequestration of LIN28 by this mechanism could buffer global changes in miRNA activity.

To test this hypothesis, we needed to disrupt the balance of miRNA and non-miRNA binding in a cell and examine the effect of this disruption on gene regulation. We constructed a non-coding RNA that could be expressed from an inducible transgene to bind and

sequester LIN28 protein (Figure 4A). This molecular sponge construct was modeled after a high-affinity target site detected in the human HMGA2 3′ UTR, which binds LIN28 preferentially at low concentrations (Figure 1D). We also incorporated pseudoknots at the 5′ and 3′ ends to confer nuclease resistance and to promote the accumulation of sponge in cells.

We confirmed that LIN28A can bind the sponge construct. We first validated previous experiments showing that LIN28A binds pre-*let*-7a miRNA with nanomolar affinity ($K_D$ ~25 nM) using an electrophoretic mobility shift assay (EMSA) (Wang et al., 2017) (Figure 4B). Given the propensity of LIN28A to form higher-order ribonucleoprotein complexes above 50 nM, which tend to aggregate in the well of a polyacrylamide gel, we also characterized this association using a filter binding assay (Rio, 2012) (Figure 4C). A filter binding assay employs a sandwich of three membranes through which a sample is passed under a vacuum, with each membrane screening for a specific type of protein-RNA interaction. The first layer (polyethersulfone [PES]) retains all higher-order aggregated ribonucleoprotein complexes. The second layer (nitrocellulose) retains free protein and any RNA bound to these proteins. The third layer (nylon) retains all free RNA. Using this system, we see a transition between unbound and bound pre-*let*-7a miRNA when LIN28A reaches a concentration of ~10 nM. Like the EMSA, LIN28A concentrations of 50 nM and above caused pre-*let*-7a miRNA to be retained in the PES layer as an aggregate, as shown by the lack of any signal on nitrocellulose or nylon.

Our LIN28 sponge construct serves as an effective competitor for LIN28A binding, as shown by the displacement of pre-*let*-7a miRNA from LIN28A when present at roughly equivalent concentrations (1 nM and above) (Figure 4D). Interestingly, a random short hairpin RNA construct used as a control was also found to displace pre-*let*-7a in a gel shift assay (Figure 4E), even though the same construct could not influence LIN28 binding when pre-*let*-7a was presented as a tethered bait (Figure S2A). This short hairpin construct contains ‘‘GNNG’’ sequences and is also distinguished by a well-defined hairpin structure, which could explain its ability to interact with LIN28A, albeit inconsistently. By contrast, repetitive and unstructured polyadenylic acid (polyA) with an average molecular weight of 250 kDa could not displace LIN28A from pre-*let*-7 miRNA (Figure 4D).

A filter binding assay largely reproduces these gel shift results, showing the depletion of probe from nitrocellulose and its accumulation on nylon in the presence of competitors (unlabeled *let*-7a, sponge, shRNA), but not in the presence of polyA RNA (Figure 4F). Interestingly, these data indicate the sponge may bind LIN28A more strongly than pre-*let*-7a miRNA. However, we cannot comment on the exact affinity of LIN28A for the sponge competitor because the target site from the human HMGA2 3′ UTR may potentially account for several different association sites. The generic binding motifs and the preference of LIN28 for structured RNA also permit the protective elements at the 5′ and 3′ ends to play a role in siphoning LIN28 protein. The sponge backbone (absent any LIN28 binding site) can compete as effectively as unlabeled pre-*let*-7a miRNA (Figure 4F). Nonetheless, the sponge appears to be specific for LIN28 as evidenced by significant changes in mature *let*-7 expression (Figure S2B), but we cannot exclude the possibility that the sponge (and its backbone) impacts other miRNA-associated RBPs.

### Sequestration of LIN28 protein alters miRNA regulation

We next examined whether the sponge construct, when expressed in cells, could be used to disrupt LIN28 activity. We stably expressed the sponge in mouse embryonic stem cells (V6.5 mESCs), a cell type that maintains high levels of Lin28A (Shyh-Chang and Daley, 2013), and confirmed that sponge expression enhanced the biogenesis of *let*-7 family miRNAs (Figure S2B). We created eCLIP sequencing libraries to monitor changes in transcriptome-wide Lin28A occupancy and polysome sequencing libraries to track changes in gene expression in response to sponge expression.

We detected 1,449 binding sites before sponge expression and 185 binding sites after sponge expression, with an overlap of 131 sites (71%) (Figure 5A). The reduction in detectable binding sites after sponge expression was expected because sequestration of LIN28 by the sponge would preferentially impact occupancy at lower-affinity sites, resulting in relatively more frequent binding at higher-affinity sites and a redistribution of eCLIP sequencing reads (Figure 5B). Indeed, we confirmed that binding sites detected after sponge expression had a greater potential to form structure (i.e., lower free energy) than sites detected in the absence of sponge (Figure 5C).

Importantly, by displacing LIN28 from pri-/pre-*let*-7 miRNAs, our sponge competitor caused an increase in cellular *let*-7 miRNA activity, as gauged using a fluorescent reporter with a fully complementary *let-7* miRNA target site (Figure 5D). These changes in miRNA activity permitted the separation of the *wild-type* (*WT*) mESCs transcriptome into two translational subpopulations (Figure 5E). In previous work (Tan et al., 2019), we showed these LIN28-dependent subpopulations are associated with genes that are either enriched for *let*-7 target sites (repressed genes) or genes that are enriched for all other non-*let*-7 miRNA target sites (activated genes). By contrast, in the absence of miRNA activity (i.e., *Dicer* knockout), we found that sponge expression had no effect on gene expression (Figure 5E). These data are consistent with the role of non-miRNA binding sites as competitors for LIN28 binding, which primarily affect gene expression by modulating miRNA pathway activity.

Sequestration of LIN28 protein indirectly regulates gene expression by enhancing *let*-7 miRNA biogenesis. The changes in gene expression observed in *WT* mESCs are consistent with an increase in *let*-7 miRNA activity, because transcripts enriched in *let*-7 family target sites are repressed (high rank) (Figure 5F), and given rate-limiting expression of miRNA pathway components, increased *let*-7 activity negatively impacts the ability of all other non-*let*-7 miRNAs to regulate their target genes. Consequently, the expression of transcripts enriched in non-*let*-7 miRNA target sites is enhanced (low rank) (Figure 5F). Together, these observations underscore the central importance of miRNAs in mediating LIN28 regulation and the ability of non-miRNA binding to modulate their activity.

## DISCUSSION

LIN28 is best known for its regulation of miRNA biogenesis, but miRNAs account for less than 1% of its interactions. The principal target of this regulation is *let*-7, and its suppression by LIN28 influences the expression of nearly every mRNA through global

changes in miRNA activity (Tan et al., 2019). However, the vast majority (~75%) of LIN28 binding sites are found on coding and ribosomal RNAs, but their functional relevance has remained unclear. Here, we reexamined transcriptome-wide binding by LIN28 at different LIN28 protein concentrations and discovered that non-miRNA targets are primarily acting as competitive inhibitors of LIN28 regulation. We showed that binding of non-miRNA targets does not result in direct post-transcriptional regulation of a bound transcript, but rather, because of their sheer abundance, they can actively sequester LIN28 from miRNAs. In doing so, non-miRNA sites indirectly influence gene expression by modulating miRNA activity. MicroRNA families that will be negatively impacted by the sequestration of LIN28 are the *let*-7 family, as well as others such as miR-1 and miR-9, which are also bound and regulated by LIN28 (Nowak et al., 2014; Rau et al., 2011). Conversely, the activity of all other miRNAs that are not bound or regulated by LIN28 are positively regulated by this mechanism.

As competitive inhibitors, the extent to which non-miRNAs post-transcriptionally regulate gene expression depends on their abundance. Because the expression of ribosomal and coding transcripts can vary significantly between cells (Khajuria et al., 2018), the magnitude of sequestration may change dynamically throughout development and disease. Our results indicate that changes in gene expression are still correlated most to differences in miRNA activity and not to differences in LIN28 binding frequency, and together with our previous work on the LIN28 mechanism (Tan et al., 2019), we can now show that at least three variables determine LIN28 regulation: the relative abundance of *let*-7 miRNA (and other miRNA binders of LIN28), the concentration of LIN28 protein, and the binding characteristics of the target transcriptome.

These findings may provide an explanation for why LIN28 levels are dynamically regulated in different cell types, which must balance the post-transcriptional control of *let*-7 with the binding characteristics of a changing transcriptome. Alternatively, the existence of non-miRNA binding sites may have been selected, during the molecular evolution of mammalian genomes, for their ability to appropriately restrain LIN28 activity. Because if non-miRNA binding sites did not exist, small changes in LIN28 level could easily saturate the amount of pri-/pre-*let*-7 miRNA present in a cell, leading to switch-like transitions between *let*-7 ‘‘ON’’ and *let*-7 ‘‘OFF’’ gene expression states (Bissels et al., 2009; Hafner et al., 2013; Powers et al., 2016). Our electromobility shift and filter binding assays confirm that, in the absence of competitor RNA, a sharp transition between bound and unbound pre-*let*-7 miRNAs occurs at a LIN28 concentration between 10 and 25 nM. Sequestration of LIN28 by non-miRNA sites may thus permit the regulation of pri-/pre-*let*-7 miRNA biogenesis to exhibit greater dose dependency, which could be critical for the physiological discrimination of *let*-7 target genes based on pairing homology and binding strength (Grimson et al., 2007; Shu et al., 2012; Wu et al., 2015).

Together, these studies reveal that the binding of an RBP to a transcript may not translate into any direct post-transcriptional regulation. Although many high-throughput studies attempt to relate what and where an RBP binds to a specific biological function, these data provide support that this assumption is not always correct. Functional characterization of any RBP should apply a more balanced approach that takes into account known biochemistry

and the potential existence of uncharacterized biology. High-throughput assays that measure global changes in regulation, such as differences in RNA stability (via RNA sequencing), translation (via polysome sequencing [Polysomeseq] or ribosome profiling [Ribo-seq]) or cellular localization (via fractionation sequencing [Fractionation-seq]), can together, with eCLIP, provide a more complete assessment of RBP function.

## STAR★METHODS

### RESOURCE AVAILABILITY

**Lead contact—**Further information and requests for reagents may be directed to and will be fulfilled by the lead contact, Gene Yeo (geneyeo@ucsd.edu).

**Materials availability—**This study did not generate new unique reagents. Nonetheless, any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

**Data and code availability—**RNA sequencing data have been deposited at GEO and are publicly available as of the date of publication under accession number: GSE178259. Unprocessed western blot images associated with Figure 4 were deposited on Mendeley at: https://dx.doi.org/10.17632/g5sznxfzfw.1. This paper does not report any original code. Nonetheless, any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

**Cell lines and culture—**HEK293 stably expressing variable levels of LIN28B from a stably integrated PiggyBAC transgene were previously derived (Tan et al., 2019). V6.5 mESCs were initially thawed on feeders and were gradually acclimated to gelatin only culture over 2 passages. *Dicer* knockout mESCs were kindly donated by Dr. Anthony Leung (JHU), with methods used for their derivation previously described (Calabrese et al., 2007). *Dicer* knockout mESCs were similarly acclimated to gelatin culture over 2 passages. PiggyBACs for LIN28 sponge expression were introduced into V6.5 and *Dicer* knockout mESCs via transfection and cells with stable expression were sorted in bulk by flow cytometry. Complete mESC media included 15% FBS, NEAA, Na-Pyruvate and ESGRO LIF (1000U/ml) in High Glucose DMEM.

### METHOD DETAILS

**Polysome fractionation—**To obtain crude lysates, cell cultures were washed once with PBS containing Cycloheximide (CHX; 100ug/ml), harvested by cell scraping and then lysed on ice using 20mM Tris HCl pH7.4, 150mM NaCl, 5mM MgCl2, 1mM DTT with 1% Triton-X + Protease Inhibitors + RNase inhibitors + CHX (100ug/ml). Nuclei and debris were separated from crude lysate by brief centrifugation at 15,000x*g* at 4°C. Sucrose gradients (10%–50%) were prepared in 20mM Tris HCl pH7.4, 150mM NaCl, 5mM MgCl2, 1mM DTT + RNase inhibitors + CHX (100ug/ml) using a Biocomp Model 108 gradient master. Crude cellular lysates were then loaded onto gradients and separated by centrifugation at 110,000x*g*, 3 hours at 4°C and fractionated into 0.5mL aliquots

using a Biocomp Model 152 Piston Fractionator. Polysome fractions (typically fractions #10 through #20) were pooled and RNA extraction/purification was performed for the preparation of sequencing libraries.

**Enhanced Crosslinking and Immunoprecipitation—**Approximately $1 \times 10^7$ cells still attached to culture plates were crosslinked using UV irradiation (254nm, 4000J/$cm^2$) to capture instances of protein/RNA binding. Cell cultures were scraped in PBS, pelleted and snap frozen. Crosslinked cells were lysed in iCLIP lysis buffer (50 mM Tris-HCl at pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, 1:200 protease inhibitor cocktail III) supplemented with 440 units of murine RNase inhibitor and subjected to brief RNase I digestion (5 min) prior to immunoprecipitation with either anti-LIN28B (Cell Signaling, #11965) or anti-LIN28A (Cell Signaling, #8706). LIN28/RNA complexes were isolated using magnetic beads (M-280 sheep anti-Rabbit IgG Dynabeads or M-280 sheep anti-Mouse IgG, Thermo) and beads were subjected to stringent wash in high-salt buffer (50 mM Tris-HCl at pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) and low-salt buffer (20 mM Tris-HCl at pH 7.4, 10 mM MgCl2, 0.2% Tween-20). Barcoded RNA adapters were ligated to the 3′ end. Samples were then run on 4%–12% Bis-Tris Protein Gels, transferred to nitrocellulose membranes, and the region from 25 to 100 kDa corresponding to LIN28A/B and LIN28A/B–RNA complexes was excised and proteinase K-treated to isolate only RNA. Isolated RNA fragments were reverse transcribed and excess oligonucleotides were removed. A second DNA adaptor containing a random-mer of five or ten random bases at its 5′ end was ligated to the 3′ end of cDNA fragments. Fragments were then PCR amplified and size selected on agarose gel and pooled for high-throughput sequencing.

**Electrophoretic Mobility Shift Assay—**Sample reactions consisted of 1nM of Biotin labeled *let*-7a RNA was incubated with LIN28-TAT recombinant protein (Peprotech) at various concentrations in 20mM Tris pH7.5, 50mM KCl, 5mM $MgCl_2$, 1mM DTT, 5% Glycerol on ice for 2 hours. Pre-casted 10% TBE polyacrylamide gels were run without any applied samples in 0.5x TBE for 1 hour, 250V at 4°C. Loading dye (2uL of 50% glycerol/ 0.01% bromophenol blue) was added to all samples and mixtures were profiled on pre-run gels, 150V at 4°C. Samples were transferred to Hybond Nylon-N+ (GE Lifesciences) in 0.5x TBE for 1 hour, at 150V at 4°C. Membranes were crosslinked using a Stratalinker (4000J). Blots were visualized with Streptavidin HRP (Chemiluminescent Nucleic Acid Detection Module, Thermo-Pierce).

**Filter Binding Assay—**Following the protocol in Rio (2012), a sandwich of three membranes was assembled in a dot blot apparatus (Biorad) consisting of a top layer of Polyethersulfone (Millipore PES, 0.45um pore size), middle layer 100% Nitrocellulose (GE, Hybond ECL Nitrocellulose) and bottom layer Nylon (GE, Hybond Nylon-N+). Membranes were washed twice with 20mM Tris, 200mM KCl prior before and after the application of samples. Membranes were crosslinked using a Stratalinker (4000J). Blots were visualized with Streptavidin HRP (Chemiluminescent Nucleic Acid Detection Module, Thermo-Pierce).

**Pulldown Assay—**Biotin conjugated pre-let-7a miRNA (10 femtomoles) was tethered to Streptavidin-C1 MyOne magnetic beads. Beads were resuspended in 20mM Tris pH7.5, 50mM KCl, 5mM $MgCl_2$, 1mM DTT, 5% Glycerol containing 100nM LIN28A-TAT (Peprotech) and a variable amount of competitor (*let*-7a, sponge construct, shRNA or polyA RNA; 5pM, 50pM, 500pM or 5nM). Beads were washed twice with 20mM Tris pH7.5, 50mM KCl, 5mM MgCl2 and 5% Glycerol and resuspended in 10ul of 1x LDS Sample loading buffer (Thermo) in RIPA. Bead resuspensions were boiled at 95°C for 5 minutes and profiled for the retention of LIN28A protein.

**Antibodies—**LIN28 Antibody (Proteintech, #11724–1), LIN28A (using Cell Signaling, #8706), LIN28B (using Cell Signaling, #11965).

**Quantitative PCR of mature miRNAs—**Total RNA was extracted from controls and samples using Trizol reagent (Invitrogen). Mature miRNAs were reverse transcribed using the miSCRIPT II RT Kit (QIAGEN) and their abundance quantified using miSCRIPT SYBR Green PCR Kit (QIAGEN) and suitable miSCRIPT primer assay for each target miRNA of interest.

**Previously generated datasets—**Polysome fractionation data from Tan et al. (2019) was used for comparison with eCLIP libraries generated from the same cell lines derived in the previous study.

## QUANTIFICATION AND STATISTICAL ANALYSIS

**Change in polysome enrichment -—**Only transcripts with RPKM > 1 were considered. All statistics were computed using length and depth normalized measurements (RPKM) and adjusted for sequencing bias using SIRV Set3 Spike-Ins (Lexogen). The change in polysome enrichment of any sample condition [ i ] relative to any control condition [ o ] can be represented by a ratio of ratios. More specifically, we have ratios representing the polysome enrichment in condition P[ i ] is normalized by polysome enrichment in control condition P [ o ].

$$P[i] = \frac{\text{Polysome RNA}[i]}{\text{Total RNA}[i]}, \quad P[o] = \frac{\text{Polysome RNA}[KO]}{\text{Total RNA}[KO]}$$

The change in polysome enrichment is the ratio P [ i ] / P [ o ].

**Random Forest Classifiers—**We developed Random Forest Classifiers to examine whether the sequence characteristics of LIN28 binding sites could be used to distinguish high from low affinity binding locations (Breiman, 2001). More specifically, these Classifiers were trained and tested on the co-variance between 6-mers found in high and low binding sites and the relative affinity of each site, represented by the regression formed by change in eCLIP read density (μ).

**Data and Software Availability—**Computer code used for bioinformatics processing and analysis of RNA sequencing datasets will be made available upon request. The datasets generated in the current study will be deposited in the Gene Expression Omnibus.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

We would like to thank Dr. Anthony Leung and Dr. Mauro Calabrese for technical assistance and reagents related to *Dicer* null mESCs. F.E.T. was supported by a postdoctoral research fellowship (Grant 129547-PF-16-060-01-RMC) from the American Cancer Society. E.C.W. was supported by grants from the University of California, San Diego, Genetics Training Program (T32, GM008666) and the NSF Graduate Research Fellowship Program. G.W.Y. was partially supported by grants from the NIH (HG007005 and HG004659).

## REFERENCES

Bissels U, Wild S, Tomiuk S, Holste A, Hafner M, Tuschl T, and Bosio A (2009). Absolute quantification of microRNAs by using a universal reference. RNA 15, 2375–2384. [PubMed: 19861428]

Breiman L (2001). Random Forests. Mach. Learn 45, 5–32.

Calabrese JM, Seila AC, Yeo GW, and Sharp PA (2007). RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. Proc. Natl. Acad. Sci. USA 104, 18097–18102. [PubMed: 17989215]

Cho J, Chang H, Kwon SC, Kim B, Kim Y, Choe J, Ha M, Kim YK, and Kim VN (2012). LIN28A is a suppressor of ER-associated translation in embryonic stem cells. Cell 151, 765–777. [PubMed: 23102813]

Grimson A, Farh KK-H, Johnston WK, Garrett-Engele P, Lim LP, and Bartel DP (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell 27, 91–105. [PubMed: 17612493]

Gruber AR, Lorenz R, Bernhart SH, Neuböck R, and Hofacker IL (2008). The Vienna RNA websuite. Nucleic Acids Res 36, W70–W74. [PubMed: 18424795]

Hafner M, Max KEA, Bandaru P, Morozov P, Gerstberger S, Brown M, Molina H, and Tuschl T (2013). Identification of mRNAs bound and regulated by human LIN28 proteins and molecular requirements for RNA recognition. RNA 19, 613–626. [PubMed: 23481595]

Jankowsky E, and Harris ME (2015). Specificity and nonspecificity in RNA-protein interactions. Nat. Rev. Mol. Cell Biol 16, 533–544. [PubMed: 26285679]

Jin J, Jing W, Lei X-X, Feng C, Peng S, Boris-Lawrie K, and Huang Y (2011). Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes. Nucleic Acids Res 39, 3724–3734. [PubMed: 21247876]

Khajuria RK, Munschauer M, Ulirsch JC, Fiorini C, Ludwig LS, McFarland SK, Abdulhay NJ, Specht H, Keshishian H, Mani DR, et al. (2018). Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis. Cell 173, 90–103.e19. [PubMed: 29551269]

Lee FCY, and Ule J (2018). Advances in CLIP Technologies for Studies of Protein-RNA Interactions. Mol. Cell 69, 354–369. [PubMed: 29395060]

Loughlin FE, Gebert LFR, Towbin H, Brunschweiger A, Hall J, and Allain FH-T (2011). Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. Nat. Struct. Mol. Biol 19, 84–89. [PubMed: 22157959]

Nam Y, Chen C, Gregory RI, Chou JJ, and Sliz P (2011). Molecular basis for interaction of let-7 microRNAs with Lin28. Cell 147, 1080–1091. [PubMed: 22078496]

Nostrand ELV, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen J-Y, Cody NAL, Dominguez D, et al. (2018). A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. bioRxiv

Nowak JS, Choudhury NR, de Lima Alves F, Rappsilber J, and Michlewski G (2014). Lin28a regulates neuronal differentiation and controls miR-9 production. Nat. Commun 5, 3687. [PubMed: 24722317]

Powers JT, Tsanov KM, Pearson DS, Roels F, Spina CS, Ebright R, Seligson M, de Soysa Y, Cahan P, Theißen J, et al. (2016). Multiple mechanisms disrupt the let-7 microRNA family in neuroblastoma. Nature 535, 246–251. [PubMed: 27383785]

Rau F, Freyermuth F, Fugier C, Villemin J-P, Fischer M-C, Jost B, Dembele D, Gourdon G, Nicole A, Duboc D, et al. (2011). Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. Nat. Struct. Mol. Biol 18, 840–845. [PubMed: 21685920]

Rio DC (2012). Filter-Binding Assay for Analysis of RNA–Protein Interactions. Cold Spring Harb. Protoc 2012, 1078–1081. [PubMed: 23028069]

Schröder K, Graumann P, Schnuchel A, Holak TA, and Marahiel MA (1995). Mutational analysis of the putative nucleic acid-binding surface of the cold-shock domain, CspB, revealed an essential role of aromatic and basic residues in binding of single-stranded DNA containing the Y-box motif. Mol. Microbiol 16, 699–708. [PubMed: 7476164]

Shinoda G, Shyh-Chang N, Soysa TY, Zhu H, Seligson MT, Shah SP, Abo-Sido N, Yabuuchi A, Hagan JP, Gregory RI, et al. (2013). Fetal deficiency of lin28 programs life-long aberrations in growth and glucose metabolism. Stem Cells 31, 1563–1573. [PubMed: 23666760]

Shu J, Xia Z, Li L, Liang ET, Slipek N, Shen D, Foo J, Subramanian S, and Steer CJ (2012). Dose-dependent differential mRNA target selection and regulation by let-7a-7f and miR-17–92 cluster microRNAs. RNA Biol 9, 1275–1287. [PubMed: 22995834]

Shyh-Chang N, and Daley GQ (2013). Lin28: primal regulator of growth and metabolism in stem cells. Cell Stem Cell 12, 395–406. [PubMed: 23561442]

Tan FE, Sathe S, Wheeler EC, Nussbacher JK, Peter S, and Yeo GW (2019). A transcriptome-wide translational program defined by LIN28B expression level. Mol. Cell 73, 304–313.e3. [PubMed: 30527666]

Ustianenko D, Chiu H-S, Treiber T, Weyn-Vanhentenryck SM, Treiber N, Meister G, Sumazin P, and Zhang C (2018). LIN28 Selectively Modulates a Subclass of Let-7 MicroRNAs. Mol. Cell 71, 271–283.e5. [PubMed: 30029005]

Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K, et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods 13, 508–514. [PubMed: 27018577]

Wang L, Nam Y, Lee AK, Yu C, Roth K, Chen C, Ransey EM, and Sliz P (2017). LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. Cell Rep 18, 2664–2675. [PubMed: 28297670]

Wilbert ML, Huelga SC, Kapeli K, Stark TJ, Liang TY, Chen SX, Yan BY, Nathanson JL, Hutt KR, Lovci MT, et al. (2012). LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. Mol. Cell 48, 195–206. [PubMed: 22959275]

Wu L, Nguyen LH, Zhou K, de Soysa TY, Li L, Miller JB, Tian J, Locker J, Zhang S, Shinoda G, et al. (2015). Precise let-7 expression levels balance organ regeneration against tumor suppression. eLife 4, e09431. [PubMed: 26445246]

## Highlights

- LIN28 directly and indirectly regulates miRNA activity
- Most LIN28 binding sites are found on non-miRNAs
- Non-miRNAs and miRNAs are bound with comparable affinity
- Non-miRNA binding sites sequester and inhibit LIN28 activity

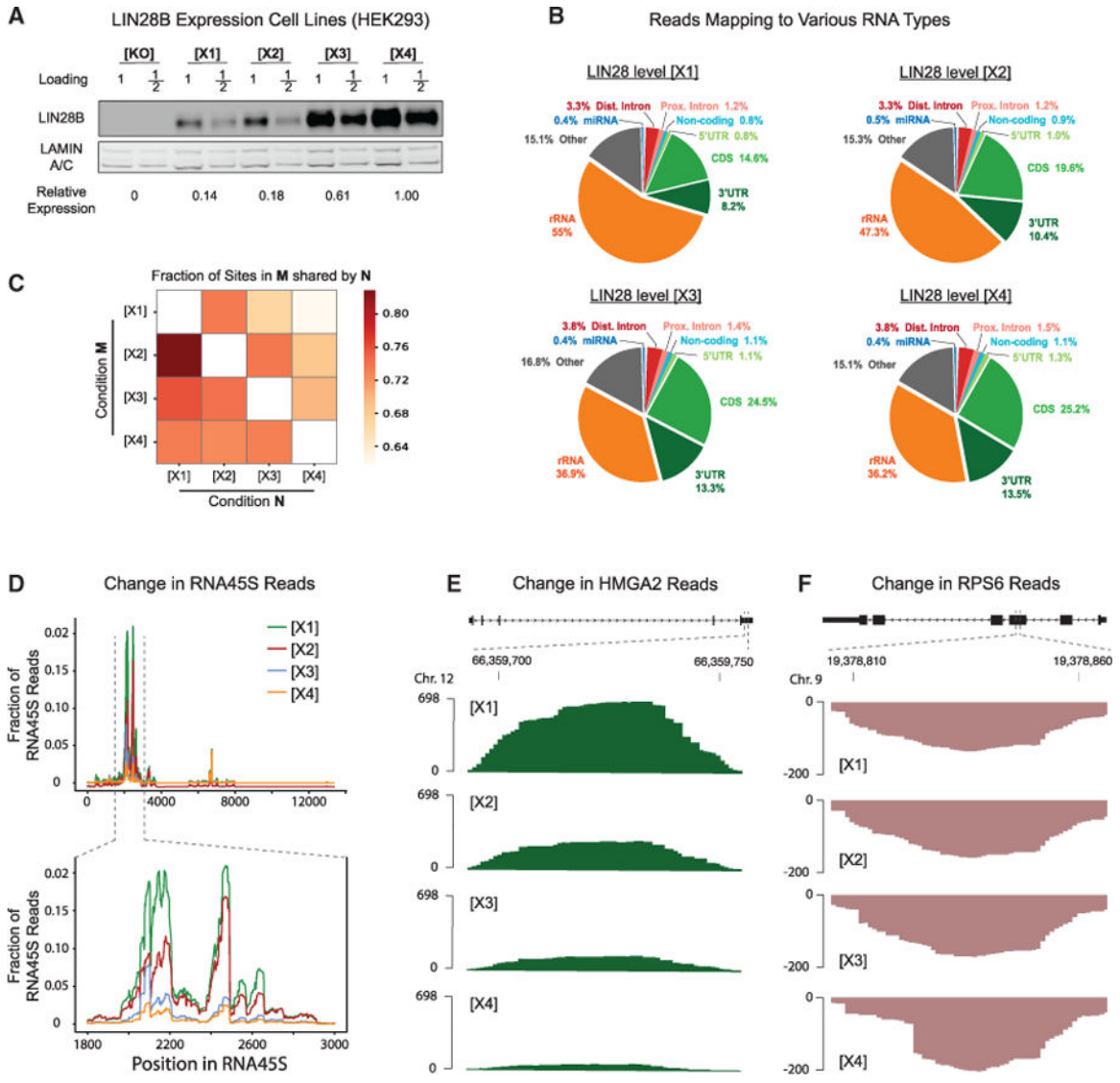



**Figure 1. eCLIP sequencing at different LIN28B concentrations in HEK293**

(A) Western blot showing transgene-mediated LIN28B expression in stable cell lines [X1] through [X4], relative to HEK293 LIN28B knockout (KO) cell line.

(B) Pie charts showing the allocation of reads to different RNA classes.

(C) Binding sites that are shared between different LIN28B expression conditions. The heatmap compares the sites detected in condition M (vertical axis) with the sites found in condition N (horizontal axis).

(D) Reads mapping to RNA45S, which was the only ribosomal transcript bound by LIN28B.

(E) Changes in read density at a binding site in the HMGA2 3′ UTR.

(F) Changes in read density at a binding site in the coding region of RPS6.

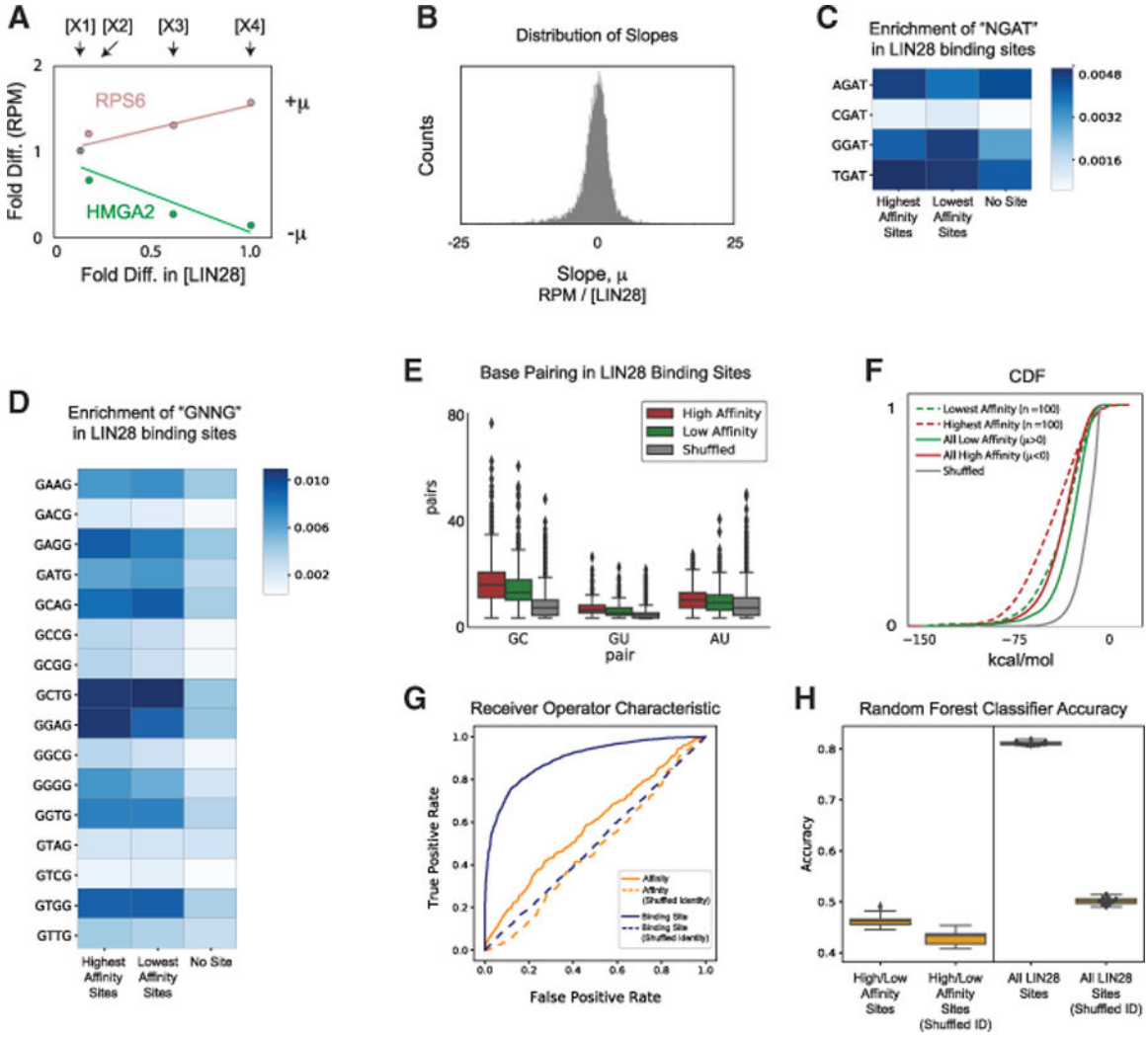



**Figure 2. Characteristics of LIN28B binding on non-miRNAs**

(A) Changes in read density at any binding site can be plotted as a function of LIN28B expression level, yielding a regression and slope. The slope is correlated to strength of LIN28B binding.

(B) Distribution of slopes (μ), encompassing 11,433 target sites.

(C) Frequency of known LIN28 motifs, ‘‘UGAU,’’ in eCLIP-identified binding sites.

(D) Frequency of known LIN28 motifs, ‘‘GNNG,’’ in CLIP-identified LIN28 binding sites.

(E) Boxplots showing the distribution of GC, GU, AU, and GG pairs found in RNA structures formed in LIN28 binding sites.

(F) Cumulative distribution functions showing the minimum free energy of the ensemble for all higher-affinity binding sites ($m < 0$, solid red), the 100 highest-affinity binding sites (dashed red), all lower-affinity sites ($m > 0$, solid green), and the 100 lowest affinity binding sites (dashed green) compared with randomly chosen sequences outside of eCLIP-identified binding sites (gray).

(G) Receiver-operator characteristic for two random forest classifiers that were developed to predict the affinity of a LIN28 binding site (blue) or the location of a LIN28 binding site (yellow). Predictions against data with shuffled identifiers were used as a control (dashed). (H) Boxplots showing the accuracy of random forest classifier predictions compared with performance on a dataset with shuffled identifiers.

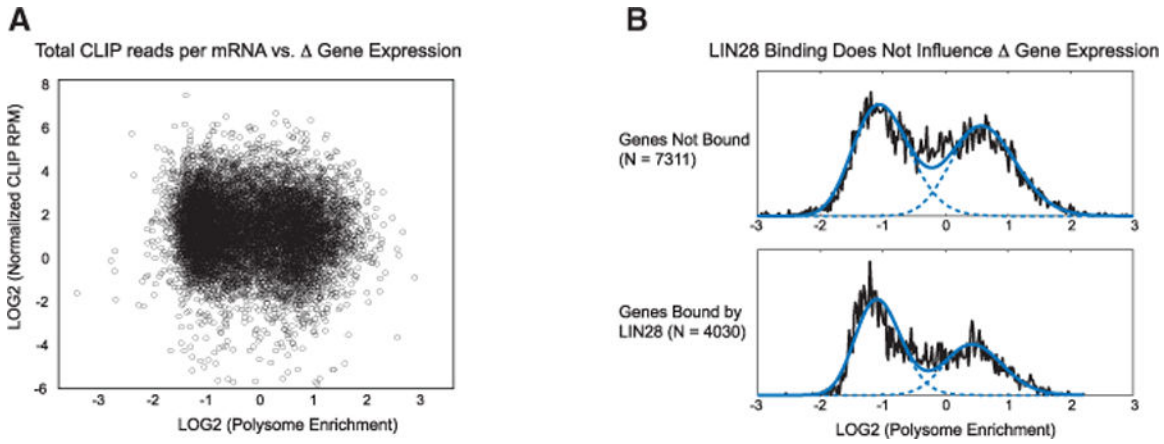



**Figure 3. Direct LIN28B binding and LIN28B-dependent changes in gene expression are uncorrelated**

(A) Scatterplot showing that total mapped reads per transcript normalized for abundance (normalized eCLIP RPM) is uncorrelated to the fold change in polysome enrichment per transcript, at LIN28 expression level [X4] compared with control HEK293 that do not express LIN28B.

(B) The same dataset plotted in (A) but split into gene sets that are bound or not bound by LIN28B and plotted as distributions relative to the fold change polysome enrichment per transcript at LIN28B expression level [X4] compared with control HEK293 that do not express LIN28B.

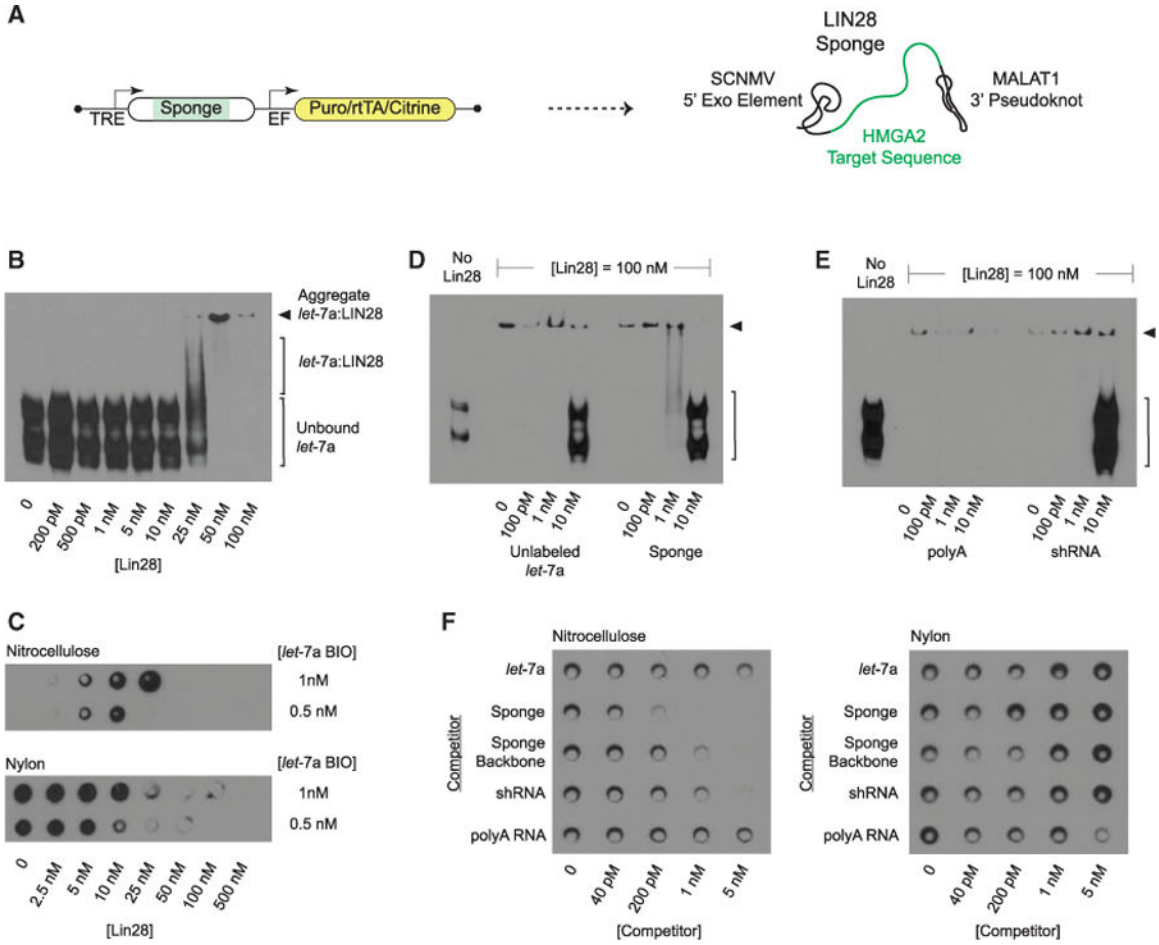



**Figure 4. An RNA binding protein sponge construct competes for LIN28 binding**

(A) Design of gene expression construct used to express an RNA construct that can sequester LIN28A protein (sponge).

(B) Electrophoretic mobility shift assay showing the association of LIN28A with biotin-labeled *let*-7a with a $K_D$ of ~25 nM.

(C) Filter-binding assay showing the association of LIN28A with biotin-labeled *let*-7a with a $K_D$ of ~10 nM.

(D) Electrophoretic mobility shift assay showing LIN28A in complex with biotin-labeled *let*-7a and its dissociation by unlabeled *let*-7a competitor and sponge construct competitor.

(E) Electrophoretic mobility shift assay showing LIN28A in complex with biotin-labeled *let*-7a and its dissociation by an unrelated shRNA construct designed to target YTHDF1, but not by polyadenylic acid (polyA RNA).

(F) Filter-binding assay showing the dissociation of biotin-labeled *let*-7a from LIN28A in the presence of various competitors (*let*-7a, Sponge, Sponge Backbone, shRNA, and polyA RNA).

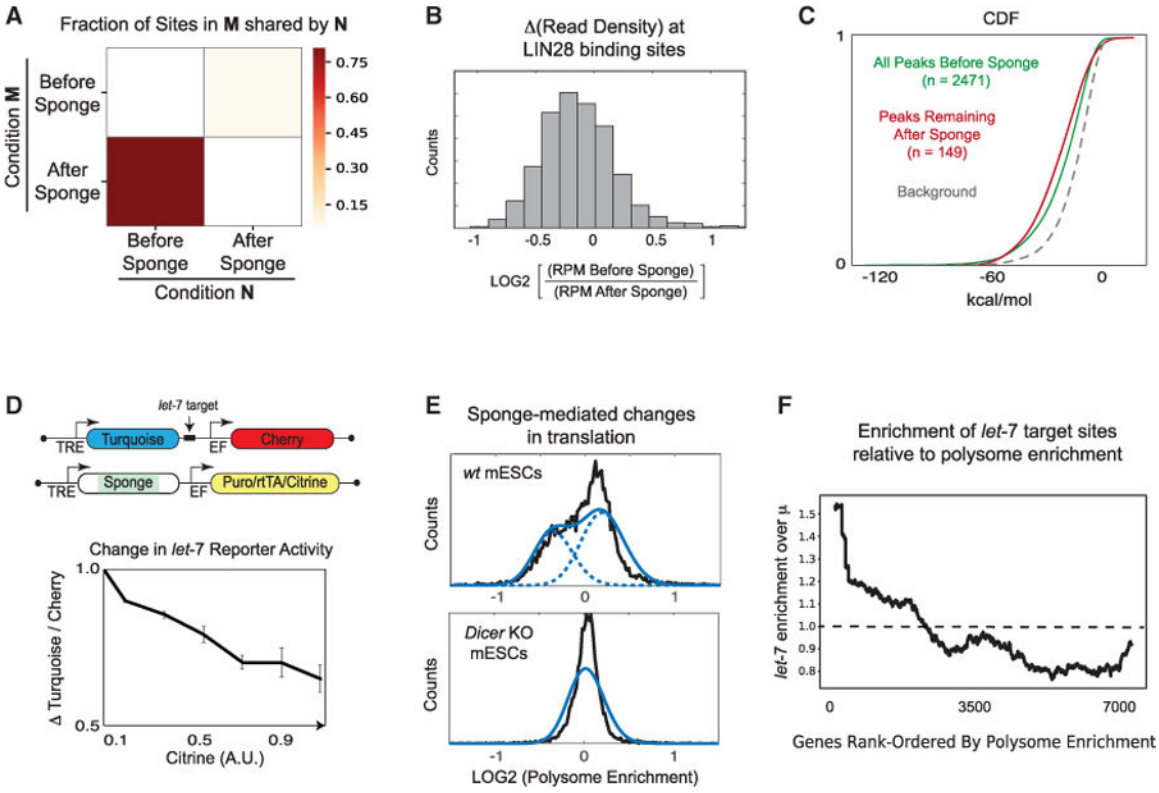



**Figure 5. Expression of a LIN28 sponge construct alters gene expression in mouse embryonic stem cells**

(A) Binding sites that are shared before and after sponge expression. The heatmap compares the sites detected in condition M (vertical axis) with the sites in condition N (horizontal axis).

(B) Distribution of ratios that reflect changes in read density at all detected binding sites in response to the expression of a LIN28 sponge construct.

(C) Cumulative distribution functions showing the minimum free energy of the ensemble for all binding sites detected before sponge expression (green; n= 992 sites) and after sponge expression (red; n = 70 sites) compared with randomly chosen sequences outside of eCLIP-identified binding sites (dashed gray).

(D) Design of gene expression constructs used to track changes in *let*-7 miRNA activity in response to the depletion of LIN28 by sponge in single cells, as measured by analytical flow cytometry. Differences in miRNA activity are measured via changes in turquoise-to-cherry fluorescence as a function of sponge expression (citrine). All fluorescence data are binned, and trends map the average response to LIN28B expression (N = 3, ± SEM).

(E) Changes in polysome enrichment in V6.5 mESCs and *Dicer* knockout mESCs after the expression of LIN28 sponge.

(F) A rolling mean plot showing the relationship between the change in polysome enrichment (rank ordered from the most repressed to the most activated by sponge construct) and the frequency of *let*-7 target sites per gene.

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Anti-LIN28B (D4H1) | Cell Signaling | 11965; RRID: AB_2750978 |
| Anti-LIN28A (6D1F9) | Cell Signaling | 5930; RRID: AB_1903976 |
| Anti-V5 | MBL | M167-3; RRID: AB_1953024 |
| Chemicals, peptides, and recombinant proteins | | |
| Doxycycline hydrochloride | Sigma | D3447 |
| LIN28-Tat | Peprotech | 110-06 |
| Critical commercial assays | | |
| Illumina Stranded mRNA Prep | Illumina | 20040532 |
| miScript Primer Assays | QIAGEN | 218300 |
| Deposited data | | |
| RNA sequencing data for Polysome Fractionation | Tan et al. (2019) | GEO: GSE109423 |
| RNA sequencing data | This Manuscript | GEO: GSE178259 |
| Uncropped Western Blots | Mendeley | https://data.mendeley.com/datasets/g5sznxfzfw/1 |
| Experimental models: Cell lines | | |
| Human HEK293A | This Manuscript | N/A |
| LIN28B knockout (human HEK293A) | Tan et al. (2019) | N/A |
| Dicer Knockout mouse embryonic stem cells | Dr. Anthony Leung (JHU) | N/A |
| V6.5 mouse embryonic stem cells | This Manuscript | N/A |
| Recombinant DNA | | |
| LIN28B transgene PiggyBAC | This Manuscript | N/A |
| LIN28B sponge transgene PiggyBAC | This Manuscript | N/A |
| Software and algorithms | | |
| Random Forest Classifiers | Breiman (2001) | https://www.stat.berkeley.edu/~breiman/RandomForests/ |