Landes Bioscience www.landesbioscience.com

Supplemental Material to:

Robin Graf, Mathias Munschauer, Guido Mastrobuoni, Florian Mayr, Udo Heinemann, Stefan Kempa, Nikolaus Rajewsky, Markus Landthaler

Identification of LIN28B-bound mRNAs reveals features of target recognition and regulation

2013; 10(6)
http://dx.doi.org/10.4161/rna.25194

www.landesbioscience.com/journals/rnabiology/article/25194/

Supplemental Material

1. 2013RNABIOL0082R-suptbl2
2. 2013RNABIOL0082R-suptbl3

SUPPLEMENTARY DATA

Identification of LIN28-bound mRNAs reveals features of target recognition and regulation

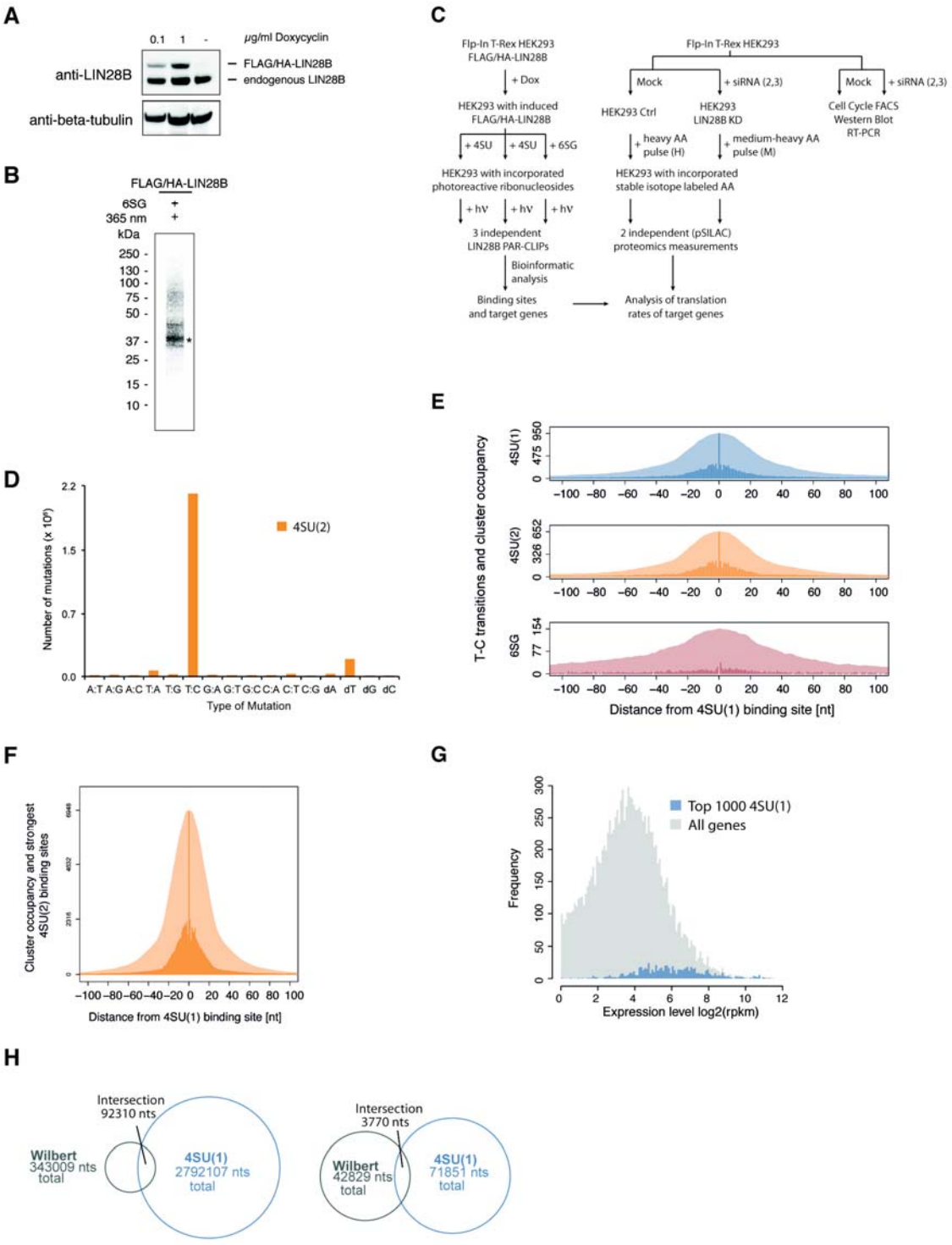
Robin Graf, Mathias Munschauer, Guido Mastrobuoni, Florian Mayr, Udo Heinemann, Stefan Kempa, Nikolaus Rajewsky, and Markus Landthaler

TABLE AND FIGURES LEGENDS

Supplementary Table S1: Overview of sequence read numbers from indicated PAR-CLIP experiments obtained after illumina sequencing and mapping of to the spliced transcriptome.

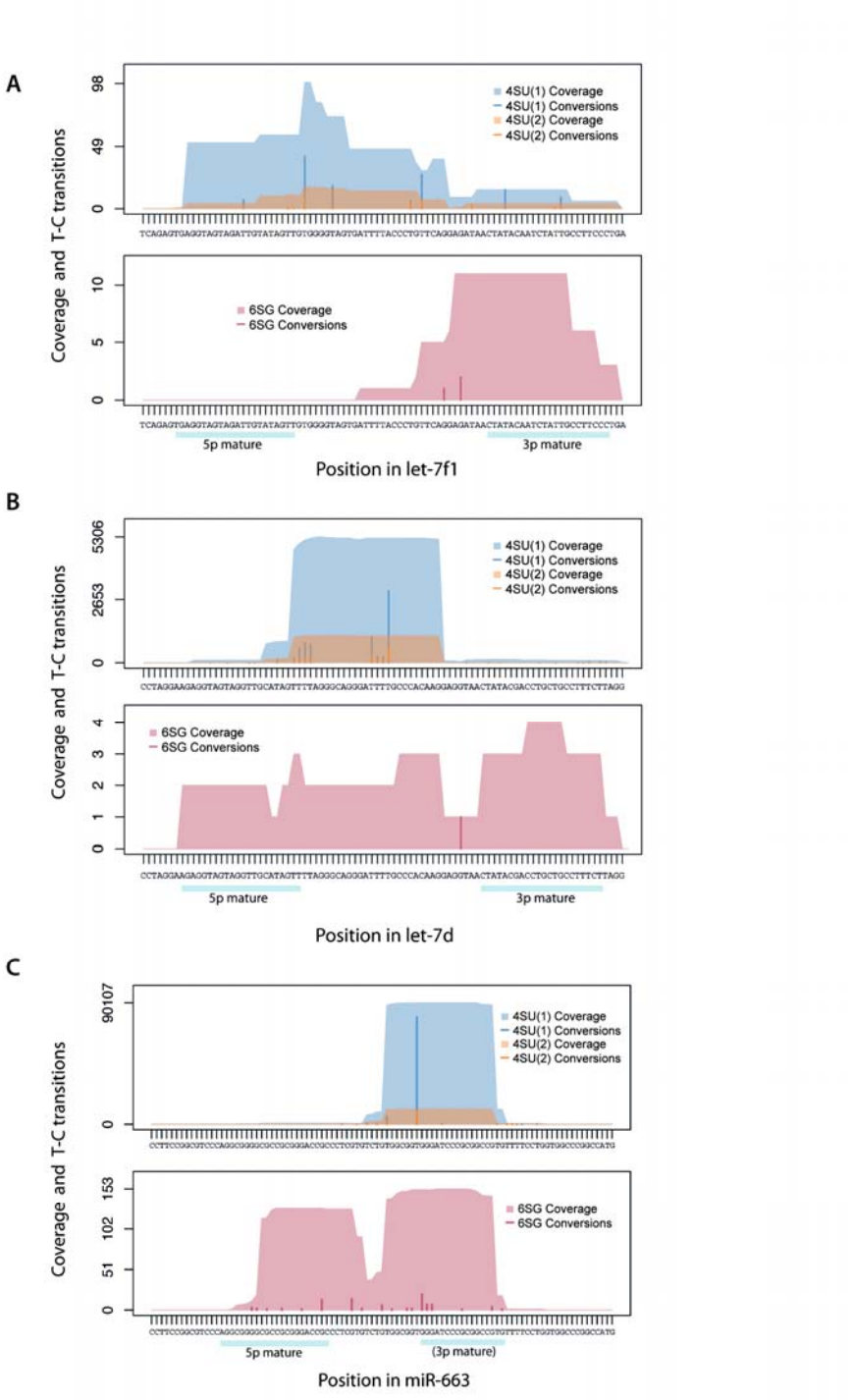
Supplementary Table S2: List of conservative LIN28B PAR-CLIP targets, target clusters and Log2 fold-change values obtained from pSILAC measurements of LIN28B siRNA knockdown experiments using 2 different siRNAs in biological replicate experiments.

Supplementary Table S3: Gene Ontology analysis

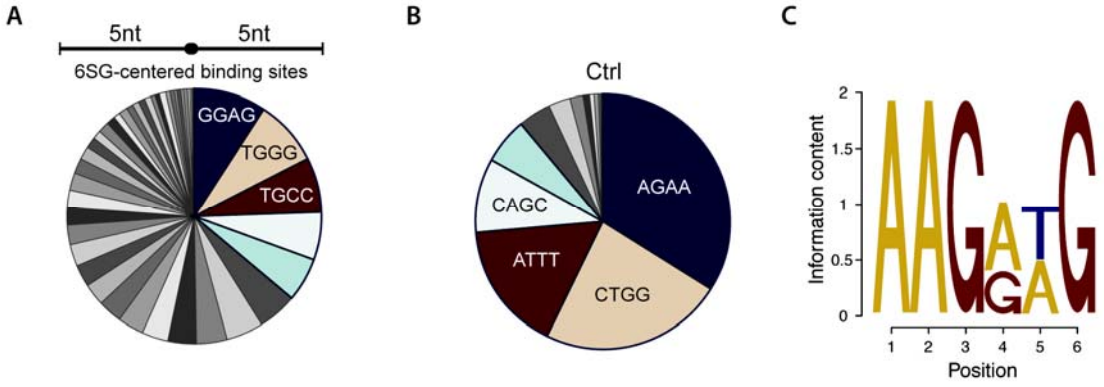


Supplementary Figure 1: (A) Western blot showing induction of HIS/FLAG/HA-LIN28B protein at different doxycyclin concentrations. LIN28B antibody was used for detection of FLAG/HA-tagged and endogenous LIN28B. Beta-tubulin antibody was used to control for equal loading. (B) Autoradiogram of SDS-PAGE, transferred to nitrocellulose membrane. Protein-RNA complexes

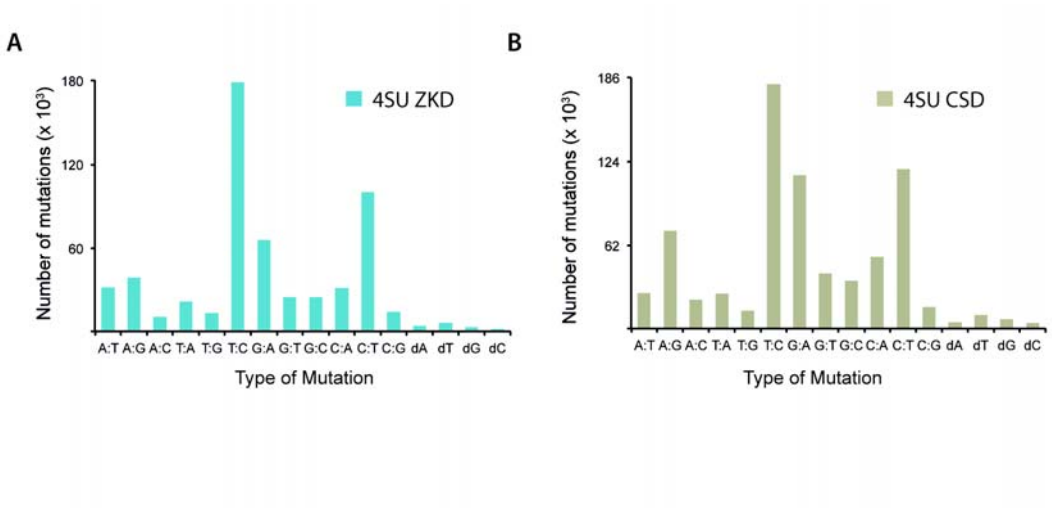
migrating at 39 kDa correspond to epitope-tagged LIN28B crosslinked to RNA using 6SG as photoreactive nucleoside. (C) Flowchart of experimental design. (D) Frequency of nucleotide mutations detected in 4SU(2) PAR-CLIP libraries after alignment to spliced human transcriptome (E) T-C transition frequency and cluster occupancy (defined as number of clusters in respective transcript region) in 4SU(1), 4SU(2) and 6SG PAR-CLIP libraries, centered on the strongest crosslinking site observed in 4SU(1) PAR-CLIP. (F) Distribution of the strongest 4SU(2) binding sites centered on the strongest 4SU(1) binding sites. T-C transition frequency and cluster occupancy is shown. (G) Histogram of RPKM values of genes covered by the top 1000 4SU (1) PAR-CLIP binding sites and of all genes. (H) Intersection of covered nucleotides in all (left) and top 1000 (right) LIN28A clusters identified by CLIP-seq in Wilbert et al. compared to nucleotides covered in 4SU(1) PAR-CLIP library.



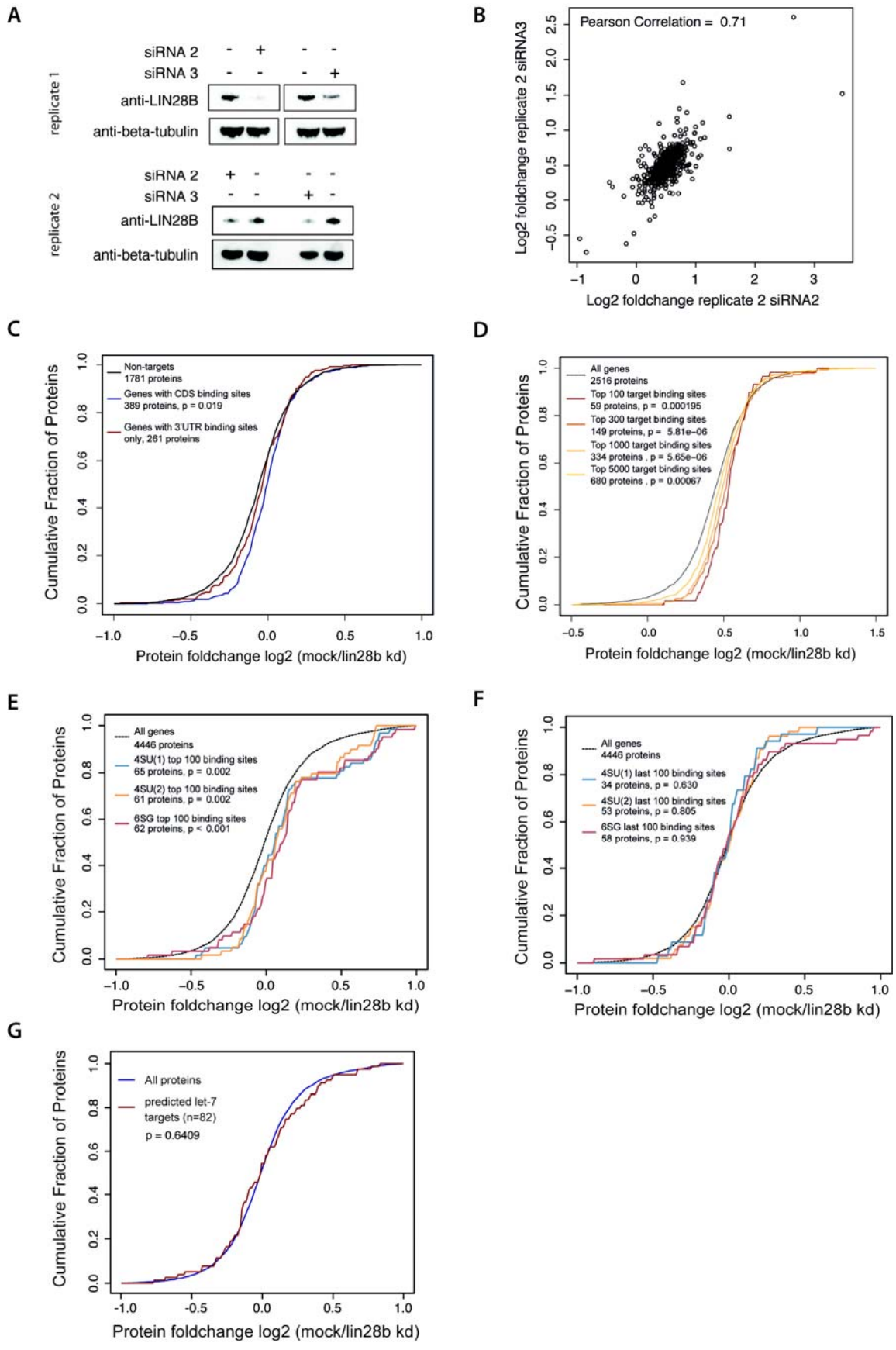
Supplemental Figure 2: (A-C) Alignment of sequence coverage signal and diagnostic nucleotide transitions observed in 4SU (1) (blue), 4SU (2) (orange) and 6SG (dark red) libraries to genomic regions encoding let-7f1, let-7d and miR-663 precursors.



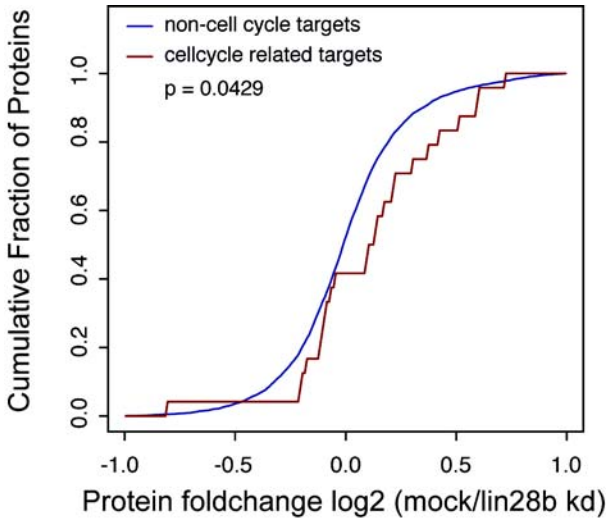
Supplemental Figure 3: (A) Most frequent tetramers in 6SG centered 3'UTR binding sites (extended by 5 nt upstream and downstream) within conservative cluster set. (B) Most frequent tetramer within G centered random control clusters derived from conservative cluster set. (C) Top sequence motif identified by MEME in top 300 6SG centered CDS binding sites.



Supplementary Figure 4: (A and B) Frequency of nucleotide mutations detected in ZKD (A) and CSD (B) iDo-PAR-CLIP libraries after alignment to spliced human transcriptome. dA, dT, dG and dG indicate respective nucleotide deletions.


A
replicate 1
siRNA 2 - + - -
siRNA 3 - - - +
anti-LIN28B
anti-beta-tubulin
replicate 2
siRNA 2 + - - -
siRNA 3 - - + -
anti-LIN28B
anti-beta-tubulin
B
Pearson Correlation = 0.71
Log2 foldchange replicate 2 siRNA3
Log2 foldchange replicate 2 siRNA2
C
Cumulative Fraction of Proteins
Non-targets
1781 proteins
Genes with CDS binding sites
389 proteins, p = 0.019
Genes with 3'UTR binding sites
only, 261 proteins
Protein foldchange log2 (mock/lin28b kd)
D
Cumulative Fraction of Proteins
All genes
2516 proteins
Top 100 target binding sites
59 proteins, p = 0.000195
Top 300 target binding sites
149 proteins, p = 5.81e−06
Top 1000 target binding sites
334 proteins , p = 5.65e−06
Top 5000 target binding sites
680 proteins , p = 0.00067
Protein foldchange log2 (mock/lin28b kd)
E
Cumulative Fraction of Proteins
All genes
4446 proteins
4SU(1) top 100 binding sites
65 proteins, p = 0.002
4SU(2) top 100 binding sites
61 proteins, p = 0.002
6SG top 100 binding sites
62 proteins, p < 0.001
Protein foldchange log2 (mock/lin28b kd)
F
Cumulative Fraction of Proteins
All genes
4446 proteins
4SU(1) last 100 binding sites
34 proteins, p = 0.630
4SU(2) last 100 binding sites
53 proteins, p = 0.805
6SG last 100 binding sites
58 proteins, p = 0.939
Protein foldchange log2 (mock/lin28b kd)
G
Cumulative Fraction of Proteins
All proteins
predicted let-7
targets (n=82)
p = 0.6409
Protein foldchange log2 (mock/lin28b kd)

Supplementary Figure 5: (A) Western blot of LIN28B siRNA knock down for samples used in pulsed SILAC proteomics experiments (replicate 1 and 2, siRNA 1 and 2, respectively). Anti-beta-tubulin was used to control for equal gel loading. (B) Correlation of log2 foldchanges observed in replicate pSILAC experiments (C) Cumulative density of log2 transformed changes in newly synthesized protein levels, measured by pulsed SILAC upon LIN28B knock down (pSILAC data from siRNA2, replicate 2 is shown). (D) Cumulative density of log2 transformed changes in newly synthesized protein levels upon LIN28B knock down. Genes covered by the top 5000, 1000, 300 and 100 conservative PAR-CLIP binding sites are shown (pSILAC data from siRNA3, replicate 2 is shown). (E) Cumulative density of log2 transformed changes in newly synthesized protein levels. The group of genes mapping to the top 100 binding sites in 4SU (1), 4SU (2) and 6SG PAR-CLIP are compared to all target genes. (F) As in (E), but comparing the group of genes corresponding to the 100 lowest ranked binding sites in 4SU (1), 4SU (2) and 6SG PAR-CLIPs. (G) Cumulative density of log2 transformed changes in newly synthesized protein levels. Genes harboring PicTar predicted let-7 binding sites are compared to all genes. All shown p-values are based on Wilcoxon test.



Supplementary Figure 6: Cumulative density of log2 transformed changes in newly synthesized protein levels, measured by pulsed SILAC upon LIN28B knock down. Cell cycle related targets are compared to targets unrelated to cell cycle. P-values are based on Wilcoxon test.

Supplementarty Table S1: Overview of sequence read numbers from indicated PAR-CLIP experiments obtained after illumina sequencing and mapping of to the spliced transcriptome.

PAR-CLIP	Clipped reads	Mapped reads*	Perfect mapping	Modified mapping	Modifications	T:C	G:A	delT	delG	No. of clusters
LIN28B_4SU_1	25718193	6997431	541009	6456422	8820758	6854674	83694	842730	66997	65592
LIN28B_4SU_2	7531061	1917636	118741	1798895	2605525	2113088	19611	203383	13049	72672
LIN28B_6SG	2760258	324675	162737	161938	187958	19365	48493	6156	14510	3563
LIN28B_4SU_ZKD	8133712	658559	235965	422594	598890	178837	66012	6185	3146	
LIN28B_4SU_CSD	6086297	751654	241318	510336	753825	180968	113981	9963	6740	

*to spliced RNA