Molecular Cell, Volume *69*

Supplemental Information

Systematic Discovery of RNA Binding Proteins that Regulate MicroRNA Levels

Julia K. Nussbacher and Gene W. Yeo

Figure S1. Increasing the stringency of the fold enrichment and padj cutoffs allows filtering of most robust candidates and reveals cell-type specificity of miR-RBP interactions. Related to Figure 1A-E and Figure 3.

(A) eCLIP IP/SMInput $\log_2$(fold enrichment) does not correlate with mature guide miRNA expression (RPM) for DROSHA or DGCR8.

(B) Heatmap of the number of RBPs with at least 1 (n=1) miRNA loci eCLIP clusters exceeding a fold enrichment (x-axis) and p-adjusted (y-axis) cutoff. Dark blue = 100% (126) RBPs passing (at least 1 pre-miR eCLIP cluster exceeding the fold enrichment and p-adjusted cutoffs), white = 0% passing.

(C) Heatmaps as in (B) of the number of RBPs with at least 1 (n=1), 5 (n=5) or 10 (n=10) miRNA loci eCLIP clusters passing the fold enrichment and p-adjusted cutoffs. Selected final cutoff of $\log_2$(fold enrichment) = 2 and $-\log_{10}$(padj) = 3 denoted by yellow text. The percent of RBPs passing the cutoffs: n=1 92% pass, n=5 51% pass, and n=10 25% pass.

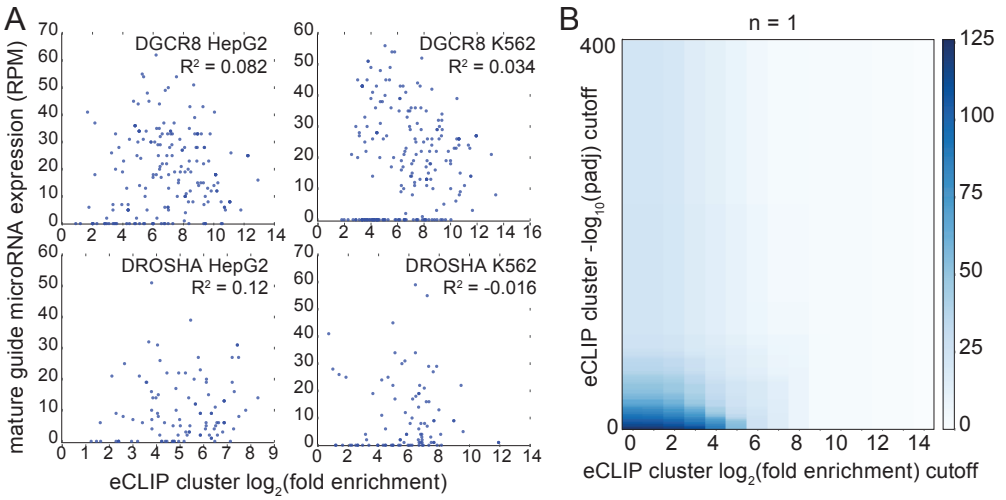
(D) Comparison of DGCR8 eCLIP cluster distribution across genic (excluding intergenic) and non-coding (excluding scRNA) regions with and without log2(fold enrichment) and $-\log_{10}$(padj) cutoffs. Distributions are compared to previously published HITS-CLIP experiments (Macias *et al*., 2012). Values are percentages, except where the true count is denoted for introns and miRNAs in parenthesis. Below the HITS-CLIP and more stringent eCLIP non-coding RNA distributions are the number of unique miRNA loci represented by the total clusters in parenthesis.

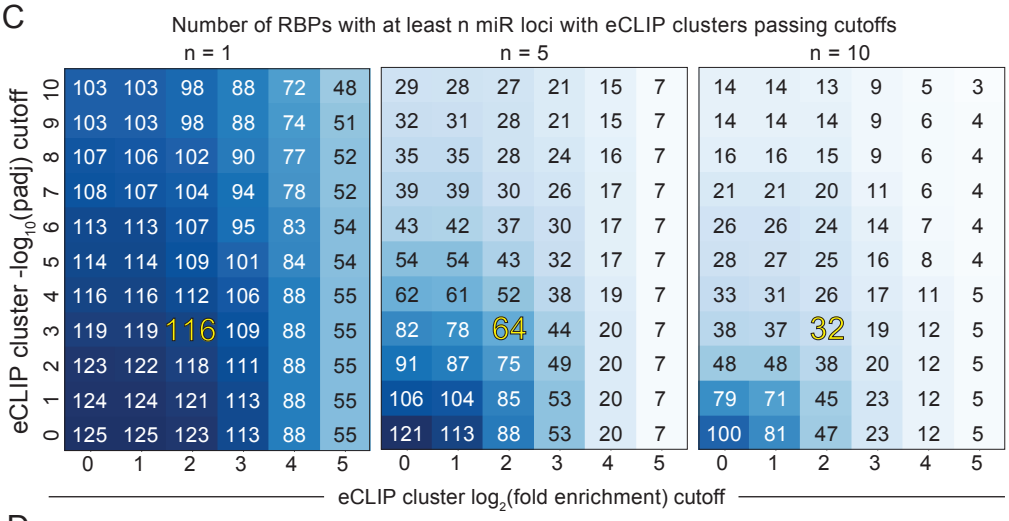
(E) Comparison of unique DGCR8 eCLIP and HITS-CLIP miR targets showing fewer targets identified by eCLIP but better overlap between cell types.

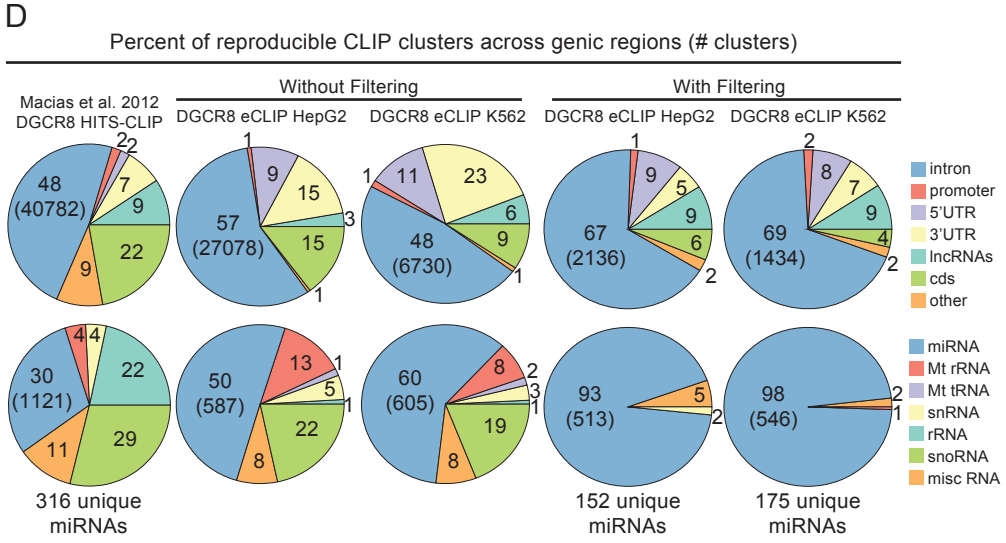
(F) Kim *et al*. further filtered their DROSHA fCLIP using a novel cut-site algorithm; we compare these more stringently filtered miRNA targets with our eCLIP clusters passing cutoffs.

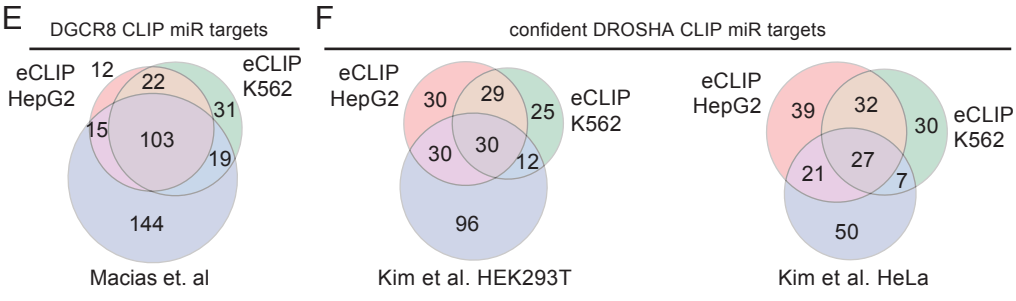
(G) Some RBP-miRNA loci interactions are cell-type specific, even when target is expressed in both contexts. Table quantifies the intersection of miRNA loci with at least 1 eCLIP cluster passing the fold enrichment and p-adj cutoffs in all encode RBPs screened. Clusters in HepG2 and K562 are considered shared if they intersect the same miR locus even if the clusters don't intersect. RBPs not displayed either had no eCLIP clusters passing cutoff in one or both of the cell lines, or only one cell line was subjected to eCLIP. The last two columns show the number of loci that are expressed in the other cell line but weren't bound.

Figure S1









G

RBP	HepG2 Only	Both	K562 Only	K562 Expressed HepG2 Bound	HepG2 Expressed K562 Bound
AGGF1	2	1	1	0/2	1/1
BUD13	22	14	18	7/22	6/18
CSTF2T	8	3	1	5/8	1/1
DDX3X	12	8	6	3/12	3/6
DDX55	2	4	2	0/2	1/2
DDX6	1	1	3	0/1	0/3
DGCR8	27	125	50	14/27	25/50
DHX30	2	1	1	1/2	1/1
DROSHA	37	43	31	27/37	19/31
EFTUD2	24	21	14	4/24	5/14
FASTKD2	3	1	13	1/3	11/13
GTF2F1	3	3	0	3/3	0/0
HNRNPM	4	0	3	0/4	0/3
IGF2BP1	2	0	2	1/2	0/2
ILF3	5	2	2	3/5	1/2
LARP4	5	2	6	2/5	5/6
LARP7	5	1	4	4/5	3/4
LIN28B	2	8	11	2/2	6/11
LSM11	1	1	3	0/1	3/3
NCBP2	17	3	2	9/17	1/2
NOLC1	9	1	3	9/9	1/3
PRPF8	39	12	11	10/39	6/11
PTBP1	2	4	1	1/2	1/1
RBFOX2	8	2	9	1/8	6/9
RBM15	1	3	14	1/1	7/14
RBM22	2	1	7	0/2	5/7
RPS3	11	2	4	4/11	2/4
RPS5	4	2	7	2/4	6/7
SF3B4	22	24	34	3/22	16/34
SLTM	2	1	2	0/2	2/2
SMNDC1	0	5	14	0/0	9/14
SND1	3	0	4	3/3	2/4
SRSF1	5	2	5	3/5	3/5
TIA1	1	2	1	0/1	1/1
TNRC6A	4	1	9	4/4	8/9
TRA2A	2	0	3	0/2	1/3
TROVE2	2	0	6	1/2	5/6
U2AF1	7	0	4	6/7	1/4
U2AF2	12	6	1	7/12	0/1
UCHL5	1	1	15	0/1	4/15
XRN2	4	2	10	1/4	7/10
ZC3H11A	2	2	1	1/2	0/1

Figure S2. Volcano plots of pre-miR eCLIP clusters in candidate RBPs. Related to Figure 3B.
(A-B) Volcano plots of eCLIP clusters at miRNA loci in HepG2 and K562 cell lines for RBP candidates. Cutoffs of eCLIP IP/SMInput $\log_2$(fold enrichment) = 2 and $-\log_{10}$(padj) = 3 are denoted by a dashed red line. Cluster are colored as in Figure 1. RBP and cell line indicated above each panel.

Figure S2

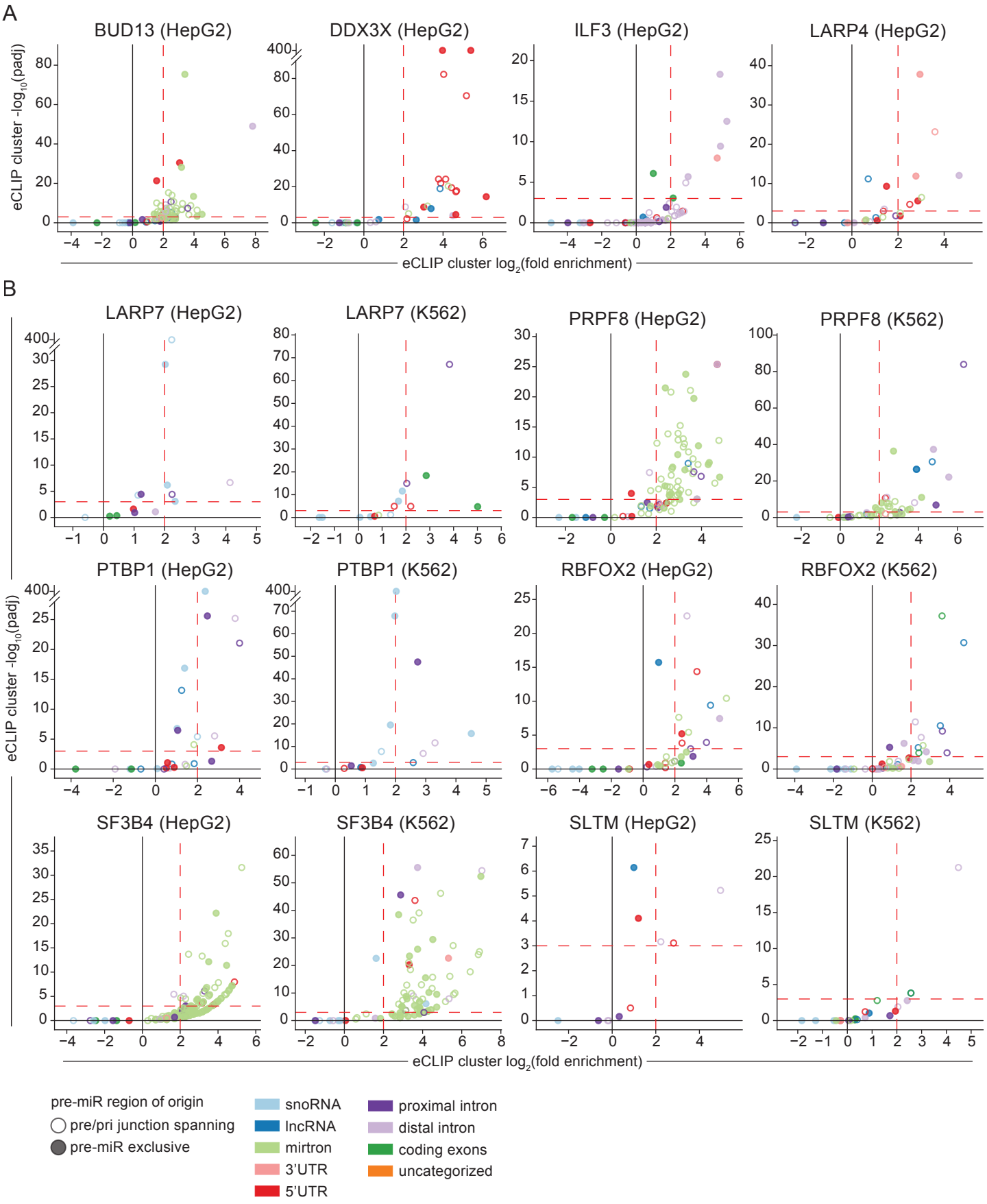

A
BUD13 (HepG2)
DDX3X (HepG2)
ILF3 (HepG2)
LARP4 (HepG2)
eCLIP cluster -log10(padj)
eCLIP cluster log2(fold enrichment)
B
LARP7 (HepG2)
LARP7 (K562)
PRPF8 (HepG2)
PRPF8 (K562)
PTBP1 (HepG2)
PTBP1 (K562)
RBFOX2 (HepG2)
RBFOX2 (K562)
SF3B4 (HepG2)
SF3B4 (K562)
SLTM (HepG2)
SLTM (K562)
eCLIP cluster -log10(padj)
eCLIP cluster log2(fold enrichment)
pre-miR region of origin
pre/pri junction spanning
pre-miR exclusive
snoRNA
lncRNA
mirtron
3'UTR
5'UTR
proximal intron
distal intron
coding exons
uncategorized

Figure S3. Western Blots of RBP knockdown and small RNA-Seq replicate correlation. Related to Figure 2 and 3.

(A-C) Western blot of RBP knockdowns in HepG2 and K562 cells by lentiviral transduction of shRNA. Controls were performed in quadruplicate and shRNA knockdowns in duplicate, controls are representative; C=control, sh=RBP knockdown. GAPDH and α-tubulin were used as loading controls where indicated.

(D) Distribution of the R^2 Pearson correlation values for shControl replicates of small RNA-Seq samples.

(E) Representative correlation plot for (D).

(F) Distribution of the R^2 Pearson correlation values for shRBP knockdown replicates of small RNA-Seq samples.

(G) Representative correlation plot for (F).

Figure S3

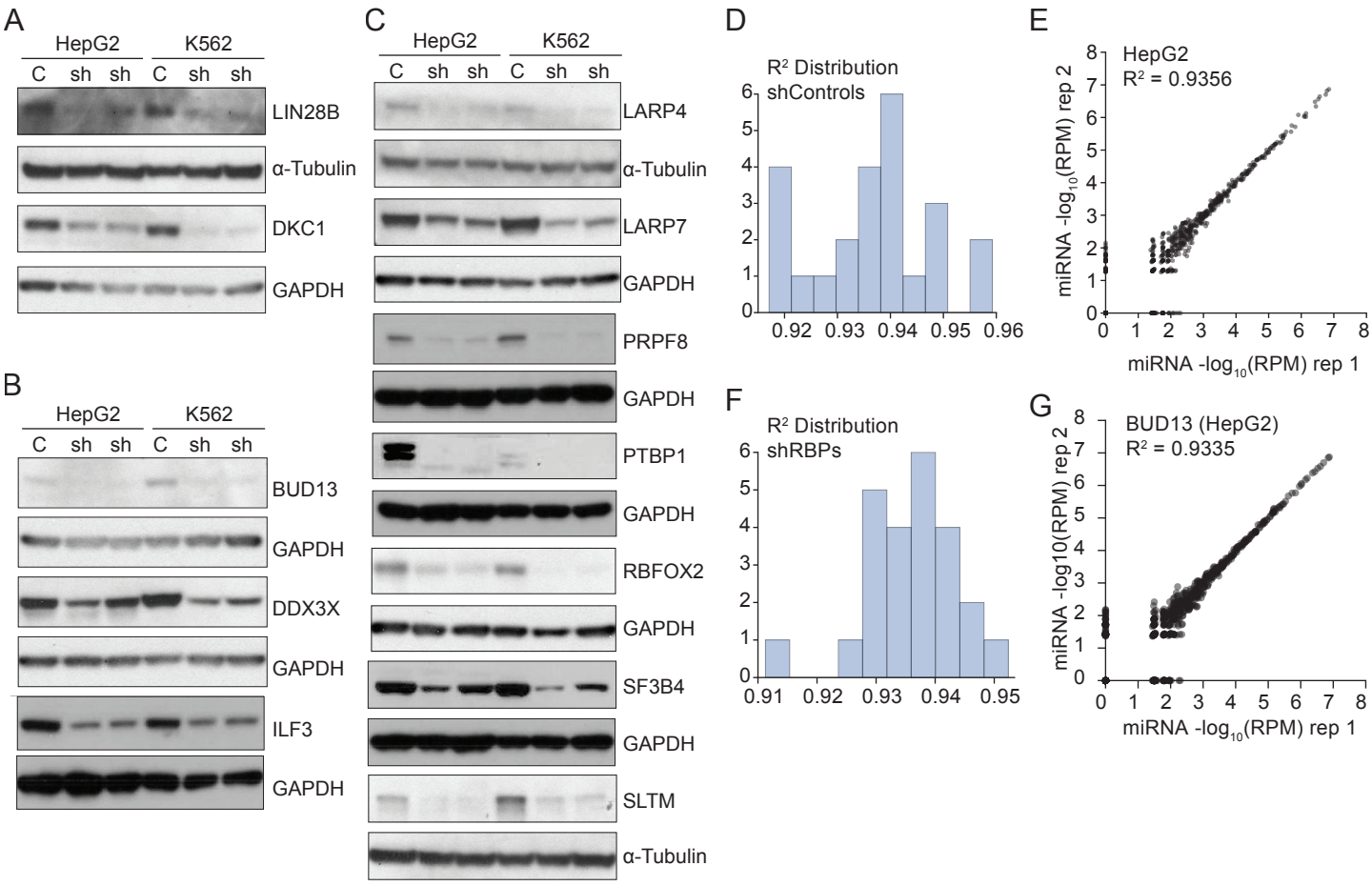

A
HepG2
K562
C sh sh C sh sh
LIN28B
α-Tubulin
DKC1
GAPDH
B
HepG2
K562
C sh sh C sh sh
BUD13
GAPDH
DDX3X
GAPDH
ILF3
GAPDH
C
HepG2
K562
C sh sh C sh sh
LARP4
α-Tubulin
LARP7
GAPDH
PRPF8
GAPDH
PTBP1
GAPDH
RBFOX2
GAPDH
SF3B4
GAPDH
SLTM
α-Tubulin
D
R² Distribution
shControls
0.92 0.93 0.94 0.95 0.96
F
R² Distribution
shRBPs
0.91 0.92 0.93 0.94 0.95
E
HepG2
R² = 0.9356
miRNA -log10(RPM) rep 2
miRNA -log10(RPM) rep 1
G
BUD13 (HepG2)
R² = 0.9335
miRNA -log10(RPM) rep 2
miRNA -log10(RPM) rep 1

Figure S4. eCLIP and small RNA-Seq integration. Related to Figure 4.
(A-B) Integration of eCLIP and small RNA-Seq data. eCLIP clusters at miRNA loci were further filtered for those intersecting miRNAs detected by small RNA-Seq. For these clusters, eCLIP cluster IP/smInput $\log_2$(fold enrichment) was plotted vs. small RNA-Seq miRNA $\log_2$(fold change). miRNAs with significant eCLIP clusters (padj < 0.05) that also changed significantly upon RBP knockdown (padj < 0.05; Benjamini-Hochberg correction) are highlighted and colored as in Figure 1.

Figure S4

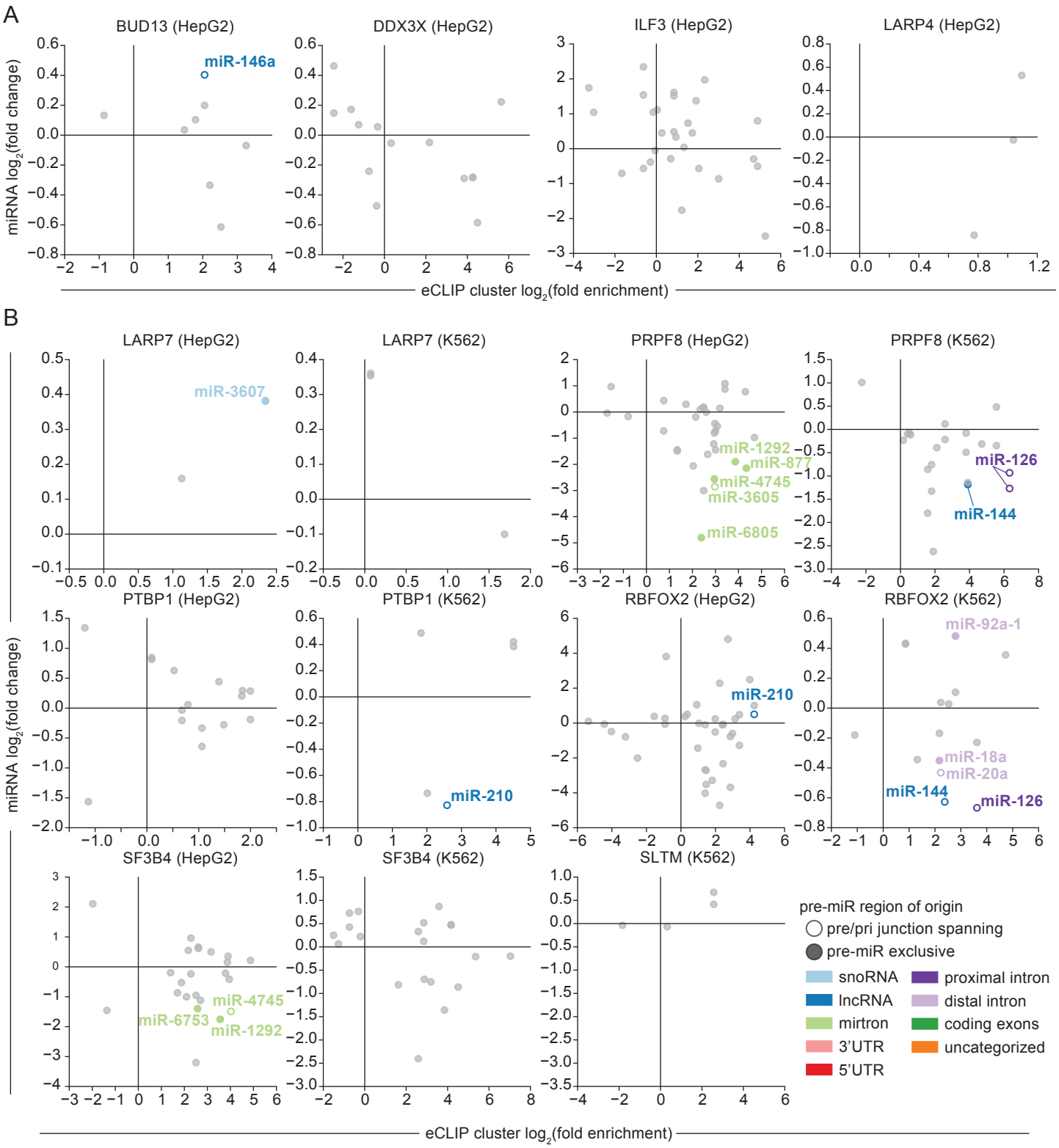
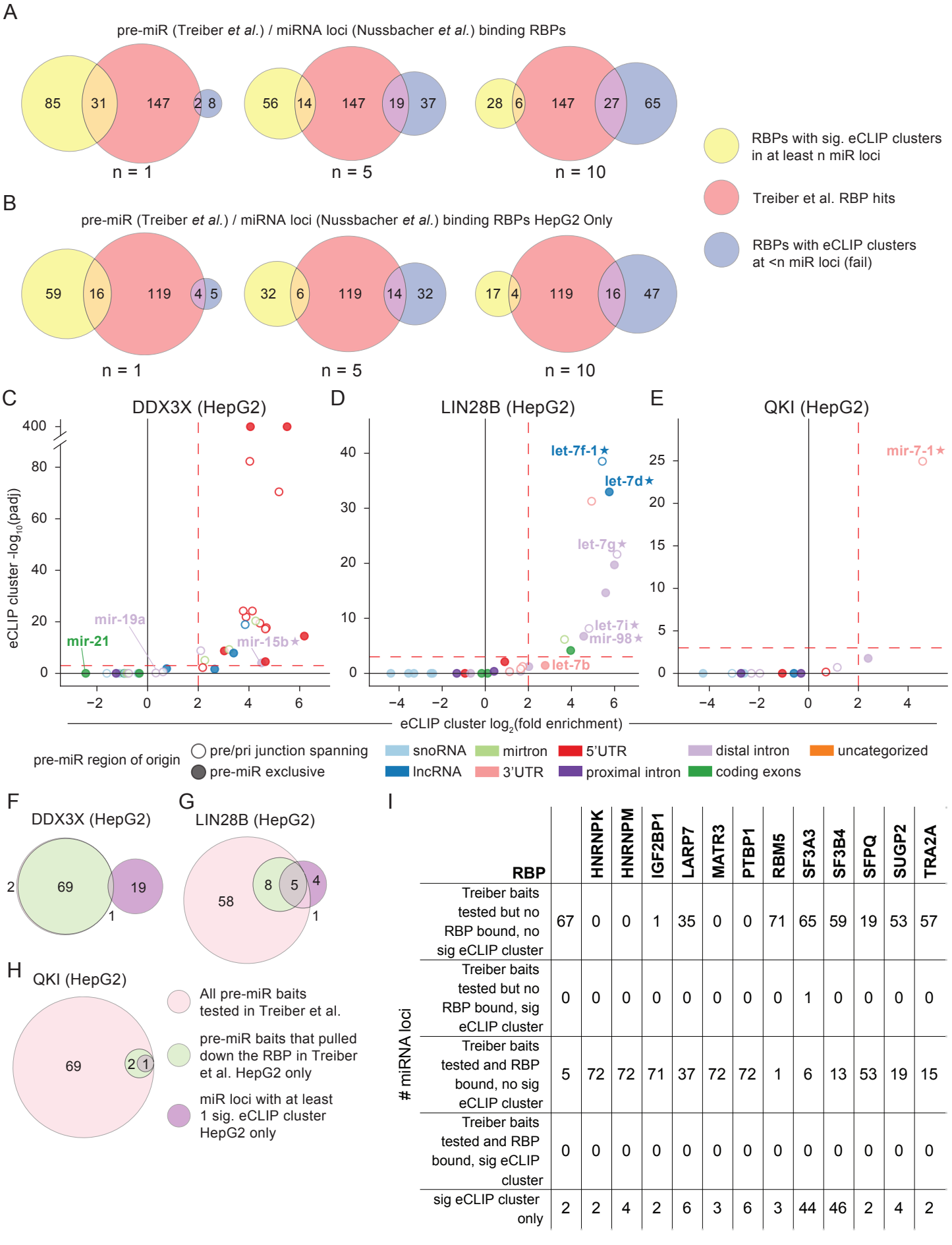

A
BUD13 (HepG2)
DDX3X (HepG2)
ILF3 (HepG2)
LARP4 (HepG2)
miR-146a
miRNA log2(fold change)
eCLIP cluster log2(fold enrichment)
B
LARP7 (HepG2)
LARP7 (K562)
PRPF8 (HepG2)
PRPF8 (K562)
miR-3607
miR-1292
miR-877
miR-4745
miR-3605
miR-6805
miR-126
miR-144
PTBP1 (HepG2)
PTBP1 (K562)
RBFOX2 (HepG2)
RBFOX2 (K562)
miR-210
miR-92a-1
miR-18a
miR-20a
SF3B4 (HepG2)
SF3B4 (K562)
SLTM (K562)
miR-6753
pre-miR region of origin
pre/pri junction spanning
pre-miR exclusive
snoRNA
lncRNA
mirtron
3'UTR
5'UTR
proximal intron
distal intron
coding exons
uncategorized

Figure S5. Comparison of findings with Treiber et al 2017. Related to Figure 3B.
(A) Comparison of RBPs identified as putative binders of pre-miRs/miRNA loci. The 180 RBPs identified by Trieber *et al.* (red) were intersected with the RBPs we identified as having at least n eCLIP cluster at a miRNA locus exceeding the fold enrichment and p-adjusted cutoffs in at least 1 cell line tested (yellow) as well as the RBPs for which <n clusters exceeded the fold enrichment and p-adjusted cutoffs in the cell line(s) tested (blue).

Figure S5



I

# miRNA loci	RBP	HNRNPK	HNRNPM	IGF2BP1	LARP7	MATR3	PTBP1	RBM5	SF3A3	SF3B4	SFPQ	SUGP2	TRA2A
Treiber baits tested but no RBP bound, no sig eCLIP cluster	67	0	0	1	35	0	0	71	65	59	19	53	57
Treiber baits tested but no RBP bound, sig eCLIP cluster	0	0	0	0	0	0	0	0	1	0	0	0	0
Treiber baits tested and RBP bound, no sig eCLIP cluster	5	72	72	71	37	72	72	1	6	13	53	19	15
Treiber baits tested and RBP bound, sig eCLIP cluster	0	0	0	0	0	0	0	0	0	0	0	0	0
sig eCLIP cluster only	2	2	4	2	6	3	6	3	44	46	2	4	2

Table S1: shRNAs for RBPs. Related to STAR Methods

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
shRNA targeting BUD13	Sigma-Aldrich	TRCN0000074896
shRNA targeting DDX3X	Sigma-Aldrich	TRCN0000000003
shRNA targeting DKC1	Sigma-Aldrich	TRCN0000039738
shRNA targeting ILF3	Sigma-Aldrich	TRCN0000329787
shRNA targeting LARP4	Sigma-Aldrich	TRCN0000161048
shRNA targeting LARP7	Sigma-Aldrich	TRCN0000122544
shRNA targeting LIN28B	Sigma-Aldrich	TRCN0000144508
shRNA targeting PRPF8	Sigma-Aldrich	TRCN0000075112
shRNA targeting PTBP1	Sigma-Aldrich	TRCN0000001062
shRNA targeting RBFOX2	Sigma-Aldrich	TRCN0000074544
shRNA targeting SF3B4	Sigma-Aldrich	TRCN0000000039
shRNA targeting SLTM	Sigma-Aldrich	TRCN0000135106