RBM17 Mediates Evasion of Pro-Leukemic Factors from Splicing-coupled NMD to Enforce Leukemic Stem Cell Maintenance

Lina Liu[1], Ana Vujovic[2,3], Nandan P. Deshpande[5,6], Shashank Sathe[4], Govardhan Anande[5,6], Joshua Xu[1,3], Gene W, Yeo[4], Ashwin Unnikrishnan[5,6], Kristin Hope[2,3,*], and Yu Lu[1,*]

[1] Department of Biochemistry and Biomedical Sciences, Faculty of Health Sciences, McMaster University, Ontario, Canada
[2] Department of Medical Biophysics, University of Toronto, Ontario, Canada
[3] Princess Margaret Cancer Centre, University Health Network, Ontario, Canada
[4] Department of Cellular and Molecular Medicine, Stem Cell Program and Institute for Genomic Medicine, University of California at San Diego, California, United States
[5] Adult Cancer Program, Lowy Cancer Research Centre, University of New South Wales, Sydney, Australia
[6] Price of Wales Clinical School, University of New South Wales, Sydney, Australia

* Corresponding authors:
Kristin Hope (kristin.hope@uhnresearch.ca)
Yu Lu (yu.lu@mcmaster.ca)

Supplementary Materials and Methods

RNA extraction, qRT-PCR

Total cellular RNA was isolated with Trizol LS reagent (Invitrogen) according to the manufacturer's instructions and cDNA was synthesized using qScript cDNA Synthesis Kit (Quanta Biosciences). For qRT-PCR, samples were prepared with iTaq Universal SYBR Green Supermix (Bio-Rad) and ran as described by the manufacturer, GAPDH or 18S rRNA was used as internal control.

Protein lysates preparation and WB

Cells were collected and then spalled by lysis buffer. Protein concentration were measured by micro BCA™ protein assay kit. After normalization, equal amount of proteins within 1x LDS loading buffer containing 10% DTT were boiled for 7 minutes at 95 °C prior to electrophoresis. Proteins were transferred onto NC membrane, blocked with 5% skim milk in 1x TBST for 1 hour at room temperature and then incubate overnight at 4°C with primary antibodies rabbit-anti-RBM17 (ab204333, Abcam), rabbit-anti-GAPDH (2118S, NEB). Following membrane washing,

secondary antibodies IRDye 680 goat-anti-rabbit (Li- COR Biosciences) was added for 1 hour at room temperature and washed 3 times with 1x TBST, then were imaged with the Odyssey Classic Imager (Li-COR Biosciences).

Generation of recombinant lentivirus

PLKO.1 lentiviral vectors that expressing DNA fragments encoding shRNAs against RBM17 or scramble (as the negative control) were purchased form Sigma. For shRNAs containing a GFP selection marker, puro was replaced by EGFP cloned from MA-1 plasmid. shRNAs oligonucleotides against EIF4A2, were ligated into PLKO.1 lentiviral vector with a GFP marker by using AgEI and ECORI sites. Recombinant vectors and packaging plasmids PSPAX2 and PMD2.G were transfected into HEK 293FT cells to produce recombinant lentivirus.

Calculation of relative engraftment potential

For AML sample#001: using flow cytometry, we first measured % of GFP+ cells within the 7-AAD- gate to determine % of shRNA-expressing cells injected (Input, %GFP-injected). After engraftment (9 weeks), we also measured % of GFP+ cells within the 7-AAD$^-$ gate and human CD45$^+$ gate to determine the % of shRNA-expressing cells engrafted (Output, %GFP-engrafted). For each experimental group (e.g. control as x and shRBM17 as group y), we first calculated engraftment value x or y using the following formula: x or y=%GFP- engrafted ÷ %GFP-injected for each mouse. This calculation yielded an array of engraftment values for both control (x1, x2, x3, x4, x5) and shRBM17 (y1, y2, y3, y4, y5) groups. To calculate the final engraftment potential score, each engraftment value was normalized by the mean of (x1, x2, x3, x4, x5) from the control group. These scores were plotted to compare the relative engraftment potential between control and shRBM17.

For AML sampel#005: Input (% GFP-injected) was over 85%, after engraftment (12 weeks), GFP was silenced, we measured % of human myeloid lineage cells (CD45$^+$CD33$^+$) within 7-AAD$^-$ gate, and the value was marked as output (% engraftment).

Cell proliferation assay

Cell proliferation were measured by counting the number of cells. In brief, GFP+ cells were sorted and seeded ($2x10^4$/ml) into 12-well plates after infection with shRNAs. Cells were stained by trypan blue and then counted by Cell Counter according to the manufacturer's instructions every two days.

RNA-seq splicing analysis and protein effects prediction

Quality control checks were performed on raw RNA-seq data using FastQC (v0.11.5). Adapter contamination and low-quality sequences in the ShRBM17 and Control (duplicate) samples were removed using the tool FastxToolkit. The quality filtered data was processed for uniform read length (73 bp) using the tool trimmomatic v0.38[1]. The pre-processed samples were individually aligned to Human genome (UCSC HG38) with default settings using the transcriptome aligner STAR v 2.7.2b, --outSAMtype BAM SortedByCoordinate). Differentially spliced events were identified in ShRBM17 samples compared to Control using the rMATS tool (v 4.0.1)[2]. rMATS was run in paired-end mode for the 73 bp uniform-length reads. Splice junction annotations for splicing events were used from ensemble GTF file (GRCh38.96). Five types of splicing events i.e., Exon skipping (CE), Intron retention (RI), Mutually exclusive exons (MXE), Alternative 3' splice site (A3'SS) and Alternative 5' splice site (A5'SS) were identified by rMATS. The differentially spliced events identified using rMATS were filtered with FDR < 0.1 in each cohort. In-house scripts were developed in python to identify splicing events which led to possible functional switches. These included biotype changes (e.g., from transcripts coding for functional proteins in one condition to the transcripts leading to proteins marked for nonsense-mediated decay or processed transcripts without a protein product in the other condition) or the events which lead to changes in protein product due to frameshift, thus, resulting in complete or partial loss of functional domains. The Bioconductor/R packages maser[3] and drawProteins[4] were employed for visualizations of the alterative splicing events in transcripts in context of their protein products.

Proteomic Sample Preparation

One million K562 cells transduced with shscramble or shRNA (#1, #2) targeting RBM17 were harvested (day 5 after transduction, replicates for each condition), washed three time with ice cold 1xDPBS. Cell pellets we lysed in 200µl of lysis buffer composing of 8M urea (Sigma-Aldrich) and 100 mM ammonium bicarbonate (Sigma-Aldrich). Cells were then vortexed at 2,800rpm using Mini S-2 Vortex Mixer (Fisher Scientific) for ten seconds, followed by ten seconds of incubation on ice. This procedure was repeated six times. The lysate was then centrifuged at 21,000xg for five minutes at 4°C. Protein reduction was conducted using 5mM of tris (2-carboxyethyl) phosphine (Sigma-Aldrich) for 45 minutes at 37°C. Subsequently, 10mM of iodoacetamide (Sigma-Aldrich) was added for protein alkylation for 45 minutes at room temperature (dark). Following alkylation, cell lysate was diluted five-fold with 100mM of ammonium bicarbonate to lower urea

concentration. Based on protein amount, Sequencing Grade Modified Trypsin (Promega) was then added in (trypsin: protein(w:w) at 1:50) for overnight digestion at 37°C. Trifluoroacetic acid (Thermo Scientific) was added to reduce pH, and desalting was conducted with SOLA Solid Phase Extraction 2mg 96-well plates (Thermo Scientific). Peptides were eluted twice using 200µL 80% Acetonitrile - 0.1% trifluoroacetic acid. Eluted peptides were speed-vacuum dried using Labconco CentriVap Benchtop Vacuum Concentrator (Kansas City, MO).

Tandem Mass Tag Six-plex (TMT 6-plex) Labeling, Liquid Chromatography and Tandem Mass Spectrometry (LC/MS/MS)

TMTsixplex Isobaric Label Reagent Set (Thermo Fisher) was resuspended in LC-MS grade anhydrous acetonitrile (Sigma-Aldrich) following manufacturer's protocol. Briefly, 0.8mg of TMT reagent was resuspended in 41µL of acetonitrile and incubated at room temperature for ten minutes. During the incubation, dried peptide samples were resuspended in 100mM of triethylammonium bicarbonate (TEAB) (Sigma-Aldrich) to 1µg/µL. Then, TMT reagents were mixed with peptide samples at 4:1 (wt/wt) ratio and incubated at room temperature for one hour. Following incubation, each TMT reaction was quenched with 8µL of 5% hydroxylamine (Sigma-Aldrich) for 15 minutes at room temperature. Labeled samples were pooled together at equal ratio and then fractionated on a home-made high-pH C18 column (200µm x 30cm, packed with Waters BEH130 C18 5µm resin) into 36 cuts. All cuts were then injected and separated on homemade trap (200µm x 5cm, packed with POROS 10R2 C18 10µm resin) and analytical column (50µm x 50cm, packed with Reprosil-Pur 120 C18-AQ 5µm resin), with 3hr reverse-phase gradient delivered by a Thermo Fisher Ultimate 3000 RSLCNano UPLC system coupled to a Thermo QExactive HF quadrupole−Orbitrap mass spectrometer. A parent ion scan was performed using a resolving power of 120,000 and then up to the 20 most intense peaks were selected for MS/MS (minimum ion count of 1000 for activation), using higher energy collision induced dissociation (HCD) fragmentation. Dynamic exclusion was activated such that MS/MS of the same m/z (within a range of 10ppm; exclusion list size=500) detected twice within 5s were excluded from analysis for 40s.

Proteomic data processing and analysis

LC-MS data generated was analyzed against a UniProt human protein database (42,173 entries) for protein identification and quantification by Thermo Proteome Discoverer (v 2.2.0). Identified proteins have at least one unique peptide, while peptide FDR (less than 0.01) and protein FDR (less than 0.01) cutoffs were used. The Significant B values were calculated using the PERSEUS (v.1.6.5) software[5]. Significance B value preset with an FDR<0.01 was used to identify proteins that are significantly differentially abundant and used for downstream integrative analysis.

Figure legends

Figure S1. RBM17 is preferentially expressed in primitive cell fraction of AML. (A) RBM17 transcript level in the primitive $CD34^+$ vs the committed $CD34^-$ subsets in OCI-AML-8227 cells measured by qRT-PCR. (B) One representative primary AML sample (#007) with intracellular flow cytometry plots showing expression profile of CD34 (left) and percentage of $RBM17^+$ cells in $CD34^+$ and $CD34^-$ fractions of this AML sample (right). (C) Intracellular flow cytometry plots showing expression profile of $RBM17^+$ cells in $CD34^+$ and $CD34^-$ fractions of OCI-AML-8227 cells.

Figure S2. RBM17 supports AML cell proliferation and inhibits AML cell differentiation. (A) Flow cytometry validation of RBM17 knockdown in $CD34^+$ OCI-AML-8227 cells using shRBM17#1 and shRBM17#2. (B-D) Growth curves of transduced K562, HL60 and NB4 cells with RBM17 knockdown and control lentivirus. (E-G) Assessment of the effects of RBM17 knockdown on myeloid differentiation of HL60 cells (E), OCI-AML-8227 cells (F) and a primary AML sample (G) as compared to control shRNA. (H) Example of colonies derived from primary AML cells transduced with a scramble (control) or RBM17-targeting shRNA (#1,2). Cells were sorted on the basis of GFP positivity 2 days after viral transduction and colonies were scored 14 d after plating in methylcellulose. (I) Flow cytometry analysis of AnnexinV signal in primary AML cells on day 7 following knockdown of RBM17. Mean ± SD, n = 3. (J-L) Raw data of engraftment experiment performed in AML sample #001 and sample #006, assessing the impact of RBM17 knockdown on their engraftment potential. For sample #001, the y-axis (%GFP^+ cells of $CD45^+$) indicates % of shRNA-positive cells in injected and engraftment human leukemia cells (J); For sample #006, shRNA infection rates are over 80% when injected (K) and the y-axis (%$CD45^+CD33^+$) indicates engraftment human leukemia cells in bone marrow (BM), right femur (RF) and spleen (L). *$p<0.05$, **$p<0.01$.

Figure S3. Identification of RBM17 RNA targets and downstream AS in AML. (A) Representative western blot showing results for the immunoblot analysis of the input and RBM17 samples, showed RBM17 is selectively immunoprecipitated. (B) Analysis of RBM17 binding gene type. (C) GO enrichment analysis of terms enriched in RNA targets bound by RBM17. (D) RT-PCR validated RBM17 regulated AS events in AML cells. (E-F) Representative RBM17-affected exon inclusion and exon skipping events validated by RT-PCR. The structure of each PCR product is indicated schematically on the right. Alternative exons affected by RBM17 are painted in orange.

Figure S4. EIF4A2 is required for AML cell survival. (A) Assessment of the effects of EIF4A2 knockdown on K562 cell growth. (B) Flow cytometric evaluation of apoptosis following EIF4A2 knockdown in HL60 cells.

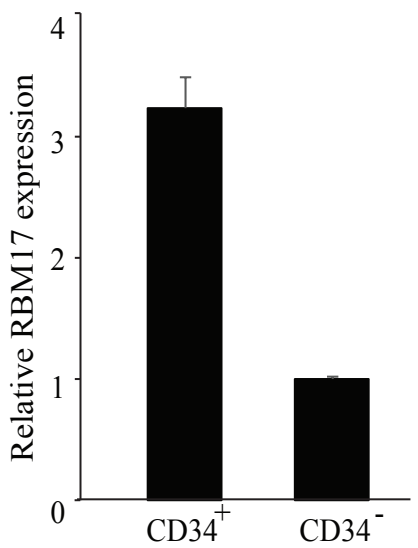
Figure S5. EIF4A2 overexpression partially rescues RBM17 knockdown-mediated apoptosis and differentiation phenotype in AML cells. (A) Representative histogram showing flow cytometry analysis of AnnexinV signal in HL60 cells on day 4 following co-expression of luciferase or EIF4A2 and knockdown of RBM17. (B) Quantification of CD11b expression in HL60 cells day 6 following co-expression of luciferase or EIF4A2 and knockdown of RBM17. Mean ± SD, n = 3.

References

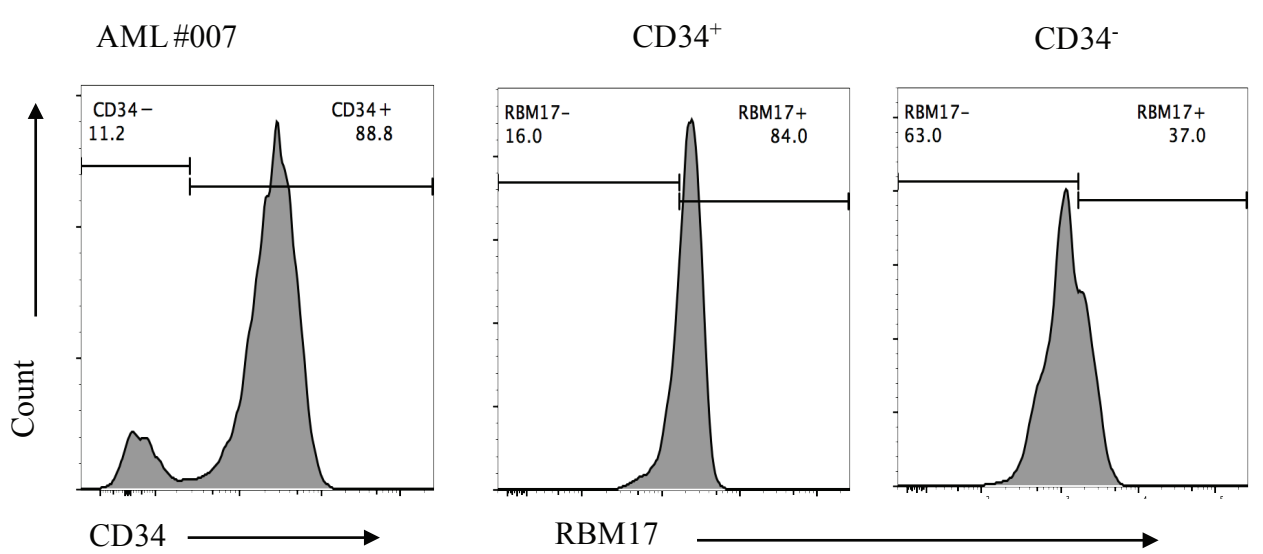
1. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
2. Shen, S. *et al.* rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci.* **111**, E5593–E5601 (2014).
3. maser. *Bioconductor* http://bioconductor.org/packages/maser/.
4. Brennan, P. drawProteins: a Bioconductor/R package for reproducible and programmatic generation of protein schematics. *F1000Research* **7**, 1105 (2018).
5. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

Figure S1

A



B



C

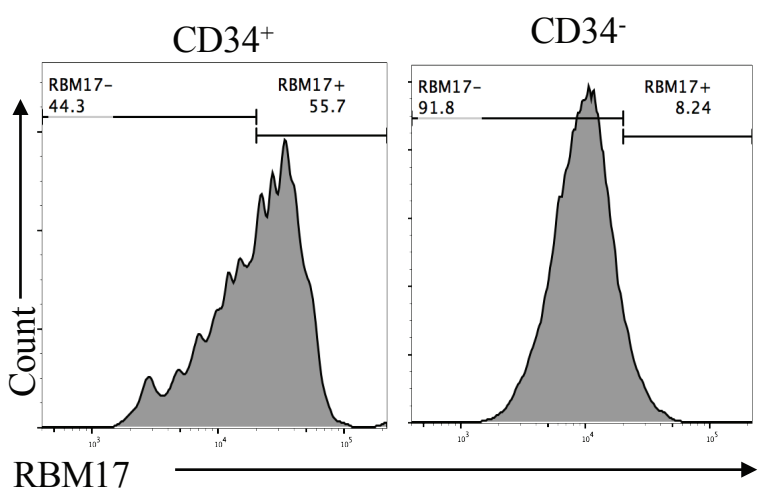


Figure S2

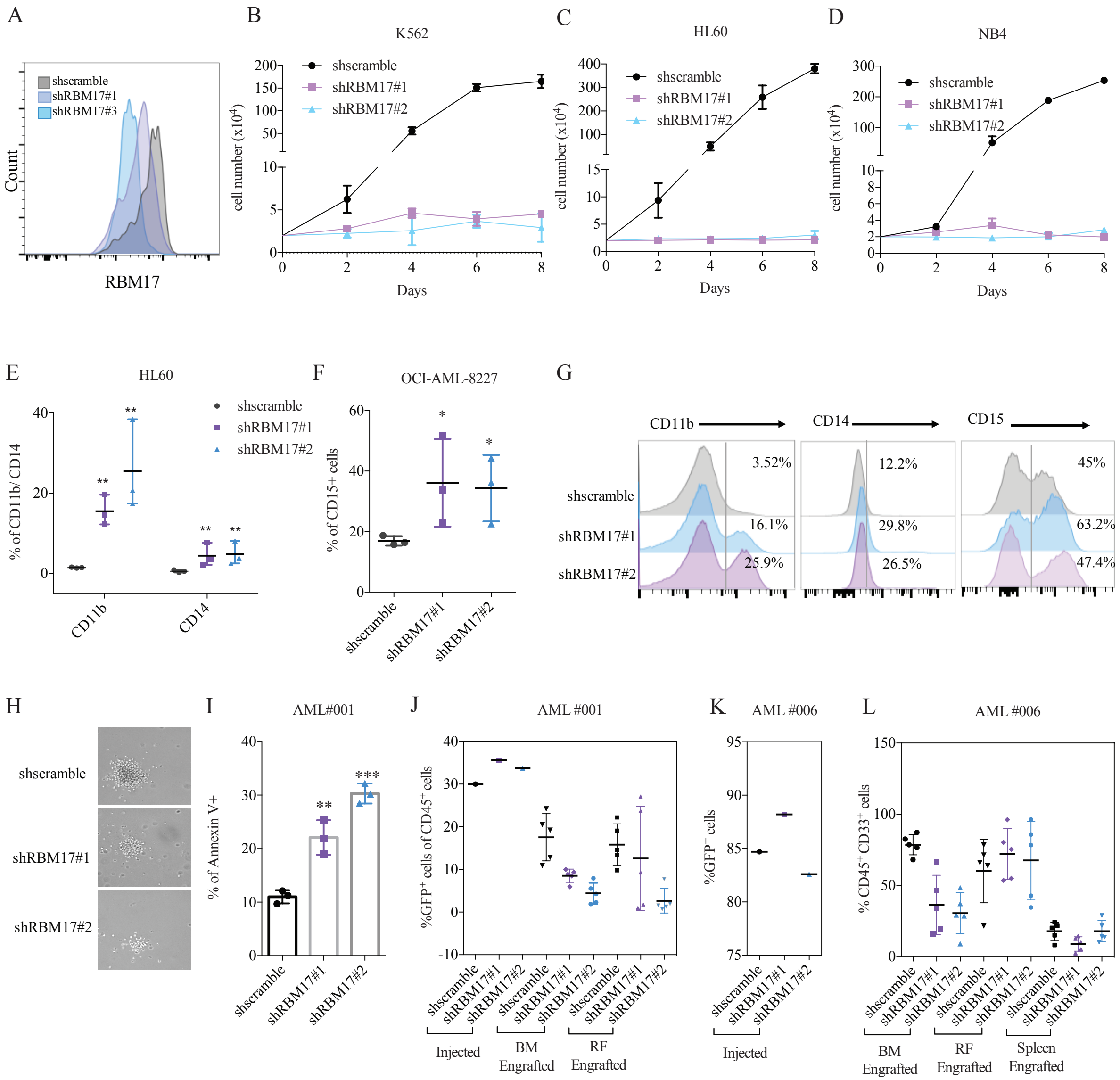
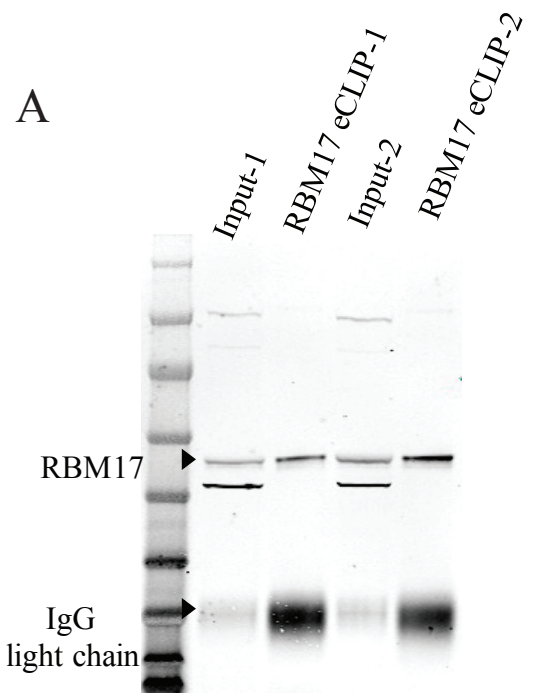
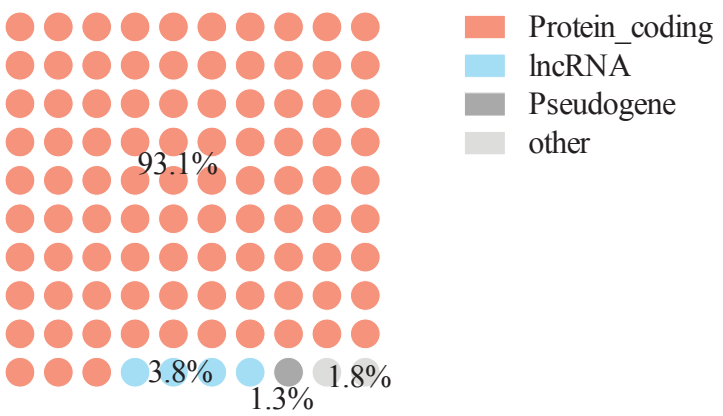

A
shscramble
shRBM17#1
shRBM17#3
Count
RBM17
B
K562
C
HL60
D
NB4
shscramble
shRBM17#1
shRBM17#2
cell number (x10^4)
Days
E
HL60
% of CD11b/ CD14
CD11b
CD14
F
OCI-AML-8227
% of CD15+ cells
G
CD11b
CD14
CD15
3.52%
12.2%
45%
16.1%
29.8%
63.2%
25.9%
26.5%
47.4%
H
shscramble
shRBM17#1
shRBM17#2
I
AML#001
% of Annexin V+
J
AML #001
%GFP+ cells of CD45+ cells
Injected
BM Engrafted
RF Engrafted
K
AML #006
%GFP+ cells
L
AML #006
% CD45+ CD33+ cells
Spleen Engrafted

Figure S3

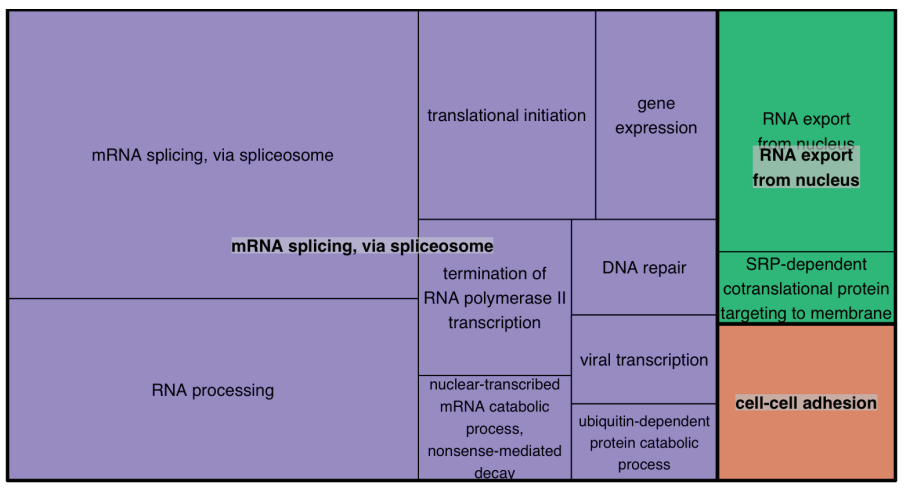
A



B



C

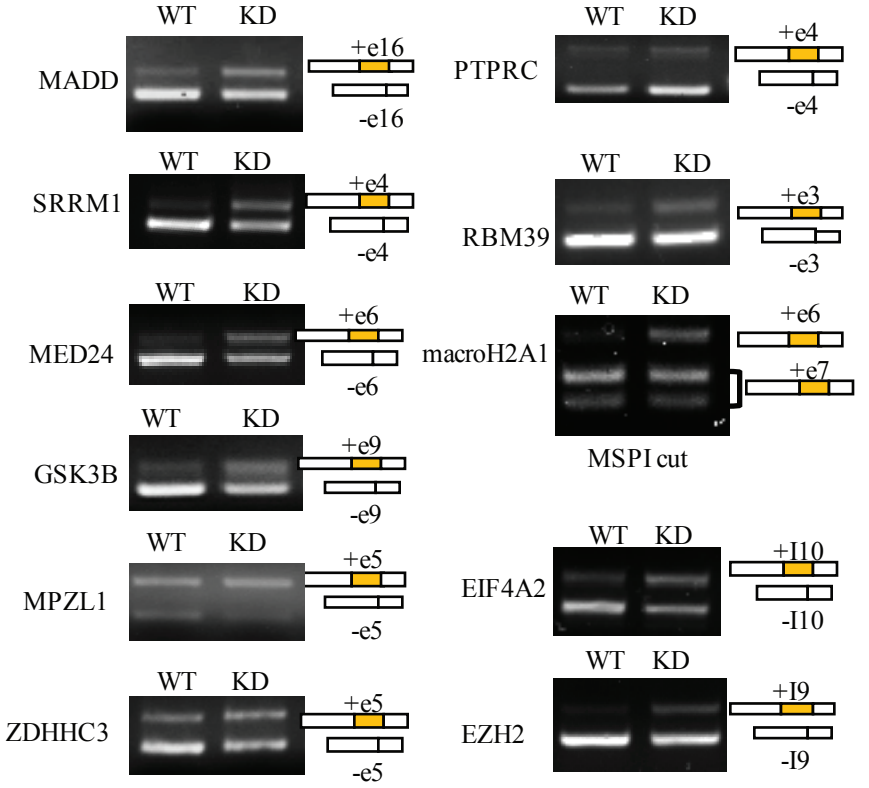


D

Validated 16 AS events impacted by RBM17 knockdown

Gene Symbol	AS event	Gene Symbol	AS event
MADD	Exon 16 exclusion	**macroH2A1**	MXE
SRRM1	Exon 4 exclusion	**EIF4A2**	Intron 10 retention
MED24	Exon 6 exclusion	**EZH2**	Intron 9 retention
GSK3B	Exon 9 exclusion	**MRPS18C**	Exon 3 inclusion
MPZL1	Exon 5 exclusion	**ZFYVE19**	Exon 7 inclusion
ZDHHC3	Exon 5 exclusion	**C4orf33**	Exon 3 inclusion
PTPRC	Exon 4 exclusion	**AP2B1**	Exon 13 inclusion
RBM39	Exon 3 exclusion	**USP33**	Alt 3' SS

E



F

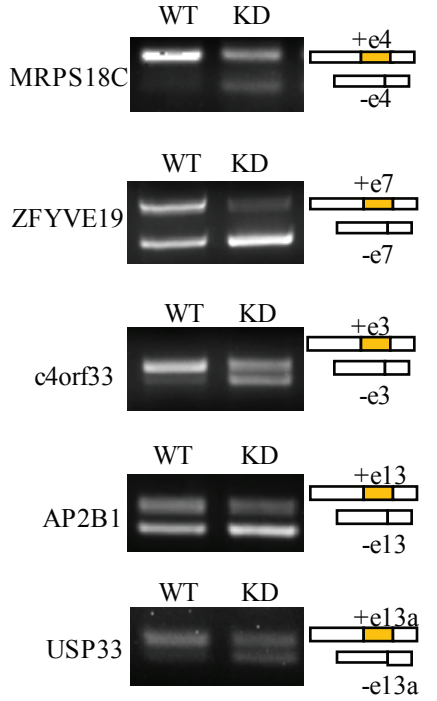
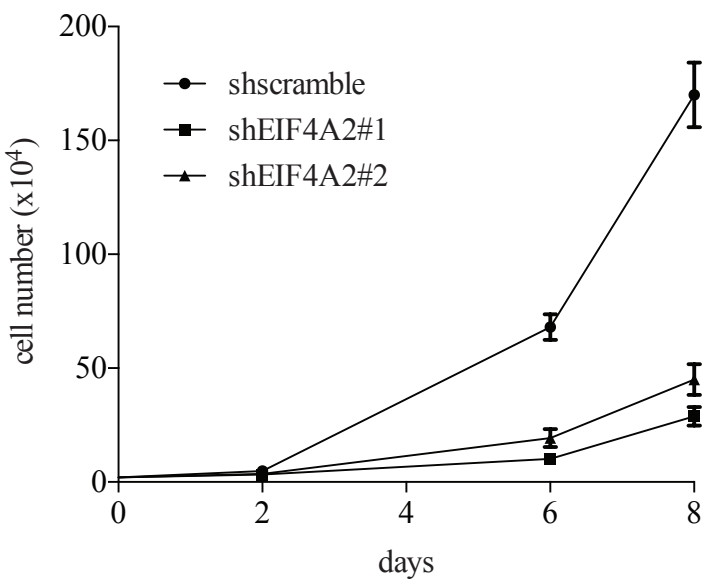


Figure S4

A


200
150
100
50
0
cell number (x10^4)
shscramble
shEIF4A2#1
shEIF4A2#2
0
2
4
6
8
days

B

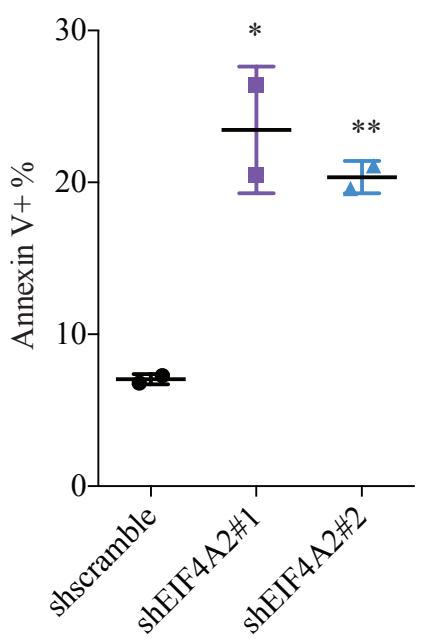
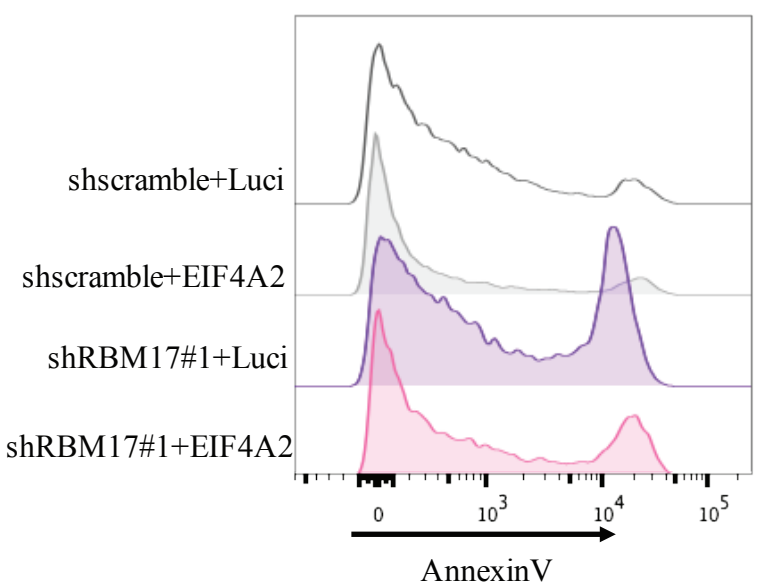
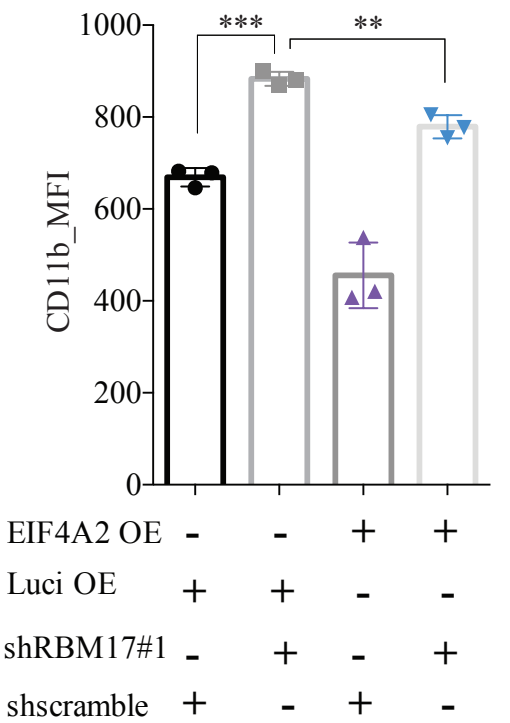

30
20
10
0
Annexin V+ %
*
**
shscramble
shEIF4A2#1
shEIF4A2#2

Figure S5

A


shscramble+Luci
shscramble+EIF4A2
shRBM17#1+Luci
shRBM17#1+EIF4A2
0
10^3
10^4
10^5
AnnexinV

B


***
**
1000
800
600
400
200
0
CD11b_MFI
EIF4A2 OE - - + +
Luci OE + + - -
shRBM17#1 - + - +
shscramble + - + -