Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

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

Lina Liu
McMaster University
Ana Vujovic
Princess Margaret Cancer Centre
Nandan Deshpande
University of New South Wales https://orcid.org/0000-0002-0324-8728
Shashank Sathe
University of California at San Diego
Govardhan Anande
University of New South Wales
Joshua Xu
Princess Margaret Cancer Centre
Mark Minden
University Health Network
Gene Yeo
University of California San Diego https://orcid.org/0000-0002-0799-6037
Ashwin Unnikrishnan
University of New South Wales https://orcid.org/0000-0001-5168-8755
Kristin Hope
University Health Network
Yu Lu (malixlu@gmail.com)
McMaster University https://orcid.org/0000-0003-4522-6823

Article

Keywords: acute myeloid leukemia (AML), chemo-resistance, leukemic stem cells (LSC)

Posted Date: July 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-728533/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Nature Communications on July 4th, 2022.
See the published version at https://doi.org/10.1038/s41467-022-31155-0.

Abstract

Chemo-resistance in acute myeloid leukemia (AML) patients is driven by leukemic stem cells (LSC) resulting in high rates of relapse and low overall survival. Here, we demonstrate that upregulation of the splicing factor, RBM17 preferentially marks and sustains LSCs and directly correlates with shorten patient survival. RBM17 knockdown in primary AML cells leads to myeloid differentiation and impaired colony formation and *in vivo* engraftment. Integrative multi-omics analyses show that RBM17 repression leads to inclusion of poison exons and production of nonsense-mediated decay (NMD)-sensitive transcripts for pro-leukemic factors and the translation initiation factor, EIF4A2. We show that EIF4A2 is enriched in LSC and its inhibition impairs primary AML progenitor activity. Proteomic analysis of EIF4A2-depleted AML cells shows recapitulation of the RBM17 knockdown biological effects, including pronounced suppression of proteins involved in ribosome biogenesis. Overall, these results provide a rationale to target RBM17 and/or its downstream NMD-sensitive splicing substrates for AML treatment.

Introduction

Acute myeloid leukemia (AML) is a malignant hematopoietic disorder with dysregulated clonal expansion of mutant undifferentiated myeloid progenitor cells, and accounts for approximately 30% of adult leukemias[1]. Despite significant advances in cancer therapeutics in recent years, adult AML patients continue to display chemo-resistance at presentation, high relapse rates, and a 5-year overall survival rate less than 25% [1]. AML is maintained by relatively rare populations of leukemic stem cells (LSC) that are responsible for seeding and propagating the disease[2–4] and possess stem cell-like characteristics including the capacity for self-renewal, differentiation potential (albeit limited), and relative quiescence[5–7]. This latter property of LSC, as well as their possession of natural resistance mechanisms such as drug efflux pumps, contributes to their intrinsic resistance to conventional chemotherapies that target proliferating cells. In addition, multiple studies have demonstrated that patients whose bulk AML cells have an elevated LSC gene expression signature have worse clinical outcomes[5,8], suggesting that heightened LSC activity correlates with poor efficacy of conventional therapy.

LSC and other primitive leukemic cells are generally thought to be transformed from hematopoietic stem cells (HSCs) or committed progenitor cells and very often share the same surface markers ($CD34^+CD38^-$ and $CD34^+$, respectively) and similar mechanisms that support the self-renewal[9] of their primitive normal counterparts. These similarities make it difficult to specifically target primitive leukemic cells for drug development. Recently a 78-patient study defined a panel of 17 LSC signature genes, whose expression levels were shown to be predictive of response to treatment and overall survival for patients treated with daunorubicin and cytarabine[8]. Despite these findings, there has been limited success in the effort to specifically target primitive leukemic cells for AML treatment. Therefore, it is essential to gain a more comprehensive understanding of the mechanistic elements that underpin primitive leukemic cell function and that, as such, may represent important and novel therapeutic targets in AML.

Alternative splicing (AS) is one of the major contributors to proteome diversity and is thus tightly controlled throughout normal development[10]. AS is a complex process that involves a variety of regulatory *trans-acting* splicing factors and responsive *cis*-acting RNA elements, which act together to determine splice site selection and alternative exon usage[11,12]. Aberrant alternative splicing is recognized as a key driver of cancer, with many of the hallmark processes of cancer being regulated by tumor-specific splice variants[13]. Dysregulated AS can either alter transcript stability, resulting in changes in protein levels or affect coding potential, leading to expression of proteins with distinctly different functions. In the context of AML, a genome-wide analysis of aberrant AS patterns showed that approximately one third of genes are differentially spliced in the primitive $CD34^+$ cells of AML patients compared to those obtained from normal controls, suggesting that such genes are involved in processes key to cellular function[14]. LSC also have a unique AS profile when compared to normal aging HSCs, including a switch to pro-survival isoforms, which enhances their maintenance[15]. For example, mis-splicing of GSK3β enhances the malignant transformation from human pre-leukemic progenitors into self-renewing LSC[16]. Aberrations in AS can result from somatic mutations in splicing factors or in *cis-acting* motifs within exons or introns, or abnormal expression of splicing factors. Analyses of the genomic landscape of AML patients have discovered recurrent mutations in splicing factors SRSF2, SF3B1 and U2AF1, however these mutations are only found in approximately 10% of AML patients studied[17–19]. Given that abnormal AS is also prevalent in AML patients with no obvious mutations in RNA splicing genes[14], it is critical to study the deregulation of splicing factor expression and their underlying mechanism in AML and primitive LSC. In the present study, we focused on aberrant AS in human primitive AML biology and show that RBM17 is preferentially expressed in progenitors and LSC and enacts within these cells an AS program that is critical for supporting their maintenance.

Results

RBM17 expression is associated with primitive AML cells and adverse AML prognosis. Previous studies examining the link between aberrant splicing and AML have focused on spliceosome genes with somatic mutations in AML patients or with abnormal expression levels in bulk AML samples[20,21]. To more broadly profile splicing factors that may mediate aberrant alternative splicing independently of mutations in AML cells, we performed a data-mining survey of 203 known mRNA splicing factors (members of the “mRNA splicing” and “mRNA alternative splicing” Gene Ontology (GO) categories[22] (Table S1). Strikingly, RNA-binding motif protein 17 (RBM17) was the only splicing factor that was both significantly elevated ($p=0.008$) in the LSC-enriched (LSC+) vs LSC-devoid (LSC-) subsets from 78 karyotypically normal AML patient samples (GSE76008)[8] and strongly linked ($p=0.00568$) to poor AML prognosis[19] (Figure 1A-C). We also analyzed the published gene expression profiles of purified LT-HSC ($Lin^-CD34^+CD38^-CD90^+$) from healthy donors and AML samples with normal karyotype (GSE35008)[23], and observed that RBM17 is expressed at significantly higher levels in AML LSC compared to normal LT-HSC (Figure 1D). Together these findings suggest that RBM17 is preferentially expressed in primitive AML cells. We went on to validate these results in 8 primary AML samples, along with a unique OCI-AML-8227 AML cell line, which

was derived from a primary AML sample and retains an LSC-driven hierarchy[24]. We found that the RBM17 transcript level is significantly upregulated in the primitive cell subset ($CD34^+$) as compared to the committed cell subset ($CD34^-$) in OCI-AML-8227 cells (Figure S1A). In keeping with the increased level of mRNA, RBM17 protein is 1.68 fold higher in the LSC-enriched primitive cell subsets ($CD34^+$) of primary AML patient samples (Table S2) compared to the more committed cell subsets ($CD34^-$) (Figure 1E, S1B-C), providing further support that elevated RBM17 preferentially marks the primitive compartments of human AML.

To further characterize RBM17 in AML patient AML cells, we analyzed its expression in the LAML-TCGA (https://portal.gdc.cancer.gov/projects/TCGA-LAML) dataset and found that RBM17 expression was significantly higher in both poor (n=38) and intermediate (n=76) molecular genetic risk groups compared with the good molecular genetic risk group (n=38) (Figure 1F). Next, to investigate the gene expression signature of AML patients with high expression of RBM17, we ranked AML patient samples from the GSE76008 dataset based on RBM17 expression level and defined the top 15% (35 of 221) as RBM17-high cases, and the bottom 15% (35 of 221) as RBM17-low cases. In total, we identified 832 differentially expressed genes (FDR$\leq$ 0.05, FC$\geq$2 or $\leq$ 0.5), including 336 transcripts more abundant, and 496 transcripts less abundant in AML patient samples with higher RBM17 expression (Figure 1G). Interestingly, of these genes, 82.1% of those co-upregulated with RBM17 in AML are more highly expressed and 73.2% of the genes anti-correlated with RBM17 are expressed at lower levels in LSC+ populations (Figure 1H), implicating that elevated expression of RBM17 and its associated co-regulated genes overlap with gene expression signature of LSC. To confirm this hypothesis, we carried out Gene Set Enrichment Analysis (GSEA) with the LSC gene set that contains upregulated (UP) and downregulated (DN) genes in LSC[8], and showed highly significant enrichment of the LSC signature gene set in genes co-upregulated with RBM17 in AML patients (Figure 1I). In addition, GSEA also revealed significant enrichment of genes involved in ribonucleoprotein complex biogenesis and spliceosomal complex assembly (FDR$\leq$ 0.05) (Figure 1J) in the set of genes co-upregulated with RBM17. These results together suggest that RBM17 could have an important role in supporting primitive leukemic cell functions.

RBM17 knockdown impairs the stem and progenitor potential of primary AML. To examine the functional roles of RBM17 in primitive AML cells, we knocked down RBM17 using short hairpin RNA (shRNA) in human AML cell lines and patient samples. Briefly, we designed lentiviruses encoding GFP (transduction marker) along with two independent shRNAs, both of which resulted in efficient knockdown of RBM17 after transduction in multiple AML cell lines (shRNA #1 and #2) (Figure 2A, S2A). Interestingly, RBM17 knockdown inhibited AML cell growth (Figure S2B-D) and induced myeloid differentiation in the HL60 AML cell line (Figure S2E), the aforementioned OCI-AML-8227 cell line (Figure S2F), and primary AML cells (Figure S2G). In addition, RBM17 knockdown in primary AML cells significantly reduced colony numbers (Figure 2B-F, Figure S2H), and impaired survival (Figure S2I), indicating that RBM17 is required for recognizing AML progenitor potential *in vitro*. Next, to directly assess the role of RBM17 in LSC growth and survival *in vivo*, we performed xenograft studies using shRBM17-transduced primary AML specimens. We conducted output/input analysis of the GFP positivity of human ($CD45^+$) cells for AML

sample #001 (Figure 2G, S2J) where we achieved a ~35% infection rate, and analysis of the percentage of total $CD45^+$ cells for AML sample #006 where transduction reached saturation (>85%) (Figure S2K-L). We observed that RBM17 knockdown in these two primary AML samples greatly impeded AML engraftment in transplanted immunodeficient mice (Figure 2H-I). Analysis of the resulting grafts revealed that RBM17 knockdown induced myeloid differentiation *in vivo* as shown by an increased percentage of mature CD14+ cells in shRBM17 grafts compared to controls (Figure 2J-K). Together, these findings indicate RBM17 depletion disrupts primitive AML cell function through enhancing differentiation and inhibiting colony-forming and engraftment capacities.

RBM17 controls alternative splicing of genes involved in multiple pathways in AML cells. To understand the molecular mechanisms that underlie the supporting role of the splicing factor RBM17 in AML, we first performed RBM17 enhanced crosslinking immunoprecipitation (eCLIP)-seq in the K562 myeloid leukemia cell line to identify genome-wide RNA targets bound by RBM17 (Figure S3A). eCLIP-seq analysis identified 866 significantly enriched reproducible binding peaks for RBM17 in the genome using a cutoff of FDR<0.05 and log2 (FC)>3 (over size-matched input control), which corresponded to 432 annotated transcripts (Figure 3A and Table S3). Of these transcripts 93.1% are protein coding genes (Figure S3B). The majority of these peaks are within the coding sequence (CDS, 20.8%), 5'-splice site (5'-SS, 21.9%), or proximal intronic regions which are closer to intron/exon boundaries (30.4%) (Figure 3B). The enrichment of RBM17 binding peaks around splice sites is consistent with its known function as a splicing regulator. Through motif analysis, we also identified that highly G-enriched motifs mapped to RBM17-binding sites (Figure 3C). GO analysis of enriched binding sites further showed that transcripts involved in mRNA splicing, RNA processing, translation initiation, DNA repair and protein ubiquitination are preferentially bound by RBM17 (Figure S3C).

To further identify possible functional consequences of these interactions, we analyzed a published ENCODE RNA-seq dataset of shRBM17 (GSE88633) vs Control (GSE88047) transduced K562 cells[25]. We discovered AS events that are affected by RBM17 knockdown in K562 cells (FDR<0.1, ΔPSI>0.05) (Table S4), among which exon inclusions of 705 splicing events were supported by RBM17 while the other 633 splicing events were repressed by RBM17 (Figure 3D). We also observed that RBM17 can be involved in many types of AS events, including Cassette exon (CE), Retained intron (RI), Alternative 3' splice site (A3SS), Alternative 5' splice site (A5SS) and Mutually exclusive exon (MXE), with cassette exons being the most affected (Figure 3E). We next validated that multiple RBM17-regulated AS events are shared by K562 and HL60 cell lines (Figure S3D-F). GO analysis revealed that RBM17-affected AS events are involved in a wide range of pathways ranging from RNA processing, cell cycle, intracellular localization, to metabolism and biosynthesis pathways[26] (Figure 3F).

Integrative multi-omics approaches indicate RBM17-mediated splicing prevents NMD of genes required for leukemic growth. Analogous to genetic mutations, inclusion or exclusion of certain exons or introns can change the reading frame, which would potentially affect conserved regions of the coded protein structure or have deleterious effect on subsequent mRNA translation. To systematically analyze the potential functional links of RBM17-affected splicing events in AML, we applied a published

bioinformatics pipeline to predict effects on the corresponding protein upon RBM17 depletion[27]. Through our analysis, we identified 88 splicing events yielding changes in “Transcript Biotype” and 70 splicing events resulting in changes in “Protein Domain” through this analysis (Figure 4A, Table S5). Intriguingly, 13.3% (21/158) of these splicing events cause a complete or partial loss of well annotated protein domains, while a further 32.3% (51/158) of the splicing events are predicted to produce nonsense-mediated decay (NMD) sensitive transcripts mainly due to the inclusion of poison exons, and the formation of premature termination codons (PTCs) (Figure 4A). By overlapping the eCLIP-seq dataset and these 51 NMD sensitive transcripts, we identified 6 alternatively spliced transcripts that are predicted to be both bound and regulated by RBM17 (Figure 4B), suggesting that RBM17 can regulate splicing through direct and specific binding to the pre-mRNA of these transcripts. Such relationships are in contrast to the previously reported role of RBM17 as a spliceosome component with no preference for its RNA substrate sequences[28]. It is particularly striking that, within these 6 direct splicing targets of RBM17 are EZH2[29,30], RBM39[20,21,31] and HNRNPDL[32], all factors that have known important roles in cancer stem cell self-renewal and myeloid malignancy. Together, these studies suggest that RBM17 knockdown leads to NMD of genes involved in leukemia propagation.

To validate whether these RBM17-mediated NMD-sensitive splicing events cause corresponding protein downregulations, we applied LC-MS proteomics to characterize proteome changes after RBM17 knockdown in the K562 cells. At day 5 after lentiviral transduction for RBM17 knockdown, we identified 741 proteins with significant changes (SigB<0.05) (Table S6). GO analysis showed that these proteins downstream of RBM17 knockdown are enriched in clusters of functional networks representing cell division, RNA processing, autophagy, DNA replication and DNA repair, translation and protein folding and vesicle organization (Figure 4C). Importantly, when we overlap RBM17-mediated NMD-sensitive splicing targets with the proteomics dataset, among 44 transcripts with measured protein expression values (the other 7 genes were not detected in the proteomics dataset), we demonstrated that 31.8% (17/44) of them are downregulated by RBM17 knockdown at their protein levels (fold change<0.9) (Figure 4D). GO analysis of these 17 proteins downregulated by RBM17 knockdown through potential NMD further revealed a cluster of RNA processing and RNA splicing genes (Table S7), suggesting that RBM17 regulates a network of RNA-processing proteins involved in RNA homeostasis that both directly and indirectly influences cancer stem cell biology. We next performed bootstrapping analysis by taking random sets (repeated 10,000 times) of 44 proteins out of the list of 8825 proteins identified in our shRBM17 versus shscramble proteomics experiments to calculate the percentage of proteins that were down-regulated (fold change<0.9) within these 10,000 randomly picked 44-protein sets. Strikingly, the median percentage of down regulated proteins from randomly-picked 10,000 runs was 22.2% (the mean percentage is 23.4%), which is significantly lower than 31.8% observed for predicted NMD sensitive transcripts ($p < 2.2e-6$) (Figure 4E), suggesting that RBM17 indeed controls protein expression through regulating alternative splicing coupled with NMD.

RBM17 suppresses EIF4A2 poison exon inclusion. Our integrative multi-omics RNA-interactome, transcriptome and proteome profiling analyses revealed that EIF4A2 (eukaryotic translation initiation

factor 4A2) is the top direct NMD-sensitive splicing target of RBM17 in leukemic cells (Figure 5A). RBM17 binds to EIF4A2 intron 10 and RBM17 knockdown promotes the inclusion of a proximal “poison” cassette exon (*chr3:186788310-186788416*), which is associated with strong depletion of EIF4A2 protein (Figure 5B). This 107-bp cryptic exon and its flanking intronic sequences are highly conserved across vertebrates (Figure 5C), strongly suggesting that it has a regulatory function[33]. Therefore, we aimed to explore its function as a potential effector of RBM17. We first validated using isoform-specific RT-PCR in AML cell lines and primary AML samples that RBM17 knockdown promoted EIF4A2 poison exon inclusion (Figure 5D-E). Next, to confirm NMD-sensitivity of this EIF4A2 transcript variant that includes the poison exon, we tracked its mRNA decay level after actinomycin D-induced transcription-halt in cells with depletion of UPF1, a protein required for NMD. We demonstrated that the mRNA level of the poison exon- included EIF4A2 variant dropped less with UPF1 knockdown compared to shLuci control (Figure 5F-G). Conversely, the mRNA level of the poison exon-skipped EIF4A2 variant dropped similarly following UPF1 knockdown as compared to control (Figure 5H). These results confirmed that RBM17 suppressed the EIF4A2 poison exon inclusion event, which would otherwise trigger EIF4A2 degradation through NMD. Lastly, we validated that RBM17 knockdown in a panel of AML cell lines consistently reduced EIF4A2 protein expression (Figure 5I). Taken together, our results demonstrate that RBM17 is required for the generation of productive protein-coding transcripts of many pro-leukemic factors and identify EIF4A2 as a *bona fide* direct downstream target of RBM17.

EIF4A2 is elevated in human LSC and is required for leukemogenesis. EIF4A2 encodes an ATP-dependent RNA helicase, which is a subunit of the EIF4F complex involved in ribosome binding to mRNA substrates and scanning for the initiator codon [34]. Intriguingly, through data analysis of the TCGA dataset[35], we found that EIF4A2 mRNA was more highly expressed in all subtypes of AML than in normal monocytes (Figure 6A). Importantly, in the context of LSC, just as in the case of RBM17, EIF4A2 is preferentially expressed in LSC-enriched cell fractions compared to LSC-devoid fractions from AML patients at both mRNA and protein levels (Figure 6B-C). Consistent with RBM17 expression in LT-HSC, EIF4A2 is also more highly expressed in AML cells with a primitive immunophenotype (Lin-CD34^{+}CD38^{-}CD90^{+}) compared to normal control HSCs (Figure 6D). Given our demonstration that EIF4A2 is downstream of RBM17, we next aimed to explore the effect of EIF4A2 knockdown on primitive AML cell function. Towards this end, we depleted EIF4A2 in AML cell lines and in primary AML samples using two independent shRNAs (#1 and #2) (Figure 6E). Strikingly, knockdown of EIF4A2 significantly inhibited AML cell growth (Figure 6F, S4A), induced myeloid differentiation (Figure 6G-H) and resulted in increased cell apoptosis (Figure 6I, S4B) as compared to a shscramble control. In addition, depletion of EIF4A2 in three primary AML samples significantly inhibited their colony-forming abilities (Figure 6J-L). These data together indicate that EIF4A2 supports the proliferation, survival and undifferentiated state of AML cells.

EIF4A2 overexpression partially rescues RBM17 knockdown-mediated phenotypes in AML cells. Through correlation analysis, we found that RBM17 supports higher expression of EIF4A2 in two different AML patient datasets (Figure 7A-B). To test the extent that the downstream effects of RBM17 knockdown are shared upon EIF4A2 knockdown, we first performed LC-MS proteomics to characterize proteome changes

induced by EIF4A2 knockdown in K562 cells. We identified a list of significantly downregulated and upregulated proteins induced by EIF4A2 knockdown (Table S8). Interestingly, these two gene sets are significantly enriched in proteins modulated downstream of RBM17 knockdown as we observed in K562 cells (Figure 7C-D), suggesting that EIF4A2 knockdown indeed largely recapitulates the biological effects of RBM17 knockdown in AML. Given the significant link between EIF4A2 and RBM17 in AML, we next tested whether restoring EIF4A2 could rescue any biological effects caused by RBM17 knockdown. Specifically, we infected HL60 cells with lentiviruses co-expressing a scramble or RBM17 targeting hairpin with either a luciferase control cDNA or the EIF4A2 cDNA (shscramble+Luci, shscramble+EIF4A2, shRBM17#1+Luci, shRBM17#1+EIF4A2) (Figure 7E). We found that overexpression of EIF4A2 efficiently reversed the adverse effects of RBM17 knockdown on AML cell apoptosis, and partially rescued AML cell differentiation induced by RBM17 knockdown (Figure 7F, S5A-B).

Interestingly, our subsequent GSEA analysis also revealed that both RBM17 and EIF4A2 knockdown in K562 cells strongly alter expression of proteins enriched in ribosome biogenesis-related gene sets (Figure 7G-H), and several known translation-related factors (Figure 7I), indicating that downregulation of RBM17 and EIF4A2 may both affect mRNA translation in leukemic cells. To confirm this finding, we performed an O-propargyl-puromycin (OPP) based protein synthesis assay and detected significantly decreased mRNA translation activity in both RBM17- and EIF4A2-knockdown AML cells, respectively (Figure 7J-K). Moreover, we demonstrated that elevation of EIF4A2 level in RBM17-knockdown AML cells partially rescued the protein synthesis rate (Figure 7L). These results together suggest that RBM17 inhibits AML apoptosis and differentiation and supports mRNA translation at least partially through enforcing the expression of an NMD-resistant transcript variant and promoting the expression of EIF4A2 protein in human leukemic cells.

Discussion

RBM17, also known as splicing factor 45kDa (SPF45), was originally identified as a component of the spliceosome complex. It co-localizes with SR proteins in nuclear speckles and regulates the second step of pre-mRNA splicing by selecting alternative AG splice acceptor sites [36]. RBM17 protein expression is limited in normal tissues and is greatly increased (5-10 fold) in solid tumors of the bladder, lung, colon, breast, ovary, pancreas, and prostate [37]. In addition, RBM17 has been previously linked to cancer chemotherapy resistance in breast and ovarian cancer cell lines through unspecified mechanisms[38,39]. However, the role of RBM17 in AML has not been explored. Interestingly, using proteomics, we previously found that RBM17 is upregulated in human PSCs compared to terminally differentiated fibroblasts and is required to support PSC self-renewal[40], a core feature shared in both normal and cancer stem cells. Our work identified RBM17 as the sole mRNA splicing factor that is both upregulated in LSC-enriched cell fractions and is significantly associated with poor prognosis of AML patients, highlighting the likelihood that RBM17 contributes to the aberrant AS program found in the primitive cells that drive the disease. Through the use of gold-standard in vivo repopulation assays with primary AML samples, we demonstrated that RBM17 depletion impairs the function of disease- and relapse-initiating primitive AML

cell compartment. Together these data position RBM17 as a novel leukemic stem cell regulator whose expression and targeting may have important implications in the diagnosis and treatment of malignant hematopoiesis.

A previous study of splicing in mouse neurons showed that RBM17 normally represses the splicing of cryptic junctions and its loss leads to the inclusion of intronic elements in mature transcripts [41]. Exonization of intronic coding cassettes normally creates frameshifts or introduces PTCs[42,43]. Our integrative multi-omics analysis of RBM17 uncovered, for the first time, that RBM17 depletion promotes inclusion of poison cassette exons or introns for a number of pro-leukemic factors and leads to their NMD-mediated mRNA degradation and subsequent protein-level downregulation. Our results identified the pro-leukemic factors RBM39, EZH2, and HNRNPDL as direct RBM17 mRNA-binding targets with these interactions serving to preserve their protein levels through exclusion of poison exons. A recent study using CRISPR/Cas9 screening demonstrated that complete loss of RBM39 suppresses AML growth both *in vitro* and *in vivo*, while pharmacologic RBM39 degradation results in broad anti-leukemic effects[20]. In the same study it was also found that RBM39 loss affects splicing of mRNAs related to RNA-splicing, export, and metablism[20]. Similarly, EZH2 is an important regulator of normal and malignant hematopoiesis[44], while HNRNPDL overexpression in CML cells has been shown to induce leukemia *in vivo*[32]. These findings indicate the possibility that RBM17 knockdown-induced inhibition of these factors contributed to anti-leukemic effects. Our work has therefore provided mechanistic insights into essential AML molecular circuitry by uncovering that the elevated expression of RBM17 serves to selectively represses the formation of PTC containing mRNAs required for supporting LSC function.

Interestingly, our proteomics data showed that RBM17 knockdown causes downregulation of its known spliceosome interactors CHERP and U2SURP (Table S7). This is in line with a previous study in human HEK293T cells showing that RBM17, CHERP and U2SURP reciprocally regulate each other's expression level and share downstream splicing targets enriched for RNA-binding proteins[28]. We speculate that RBM17, and the spliceosome complex it interacts with, collectively block the usage of cryptic splice sites around cassette exons or introns containing PTCs and skip their inclusions. Furthermore, our bootstrapping analysis indicated that RBM17 knockdown leads to downstream gene expression inhibition through splicing-coupled NMD, suggesting that RBM17 preferentially regulates splicing of mRNAs containing PTCs. Exploration of the mechanisms through which RBM17 mediates its specific splicing of NMD-sensitive transcripts will be of interest to pursue in future studies.

In the present work, we showed that RBM17 represses the inclusion of poison intron10 in the EIF4A2 pre-mRNA, which prevents EIF4A2 mRNA NMD and promotes its downstream protein synthesis. Therefore, EIF4A2 represents a novel potential therapeutics target for AML and intersection between splicing and translation control. Previously, Sadlish et al found that the natural compound rocaglamides stabilizes EIF4A-RNA interactions and interferes with the assembly of the EIF4F complex, thereby blocking translation initiation[45]. More recently, Callahan and colleagues showed that rocaglamide is able to preferentially kill functionally defined LSC, but relatively spares normal HSPCs through mechanisms

beyond simply inhibiting translation initiation[46]. Since rocaglamide does not distinguish EIF4A family members, it has not been clear which member of the EIF4A family underlies the anti-leukemia effects of the compound. Our work clearly demonstrates that EIF4A2 is more highly expressed in LSC and is required to support the proliferation, survival and undifferentiated state of AML cells, indicating the RBM17/EIF4A2 axis we have uncovered is indeed targetable for AML treatment. Importantly, a previous comparative EIF4A1 and EIF4A2 RIP-seq study in HEK293 cells also showed that 23% of EIF4A2's RNA targets are unique[47]. GO analysis of the EIF4A2-specific RNA targets further showed that these genes are involved in transcription, cell migration, cell cycle and positive regulation of GTPase activity. Many of EIF4A2's unique RNA targets, such as USP6NL (a GTPase-activating protein) and REV1, both of which are significantly upregulated in LSC, are indeed downregulated by EIF4A2 knockdown in K562 cells (Table S7). Thus, understanding the functional role of the EIF4A2-specific RNA targets in AML and LSC may provide mechanistic support for specifically targeting EIF4A2 and/or its regulated pathways for AML treatment and directing the improvement of drug target sensitivity.

Upregulation of protein synthesis has been described to occur in "pre-leukemic" myelodysplastic (MDS) stem cells[48]. AML stem cells also exhibit increased expression of ribosome pathway genes[49], indicating the potential role of ribosome biogenesis in the establishment and propagation of cancer stem cells in the blood system. Importantly, both RBM17 and EIF4A2 knockdown inhibited protein synthesis and downregulated at the protein level, the expression of factors enriched in the ribosome biogenesis pathway, suggesting a link between elevated expression of RBM17 along with its downstream target EIF4A2 and protein synthesis activation in primitive AML cells. Our EIF4A2 rescue experiments support the concept that EIF4A2 inhibition is necessary for the shRBM17-induced apoptosis and contributes to shRBM17-induced myeloid differentiation and translation inhibition. Interestingly, our proteomics analysis showed that RBM17 knockdown also led to downregulation of other translation-related factors including UBA52, EIF4H and EIF3B (Table S6), the collective loss of which may synergize with that of EIF4A2 to further solidify the translation inhibitory effects induced by RBM17 loss.

In conclusion, human AML stem and progenitor cells express abnormally high levels of RBM17 to ultimately enforce NMD-escape for a number of key pro-LSC transcripts (Figure 8). In particular, we place this mechanism of RBM17-directed control upstream of the essential process of protein synthesis in LSC and identify inhibition of the RBM17-EIF4A2 axis as a potential therapeutic avenue for AML treatment.

Materials And Methods

Mice

NOD-*scid*-*IL2Rγc*$^{-/-}$ (NSG) (Jackson Laboratory) mice were bred and maintained in the Stem Cell Unit animal barrier facility at McMaster University. All procedures received the approval of the Animal Research Ethics Board at McMaster University.

Primary CB and AML patient samples

All CB and AML patient samples were obtained with written informed consent and with the approval of the local human subject research ethics board at the University Health Network and McMaster University in accordance with Canadian Tri-Council Policy Statement on the Ethical Conduct for Research Involving Humans (TCPS). Following Ficoll-Paque separation mononuclear cells were stored in the vapor phase of liquid nitrogen in 10% DMSO, 40% FCS and alpha MEM. Primary samples were thawed in PBS 10% FBS with 100 µg/ml DNAse (AML only) prior to using in *in vitro* and *in vivo* assays.

Cell culture, cell lines and flow cytometry

The human promyelocytic leukemic cell line NB4, myeloblastic cell line HL60, and chronic myelogenous leukemic cell line K562 were cultured in RPMI-1640, supplemented with 10% fetal bovine serum. HEK 293FT and LentiX-293T cells were cultured in Dulbecco's modified Eagle's medium, supplemented with 10% fetal bovine serum. OCI-AML-8227 cells were cultured in StemSpan SFEM II (Stem Cell Technologies) with Interleukin (IL-3, 10ng/ml, R&D Systems), stem cell factor (SCF, 50ng/ml, R&D Systems), Interleukin-6(IL-6, 10ng/ml, R&D Systems), FMS like tyrosine kinase 3 ligand (FLT3, 50 ng/mL, R&D Systems), Thrombopoietin (TPO, 25 ng/mL, Peprotech), and Granulocyte Colony-Stimulating Factor (G-CSF, 10ng/ml, Peprotech). OCI-AML-8227 cells were plated 400,000 cells/ 24 well non-adherent culture plate and were split every 6-7 days. Primary AML samples were grown in X-VIVO with 20%BIT, 100 ng/mL human SCF, 100 ng/mL human FLT3, 20 ng/mL human TPO, 20 ng/mL human IL3 and 10 ng/mL human IL3. Human cord blood derived hematopoietic stem and progenitor cells (HPSCs) were cultured in StemSpan SFEM with 20ng/mL human IL-6, 100ng/mL human SCF, 20 ng/mL human TPO and 100ng/mL human FLT3. All cells were incubated at 37°C in a humidified atmosphere containing 5% CO_2. All flow cytometry analysis was performed using a BD LSRII flow cytometer, MACSQuant Analyzer and FlowJo Software (v7.6.5). Cell sorting was performed with MoFlo XDP (Beckman Coulter).

AML transduction

For AML cell lines HL60, K562 and NB4, cells were infected with lentivirus using an MOI of 10 followed by puromycin selection or 7-AAD- and GFP+ sorting. For primary AML samples, 1.5 million cells were infected with lentivirus using an MOI of 50 in 24 well ultralow attachment plates with 500ul total growth media, another 500ul of growth media were added 16 hours after infection followed by 7AAD- and GFP+ sorting.

Colony-forming unit assay

Sixteen hours (for primary AML) post lentiviral infection, freshly sorted 7- AAD- and GFP+ (healthy shRNA-expressing) cells were plated in triplicates in Human Methylcellulose Complete Media (ColonyGEL #1102) and grown in 37°C, 5% CO_2 incubator for 10-14 days before imaging and counting. For primary AML, ~7K-30K cells per ml were plated.

Xenograft studies

NSG mice were sublethally irradiated (315 cGy) 1 day prior to injection. Pre-validated engrafting AML sample were infected with shRNA for 24 hours in 24-well culture plates at a multiplicity of infection of 50. Post transduction, cells were validated for GFP expression (monitored sample of cells until 3 days after transduction), washed, resuspended in IMDM + 1% FBS. Then ~200,000 cells were injected with 30ul of IMDM + 1% FBS into the right femur of each recipient mouse, 5 mice were injected per experimental group. 9-12weeks post-transplant, mice were sacrificed and bone marrow from tibias, femurs and pelvis was harvested, crushed with mortar and pestle, filtered and red blood cell lysed using ammonium chloride buffer. Human AML engraftment was analyzed by blocking reconstituted mouse bone marrow with mouse Fc block (BD Biosciences) and human IgG (Sigma), followed by staining with fluorochrome-conjugated antibodies against human CD45-BV-421, CD33-PE, CD34-APC, CD14-APC-H7, CD11b-BV605 (BD Biosciences) for graft analysis.

Apoptosis assay and Differentiation assay

To measure apoptosis, cells were washed with PBS and incubated with V450-Annexin V (APOAF, sigma) or anti-V450-Annexin V (560506, BD) and PI or 7-AAD in the Annexin V binding buffer in a reaction volume of 100ul for 15 minutes according to the manufacture's instruction analyzed by FACS. To monitor cellular differentiation status, cells were stained with the following antibodies: PE-CD11b (557321, BD), FITC-CD14 (551376, BD), APC-H7-CD14 (561384, BD), APC-CD15 (551376, BD).

Intracellular flow cytometry

Primary AML cells and OCI-AML-8227 cells were initially stained with anti-CD34 APC (555824, BD Biosciences) antibody and LIVE/DEAD Fixable Green (L34969, Thermo Fisher) and then fixed with the Cytofix/Cytoperm kit (BD Biosciences) according to the manufacturer's instructions. Fixed and permeabilized cells were immunostained with anti-RBM17 rabbit antibody (ab204333, Abcam) and detected by Alexa-Fluor 405 goat anti-rabbit IgG(H+L) secondary antibody (Thermo Fisher) through FCAS. For RBM17 knockdown efficiency test, sorted cells were initially stained with LIVE/DEAD Fixable Near-IR (L34975, Thermo Fisher) for 30 minutes and then fixed with the Cytofix/Cytoperm kit. Fixed and permeabilized cells were immunostained with anti-RBM17 rabbit antibody and detected the same way as described above.

eCLIP-seq Preparation

RBM17 eCLIP studies were performed in duplicate according to the published eCLIP-seq experimental procedures[34]. 20 million K562 cells were washed in PBS and UV crosslinked at 400 mJoules/cm^2 with 254 nm radiation on ice, pelleted and snap frozen. Cells were then lysed in iCLIP lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, 1:100 Protease Inhibitor Cocktail) and treated with Turbo DNase and RNase I to fragment RNA, and supernatants from lysates were collected for immunoprecipitation. 10 µg of RBM17 antibody mixture (Bethyl A302- 497A and A302-498A) for each sample was then pre-coupled to 125ul of sheep anti-rabbit Dynabeads (LifeTech), added to lysate and incubate overnight at 4°C. Prior to washing, 2% of the sample (with beads-RBM17 antibody-

cell lysate) was taken as the size match input sample, with the remainder magnetically separated and washed twice with cold High salt wash buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) and wash buffer (20 mM Tris- HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20). After RNA adapter ligation, IP-western, eCLIP was performed by excising the area from 50 kDa to 125 kDa for Input and IP samples. Reverse transcription, DNA adapter ligation, and PCR amplification were performed according to the standard eCLIP experimental procedure. Three libraries (Input and RBM17 eCLIP triplicates) were prepared and sent for paired-end 75-bp Illumina sequencing.

Patient database and RNA-seq analyses

Gene expression data of 138 xenotransplant defined LSC-enriched and 89 non-LSC subsets from 78 AML were obtained from GEO (GSE76008)[8]. Gene expression data on sorted LT-HSC (most primitive hematopoietic cells) from AML patients and healthy controls were obtained from GEO (GSE35008)[23]. Protein expression of xenotransplant validated LSC-enriched and non-LSC fractions from 6 AML samples were obtained from PRIDE (Project PXD008307)[50]. Gene set enrichment analysis (GSEA) was performed by compassion of the “RBM17 high” and “RBM17 low” gene expression profile with the published LSC gene signature (GSE76008)[8]. RNA-seq data of shRBM17 or Control transduced K562 cells were downloaded from GSE88633 and GSE88047. Data was pre-processed and filtered using standard parameters. Samples were mapped to Human genome (UCSC HG38) with default settings and alternative splicing analyses were performed with rMATS. Functional switches were identified using in-house pipelines. Full methodological details are provided in the Supplemental Methods.

Proteomics analyses

Proteomics experiments were performed and analyzed using methods previously described[51]. Full methodological details are provided in the Supplemental Methods.

Statistical Analysis

Sample sizes are indicated in relevant figures. Experiments were repeated at least three times. All statistical analysis was performed using GraphPad Prism (GraphPad Software version 6.0). Unpaired student t-tests were performed with $p<0.05$ as the cutoff for statistical significance. Error bars indicate standard deviation.

Declarations

Data availability

The mass spectrometry proteomics raw data have been deposited in the ProteomeXchange Consortium via Proteomics Identification (PRIDE)[52]. The accession number of the proteomics data reported in this paper is PRIDE: PXD026780. Publicly available datasets used in this study are described in the Method

section of Patient database and RNA-seq analyses. All other data generated during this study are available from the authors on request.

Acknowledgements

J.X. was supported by a Canadian Institutes for Health Research Doctoral Research Award. A.U. is supported by research funding from the National Health & Medical Research Council (Australia, APP1163815) and Leukemia & Lymphoma Society, Leukemia Foundation and Snowdome Foundation (6589-20). K.J.H. received support from an Ontario Institute for Cancer Research Investigator Award (#IA-033) and Canadian Institutes for Health Research Project Grant (PJT-156158). K.J.H. and G.W.Y. are partially funded by NIH grant R01 HL137223. Y.L. received support from Canadian Natural Science and Engineering Research Council Discovery Grant (RGPIN-2017-06159), Cancer Research Society Operating Grant (24278), Canada Foundation for Innovation JELF (35544), Ontario Ministry of Economic Development, Job Creation and Trade ORF-RI (35544), and the Marta and Owen Boris Foundation.

Competing Interests

G.W.Y. is co-founder, member of the Board of Directors, on the Scientific Advisory Board, equity holder, and paid consultant for Eclipse BioInnovations. 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.

Author Contributions

L.L. designed and performed experiments, analyzed data and wrote the manuscript. A.V., J.X., and K.J.H. assisted with in vivo xenotransplant and mouse bone marrow processing. N.P.D., G.A., and A.S. performed RNA-seq splicing analysis and protein effects prediction. S.S. and G.W.Y. performed eCLIP–seq bioinformatic analyses and advised on data interpretation. M.D.M provided AML patient samples. Y.L. performed LC-MS and proteomics data analyses. K.J.H. and Y.L. conceived the project, supervised the study, analyzed data, and interpreted results. L.L. and Y.L. wrote the manuscript. K.J.H and Y.L. are co-corresponding authors.

References

1. Döhner, H., Weisdorf, D. J. & Bloomfield, C. D. Acute Myeloid Leukemia. *N. Engl. J. Med.* **373**, 1136–1152 (2015).

2. Bhatia, M., Wang, J. C. Y., Kapp, U., Bonnet, D. & Dick, J. E. Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proc. Natl. Acad. Sci.* **94**, 5320–5325 (1997).

3. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737 (1997).

4. Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645–648 (1994).

5. Bjerkvig, R., Tysnes, B. B., Aboody, K. S., Najbauer, J. & Terzis, A. J. A. The origin of the cancer stem cell: current controversies and new insights. *Nat. Rev. Cancer* **5**, 899–904 (2005).

6. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).

7. Tallman, M. S., Gilliland, D. G. & Rowe, J. M. Drug therapy for acute myeloid leukemia. *Blood* **106**, 1154–1163 (2005).

8. Ng, S. W. K. *et al.* A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437 (2016).

9. Yilmaz, Ö. H. & Morrison, S. J. The PI-3kinase pathway in hematopoietic stem cells and leukemia-initiating cells: a mechanistic difference between normal and cancer stem cells. *Blood Cells. Mol. Dis.* **41**, 73–76 (2008).

10. Wang, E. T. *et al.* Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature* **456**, 470–476 (2008).

11. Black, D. L. Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* **72**, 291–336 (2003).

12. Chen, M. & Manley, J. L. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **10**, 741–754 (2009).

13. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* **33**, 5311–5318 (2014).

14. Adamia, S. *et al.* A genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **20**, 1135–1145 (2014).

15. Crews, L. A. *et al.* RNA Splicing Modulation Selectively Impairs Leukemia Stem Cell Maintenance in Secondary Human AML. *Cell Stem Cell* **19**, 599–612 (2016).

16. Abrahamsson, A. E. *et al.* Glycogen synthase kinase 3 missplicing contributes to leukemia stem cell generation. *Proc. Natl. Acad. Sci.* **106**, 3925–3929 (2009).

17. Tyner, J. W. *et al.* Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).

18. Papaemmanuil, E. *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).

19. The Cancer Genome Atlas Research Network. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).

20. Wang, E. *et al.* Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. *Cancer Cell* **35**, 369-384.e7 (2019).

21. Ge, Y. *et al.* The splicing factor RBM25 controls MYC activity in acute myeloid leukemia. *Nat. Commun.* **10**, 172 (2019).

22. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

23. Shastri, A. *et al.* Antisense STAT3 inhibitor decreases viability of myelodysplastic and leukemic stem cells. *J. Clin. Invest.* **128**, 5479–5488 (2018).

24. Lechman, E. R. *et al.* miR-126 Regulates Distinct Self-Renewal Outcomes in Normal and Malignant Hematopoietic Stem Cells. *Cancer Cell* **29**, 214–228 (2016).

25. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

26. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS ONE* **6**, e21800 (2011).

27. Anande, G. *et al.* RNA Splicing Alterations Induce a Cellular Stress Response Associated with Poor Prognosis in Acute Myeloid Leukemia. *Clin. Cancer Res.* **26**, 3597–3607 (2020).

28. De Maio, A. *et al.* RBM17 Interacts with U2SURP and CHERP to Regulate Expression and Splicing of RNA-Processing Proteins. *Cell Rep.* **25**, 726-736.e7 (2018).

29. Khan, S. N. *et al.* Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. *Leukemia* **27**, 1301–1309 (2013).

30. van Vlerken, L. E. *et al.* EZH2 is required for breast and pancreatic cancer stem cell maintenance and can be used as a functional cancer stem cell reporter. *Stem Cells Transl. Med.* **2**, 43–52 (2013).

31. Lim, K. M., Yeo, W. S. & Chow, V. T. K. Antisense abrogation of DENN expression induces apoptosis of leukemia cells in vitro, causes tumor regression in vivo and alters the transcription of genes involved in apoptosis and the cell cycle. *Int. J. Cancer* **109**, 24–37 (2004).

32. Ji, D. *et al.* Oncogenic heterogeneous nuclear ribonucleoprotein D-like modulates the growth and imatinib response of human chronic myeloid leukemia CD34 + cells via pre-B-cell leukemia homeobox 1.

Oncogene **39**, 443–453 (2020).

33. Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* **21**, 708–718 (2007).

34. Hussain, T. *et al.* Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159**, 597–607 (2014).

35. Rapin, N. *et al.* Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood* **123**, 894–904 (2014).

36. Lallena, M. J., Chalmers, K. J., Llamazares, S., Lamond, A. I. & Valcárcel, J. Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. *Cell* **109**, 285–296 (2002).

37. Al-Ayoubi, A. M., Zheng, H., Liu, Y., Bai, T. & Eblen, S. T. Mitogen-activated protein kinase phosphorylation of splicing factor 45 (SPF45) regulates SPF45 alternative splicing site utilization, proliferation, and cell adhesion. *Mol. Cell. Biol.* **32**, 2880–2893 (2012).

38. Perry, W. L. *et al.* Human splicing factor SPF45 (RBM17) confers broad multidrug resistance to anticancer drugs when overexpressed–a phenotype partially reversed by selective estrogen receptor modulators. *Cancer Res.* **65**, 6593–6600 (2005).

39. Sampath, J. *et al.* Human SPF45, a splicing factor, has limited expression in normal tissues, is overexpressed in many tumors, and can confer a multidrug-resistant phenotype to cells. *Am. J. Pathol.* **163**, 1781–1790 (2003).

40. Lu, Y. *et al.* Alternative Splicing of MBD2 Supports Self-Renewal in Human Pluripotent Stem Cells. *Cell Stem Cell* **15**, 92–101 (2014).

41. Tan, Q. *et al.* Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. *Hum. Mol. Genet.* **25**, 5083–5093 (2016).

42. Sorek, R. The birth of new exons: Mechanisms and evolutionary consequences. *RNA* **13**, 1603–1608 (2007).

43. Sorek, R., Shamir, R. & Ast, G. How prevalent is functional alternative splicing in the human genome? *Trends Genet. TIG* **20**, 68–71 (2004).

44. Tanaka, S. *et al.* Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. *Blood* **120**, 1107–1117 (2012).

45. Sadlish, H. *et al.* Evidence for a functionally relevant rocaglamide binding site on the eIF4A:RNA complex. *ACS Chem. Biol.* **8**, 1519–1527 (2013).

46. Callahan, K. P. *et al.* Flavaglines target primitive leukemia cells and enhance anti-leukemia drug activity. *Leukemia* **28**, 1960–1968 (2014).

47. Wilczynska, A. *et al.* eIF4A2 drives repression of translation at initiation by Ccr4-Not through purine-rich motifs in the 5'UTR. *Genome Biol.* **20**, 262 (2019).

48. Stevens, B. M. *et al.* Characterization and targeting of malignant stem cells in patients with advanced myelodysplastic syndromes. *Nat. Commun.* **9**, 3694 (2018).

49. Majeti, R. *et al.* Dysregulated gene expression networks in human acute myelogenous leukemia stem cells. *Proc. Natl. Acad. Sci.* **106**, 3396–3401 (2009).

50. Raffel, S. *et al.* BCAT1 restricts αKG levels in AML stem cells leading to IDHmut-like DNA hypermethylation. *Nature* **551**, 384–388 (2017).

51. Kwan, V. *et al.* Proteomic Analysis Reveals Autism-Associated Gene *DIXDC1* Regulates Proteins Associated with Mitochondrial Organization and Function. *J. Proteome Res.* **20**, 1052–1062 (2021).

52. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

53. Karolchik, D. *et al.* The UCSC Genome Browser Database. *Nucleic Acids Res.* **31**, 51–54 (2003).

Figures

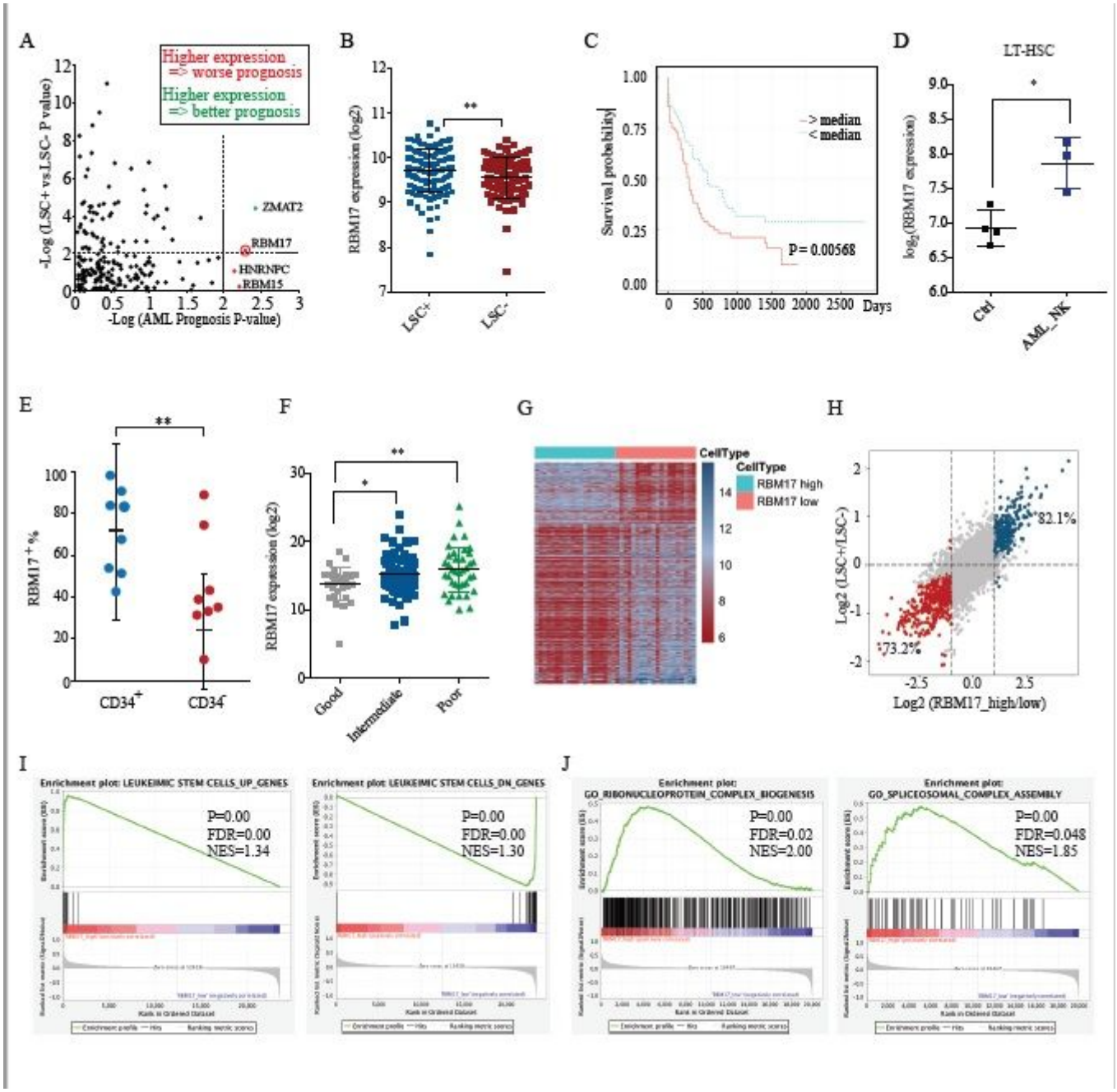

Figure 1

Heightened expression of RBM17 correlates with poor prognosis in human AML patients. (A) Fold change of expression in LSC-enriched vs LSC-depleted subsets from 78 AML patients (y axis) and their correlation with AML patients' overall survival (x axis). (B) RBM17 transcript level in LSC-enriched vs LSC-depleted subsets from 78 AML patient samples. (C) Kaplan Meier curves showing outcomes of AML patients from the TCGA with above vs below median expression of RBM17. (D) Gene expression data from sorted AML bone marrow samples were compared with data from healthy controls (Ctrl) and revealed significantly increased RBM17 expression in AML LT-HSCs (Lin−CD34+CD38−CD90+, AML with normal karyotype (NK)=3) compared with healthy control(Ctrl)=4). (E) Intracellular flow cytometric measurements of RBM17 protein level in the primitive CD34+ subset vs the committed CD34- subset in 8 primary AML samples. (F) Expression of RBM17 in 152 AML specimens and each molecular genetic risk

group from the LAML-TCGA cohort. (G) Gene expression profiles form AML samples with high and low RBM17 were compared and differentially expressed transcripts identified. (H) Expression of differential transcripts identified from (G) in LSC and non-LSC subsets of 78 AML patients. (I, J) Gene set enrichment analysis (GESA) of the gene signature of high-RBM17 AML cases compared with previously published (I) LSC signatures and (J) spliceosomal complex assembly and ribonucleoprotein complex biogenesis pathways. *p<0.05, ** p<0.01.

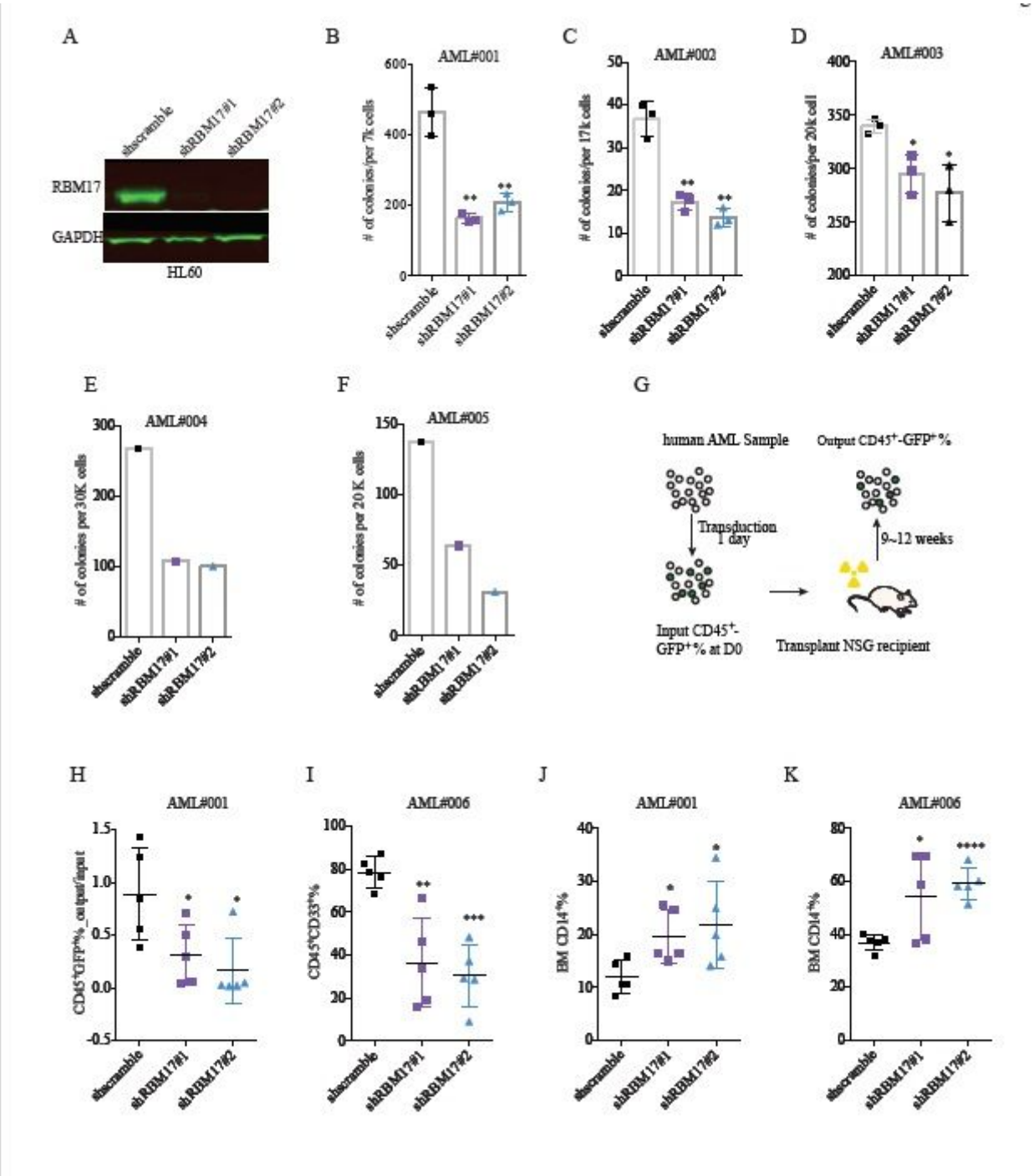

Figure 2

RBM17 is required to support primitive leukemic cell functions. (A) Western blot validation of RBM17 knockdown (KD) in transduced GFP+ HL60 cells. (B-F) Colony formation capacity assessed in 5 AML

samples following RBM17 KD. (G) Xenotransplant experimental design using GFP% output/input analysis to readout primary AML LSC reconstitution capacity after RBM17 KD. (H) AML engraftment after 9 weeks (#001) in whole bone marrow. n=5 each for shscramble, shRBM17#1, and shRBM17#2. Shown is the ratio of the GFP+ cell percentage in the human cell population post-transplant to the initial pre-transplant GFP+ cell percentage. (I) AML engraftment (#006) after 12 weeks in bone marrow. n=5 for each group. Shown is the percentage of human (CD45+) myeloid cells (CD33+) found in bone marrow. (J-K) Immunophenotyping of myeloid differentiation in /post-transplant grafts from (H) and (I). *p<0.05, **p<0.01, ***p<0.001.

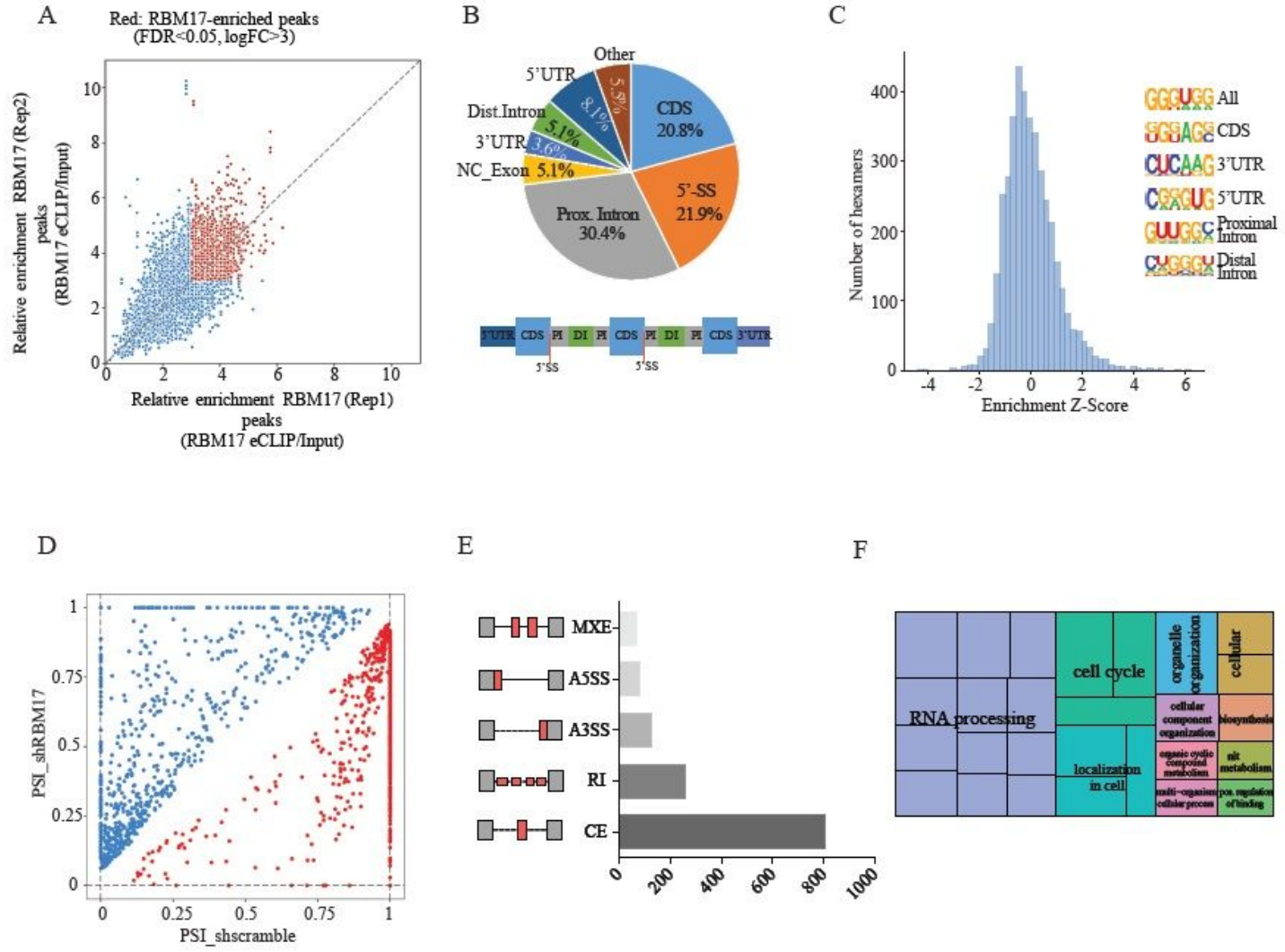

Figure 3

Identification of RNA targets of and alternative splicing variants enforced by RBM17 in AML. (A) eCLIP-seq analysis of RBM17 binding sites in K562 cells. Input-normalized peak signals are shown as log2 fold change. Red points indicate eCLIP-enriched RBM17 peaks (FDR<0.05 and log2 (FC)>3) in biological replicates. (B) Significantly enriched reproducible binding sites or RBM17 across the transcriptome. (C) Hexamer enrichment for RBM17 binding peaks in K562 cells based on eCLIP-seq. (D) RNA-seq analysis of

K562 cells showed AS events impacted by RBM17 knockdown. (E) Quantification of types of AS events affected by RBM17, as revealed by analysis of RNA-seq data. (F) Tree plot showing GO pathway associations of genes with AS events downstream of RBM17.

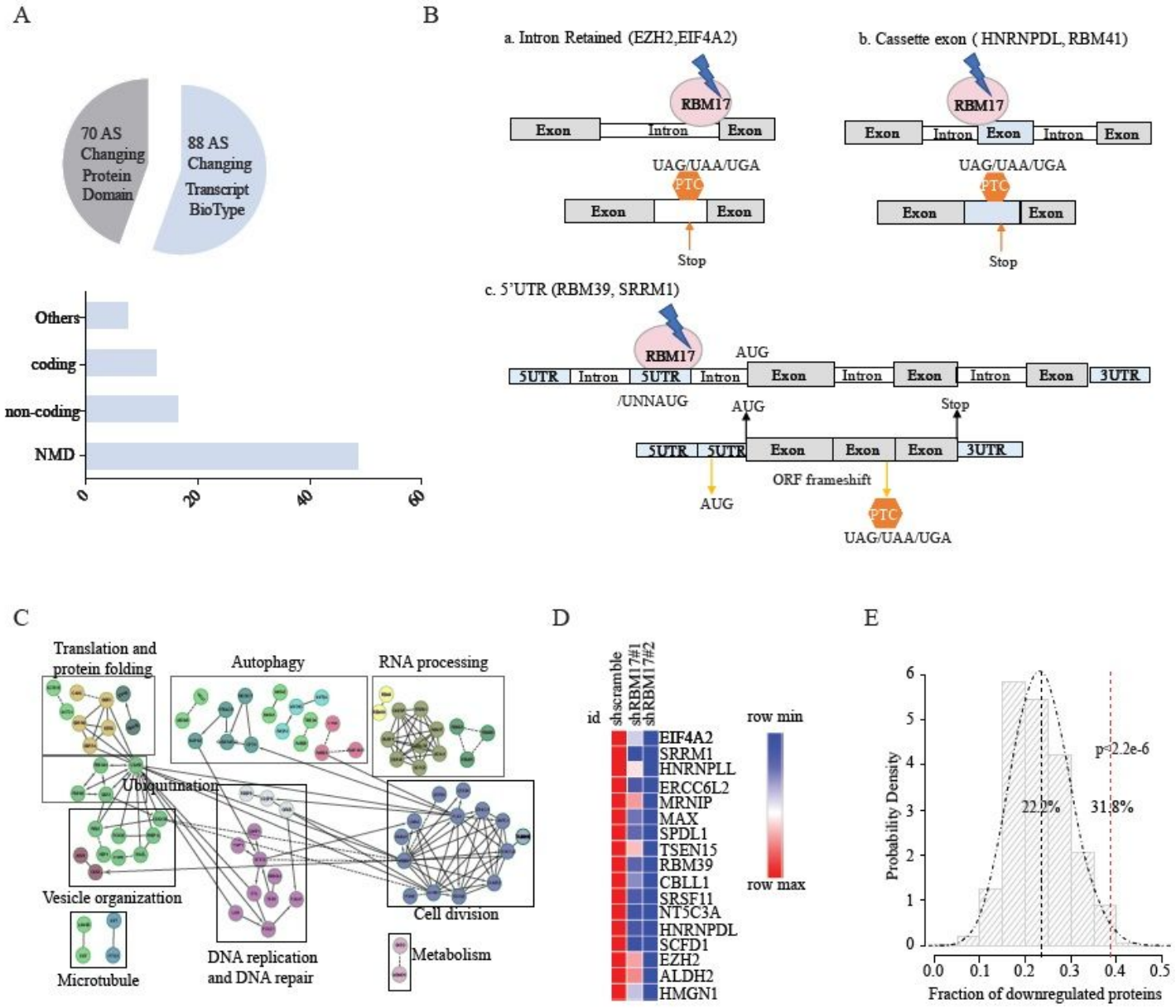

Figure 4

RBM17 knockdown leads to the production of NMD sensitive transcripts. (A) Pie chart distribution of protein domain and biotype change prediction. Bottom bar plot indicates the distribution of alternative splicing events predicted to lead to biotype changes in shRBM17 groups. (B) Depiction of (a) RBM17 binding to intronic regions of EZH2, EIF4A2 and RBM25 and resultant promotion of their retention and PTC introduction following knockdown of RBM17; (b) RBM17 binding to the cassette exon of HNRNPDL and RBM41 leading to cassette exon inclusion and introduction of PTCs post-RBM17 knockdown; (c) RBM17 binding to exons belonging to 5'UTRs of RBM39 and SRRM1 and subsequent inclusion of these exons that contain alternative start codons upon RBM17 knockdown, inducing ORF frameshifts and

PTCs. (C) Cytoscape network analysis of proteins significantly deregulated by RBM17 knockdown in K562 cells. (D) Heat map of protein expression fold change of 20 NMD-sensitive transcripts with and without RBM17 knockdown. (E) Bootstrapping analysis of 44 proteins from 8497 total proteins identified from the RBM17 knockdown proteome.

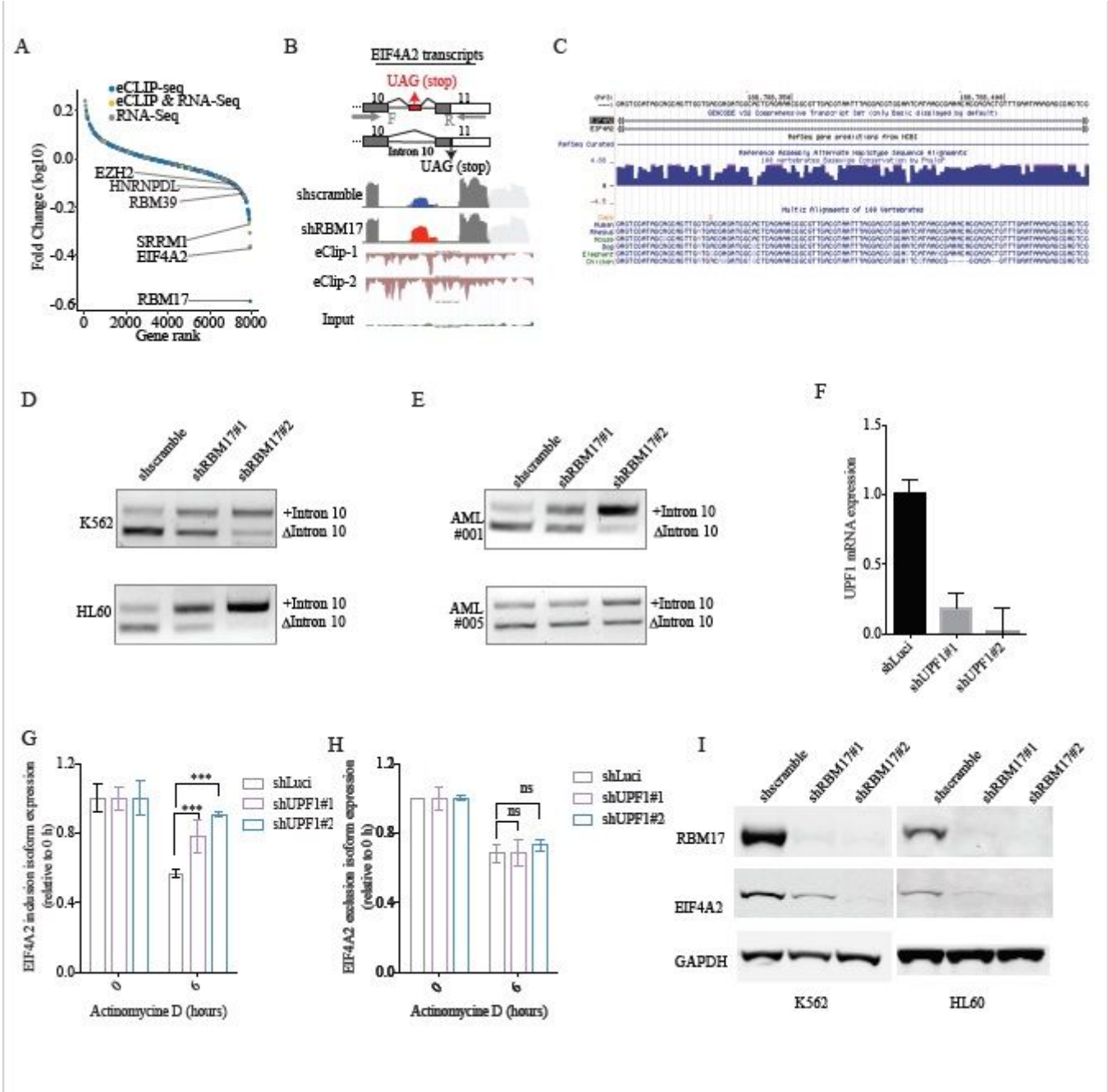

Figure 5

RBM17 suppresses EIF4A2 poison exon inclusion. (A) Ranked protein expression change curve 5 days after transduction with shRNAs against RBM17 in K562 cells. Protein names that also have significant RBM17 binding with their transcripts (eCLIP-seq) and have significant alternative splicing changes after RBM17 knockdown (at day 5 by RNA-seq) are displayed. (B) Upper: Diagram of EIF4A2 splicing variants and the primers for RT-PCR detection of cryptic exon inclusion. Bottom: IGV plot of RNA-seq data

illustrating the EIF4A2 cryptic exon promoted by RBM17 knockdown and anti-RBM17 eCLIP-seq tracks. (C) The conservation track obtained from the UCSC Genome Browser53 shows high conservation of EIF4A2 “cryptic exon” with PTC and flanking intronic sequences. (D-E) RT-PCR validation of EIF4A2 variants using RNA extracted from K562, HL60 and primary AML cells with and without RBM17 knockdown. (F) qPCR quantification of UPF1 knockdown in K562 cells. (G) EIF4A2 intron 10-included isoform in K562 cells with and without UPF1 knockdown and actinomycin D treatment. (H) EIF4A2 intron 10-excluded isoform in K562 cells with and without UPF1 knockdown and actinomycine D treatment. (I) Western blot detection of EIF4A2 after RBM17 knockdown in K562 and HL60 cell lines.

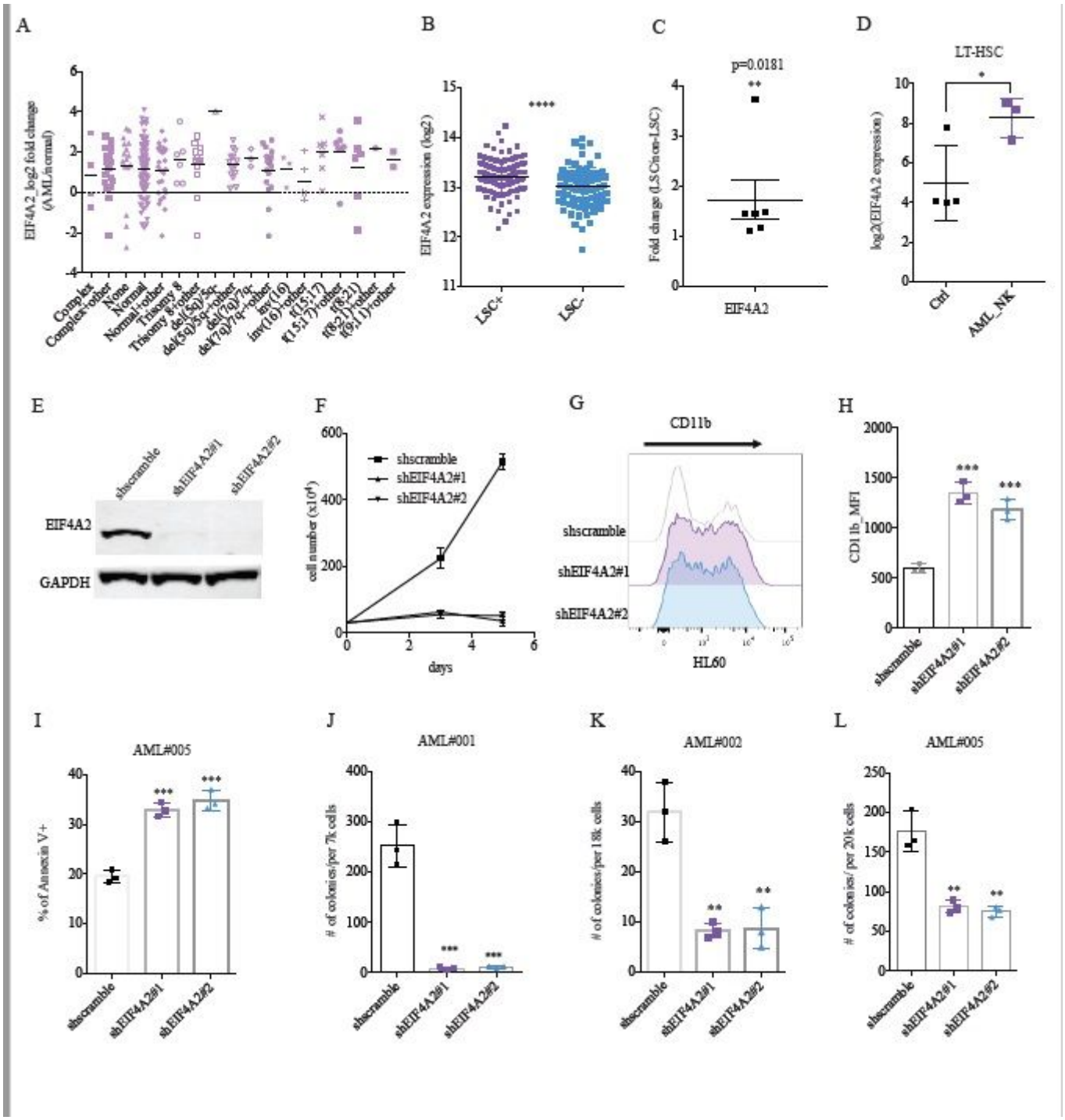

Figure 6

EIF4A2 is required for sustaining primitive leukemic cell fitness. (A) EIF4A2 mRNA expression in different subtypes of AML as compared to their nearest normal counterpart, on the basis of data from The Cancer Genome Atlas database. Data are presented as mean log2 fold change. (B) EIF4A2 transcript level in LSC-enriched vs LSC-depleted subsets from 78 AML patient samples. (C) EIF4A2 protein level as assessed proteomics analysis in functionally defined LSC and non-LSC populations from 6 AML patient samples. (D) Gene expression data from sorted AML bone marrow samples were compared with data from healthy controls (Ctrl) and revealed significantly increased RBM17 expression in AML LT-HSCs (Lin-CD34+CD38-CD90+, AML with normal karyotype (NK)=3) compared with healthy control(Ctrl=4). (E) WB validation of EIF4A2 knockdown in GFP+ transduced HL60 cells. (F-H) Assessment of the effects of EIF4A2 knockdown on cell growth (F) and myeloid differentiation (G-H) as compared to control shRNA in HL60 cells. (I) Flow cytometric evaluation of apoptosis following EIF4A2 knockdown in primary AML cells. (J-L) Colony formation capacity following EIF4A2 knockdown in 3 AML samples.

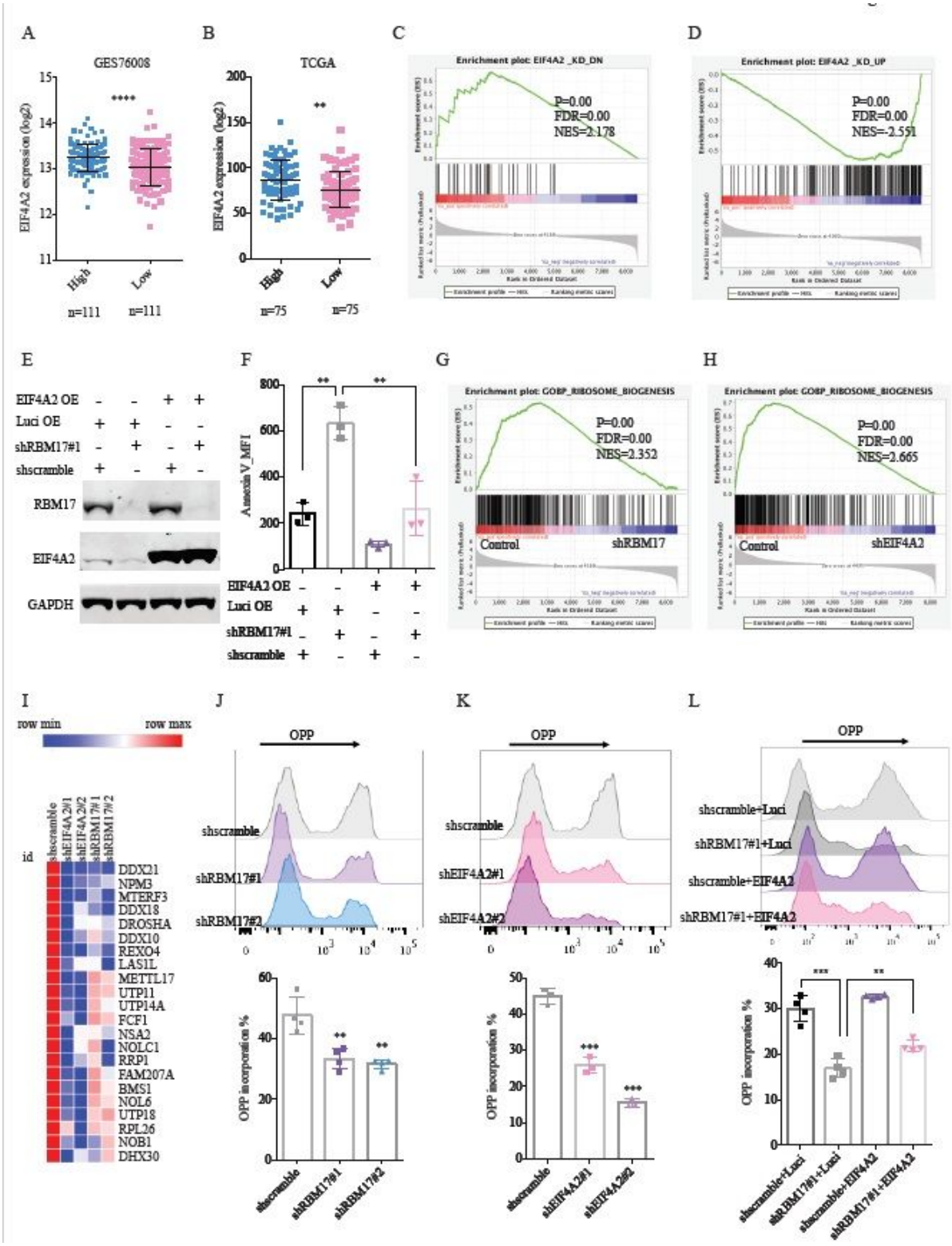

Figure 7

EIF4A2 overexpression partially rescues the RBM17 knockdown phenotype in AML cells. (A-B) Correlation analysis between RBM17 and EIF4A2 mRNA levels using the above microarray data about functionally defined 139 LSC-enriched and 83 non-LSC populations (n=222) (A) and TCGA data (B). (C-D) GSEA enrichment plots showing RBM17 knockdown in K562 cells leads to downregulation of the EIF4A2_KD_DN gene set and upregulation of the EIF4A2_KD_UP gene set. (E) WB images showing

expression of RBM17 and EIF4A2 in HL60 cells engineered to co-express luciferase or EIF4A2 with and without RBM17 knockdown. (F) Flow cytometry analysis of AnnexinV in HL60 cells on day 4 following co-expression of luciferase or EIF4A2 and knockdown of RBM17. Mean ± SD, n = 3. (G-H) GSEA enrichment plots showing EIF4A2 and RBM17 knockdown in K562 cells leads to downregulation of the GO ribosome biogenesis gene set. (I) Heat map showing downregulated proteins from the ribosome biogenesis gene set induced by EIF4A2 and RBM17 knockdown. (J-L) Representative histogram and quantification of flow cytometric detection of op-puro incorporation in HL60 cells on day5 following (J) EIF4A2 knockdown (K) RBM17 knockdown or (L) simultaneous expression of luciferase or EIF4A2 with or without RBM17 knockdown. Mean ± SD, n = 3.

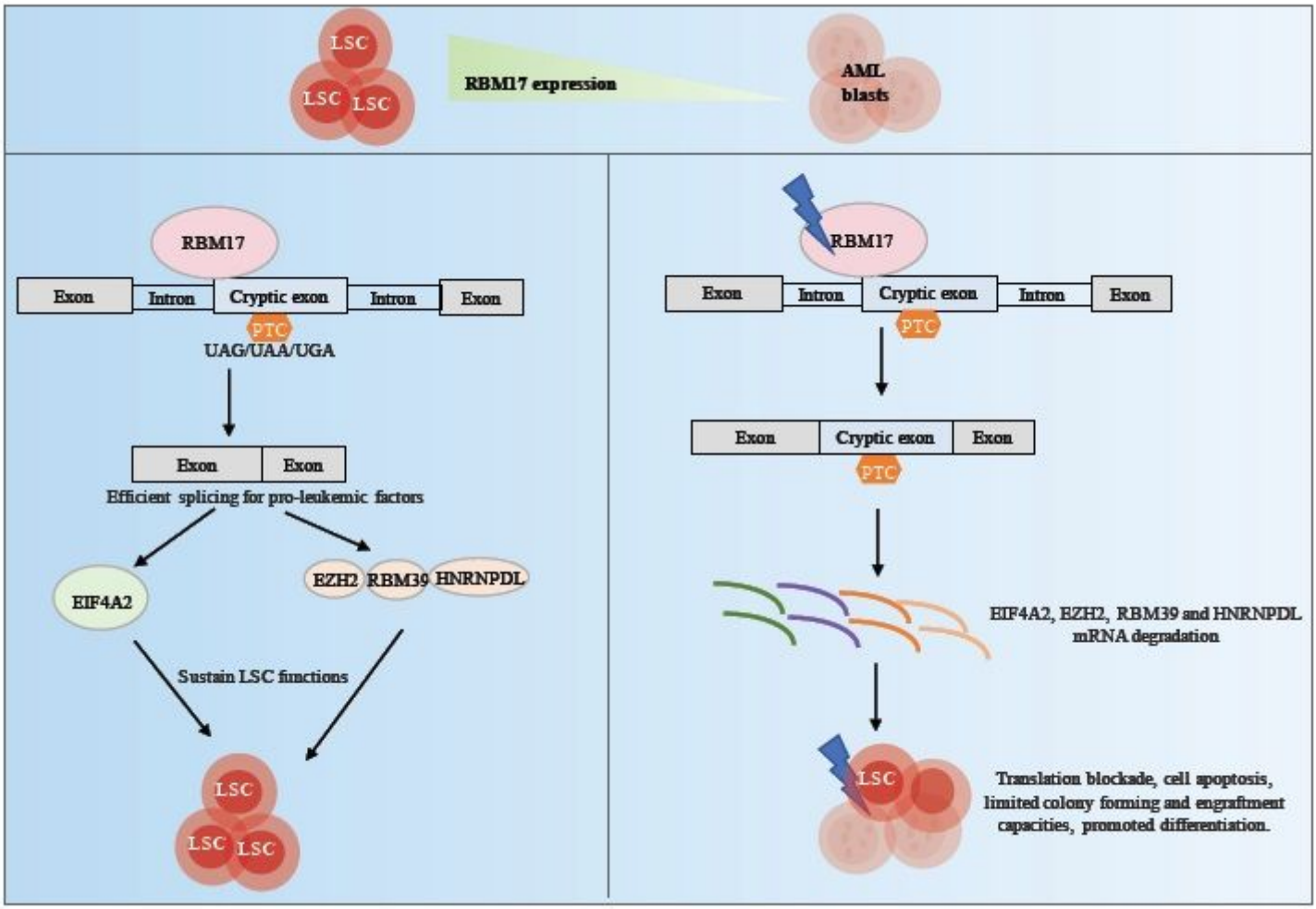

Figure 8

Schematic model depicting the role of RBM17 in primitive AML cells. RBM17 is abnormally higher expressed in the most primitive cell fractions of AML compared to AML blasts, which contributes to efficient splicing of many pro-leukemic factors EZH2, RBM39 and HNRNPDL, along with EIF4A2 that functions in translation control, to sustain LSC functions. Knockdown of RBM17 promotes inclusions of cryptic exons or introns into mRNAs of these pro-leukemic factors, leading to their mRNA degradations due to NMD and consequently resulting in translation blockade, cell apoptosis, limited colony-forming and engraftment capacities, and promoted differentiation in primitive AML cells.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1GOmRNAsplicingfactorsTCGAAMLprognosisLSC.xlsx
- TableS2AMLsampleinformation.xlsx
- TableS3RBM17ecliptargts.xlsx
- TableS4RNAseqASRBM17KDK562.xlsx
- TableS5shRBM17BiotypeandProteinDomainprediction.xlsx
- TableS6K562RBM17KDDay5TMT640ug36cutNormalized.xlsx
- TableS717NMDGO.xlsx
- TableS8K562EIF4A2KDDay5TMT640ug36cutNormalized.xlsx
- LinaLiuRBM17papersupplementaryNC.pdf