

# METTL3-Mediated m<sup>6</sup>A Modification Controls Splicing Factor Abundance and Contributes to Aggressive CLL



Yiming Wu<sup>1</sup>, Meiling Jin<sup>1</sup>, Mike Fernandez<sup>1</sup>, Kevyn L. Hart<sup>1</sup>, Aijun Liao<sup>1</sup>, Xinzhou Ge<sup>2,3</sup>, Stacey M. Fernandes<sup>4</sup>,  
Tinisha McDonald<sup>5,6</sup>, Zhenhua Chen<sup>1</sup>, Daniel Röth<sup>7</sup>, Lucy Y. Ghoda<sup>5,6</sup>, Guido Marcucci<sup>5,6,8</sup>, Markus Kalkum<sup>7</sup>,  
Raju K. Pillai<sup>9</sup>, Alexey V. Danilov<sup>8,10</sup>, Jingyi Jessica Li<sup>2,3</sup>, Jianjun Chen<sup>1</sup>, Jennifer R. Brown<sup>4</sup>, Steven T. Rosen<sup>8,10</sup>,  
Tanya Siddiqi<sup>8,10</sup>, and Lili Wang<sup>1,10</sup>

**ABSTRACT**

RNA splicing dysregulation underlies the onset and progression of cancers. In chronic lymphocytic leukemia (CLL), spliceosome mutations leading to aberrant splicing occur in ~20% of patients. However, the mechanism for splicing defects in spliceosome-unmutated CLL cases remains elusive. Through an integrative transcriptomic and proteomic analysis, we discover that proteins involved in RNA splicing are posttranscriptionally upregulated in CLL cells, resulting in splicing dysregulation. The abundance of splicing complexes is an independent risk factor for poor prognosis. Moreover, increased splicing factor expression is highly correlated with the abundance of METTL3, an RNA methyltransferase that deposits N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) on mRNA. METTL3 is essential for cell growth *in vitro* and *in vivo* and controls splicing factor protein expression in a methyltransferase-dependent manner through m<sup>6</sup>A modification-mediated ribosome recycling and decoding. Our results uncover METTL3-mediated m<sup>6</sup>A modification as a novel regulatory axis in driving splicing dysregulation and contributing to aggressive CLL.

**SIGNIFICANCE:** METTL3 controls widespread splicing factor abundance via translational control of m<sup>6</sup>A-modified mRNA, contributes to RNA splicing dysregulation and disease progression in CLL, and serves as a potential therapeutic target in aggressive CLL.

See related commentary by Janin and Esteller, p. 176.

**INTRODUCTION**

Alternative splicing is a fundamental process that generates functional diversity in the transcriptome (1, 2). Advancement of RNA sequencing (RNA-seq) technology and computational algorithms for quantitative splicing analysis coupled with clinical sample annotations has not only identified RNA splicing dysregulation as a hallmark of cancer but also revealed that it underlies transformation, disease progression, and resistance to therapy (3–10). Understanding the molecular mechanisms of splicing dysregulation is key to dissecting the origins and progression of cancer and designing better therapeutic approaches.

Chronic lymphocytic leukemia (CLL), the most common form of adult mature B-cell leukemia in North America,

displays transcriptome-wide RNA splicing defects (11, 12). Large-scale cancer genome sequencing has identified recurrent mutations of spliceosomal components (e.g. *SF3B1*) in CLL (13–16). These mutations alter splicing specificity and activate pathways that are essential for disease onset and progression (10, 16, 17). As exemplified in murine models, coexpression of *Sf3b1* mutations with *Atm* deletion in murine B cells triggers the onset of CLL (8, 10), highlighting splicing dysregulation as a central underlying pathway in CLL. However, mutations of splicing factors are present in only about 20% of CLL samples and cannot fully explain the general splicing defects observed in this disease (18). This led us to hypothesize that transcriptional and posttranscriptional regulation of RNA splicing contributes to leukemogenesis.

N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) is the most abundant and conserved internal modification of mRNA and regulates many aspects of RNA metabolism, including RNA splicing (19–24). As a key player in transcriptional and posttranscriptional control of gene expression, m<sup>6</sup>A modification is dynamically regulated by writers (METTL3/METTL14/WTAP, METTL16), erasers (ALKBH5, FTO), and readers (YTHDF1/2/3, YTHDC1/2, IGFBP1/2/3, HNRNPC, HNRNPG; refs. 22, 25–29). Emerging evidence reveals m<sup>6</sup>A modification is required for normal hematopoiesis, and alteration of m<sup>6</sup>A-modified transcripts leads to hematologic malignancies (30–32). However, the exact roles of m<sup>6</sup>A modification and its critical regulators in RNA splicing dysregulation in CLL are largely unknown. In the current study, we set out to examine the nongenetic regulation of RNA splicing dysregulation in CLL by METTL3 and m<sup>6</sup>A modification.

**RESULTS****Posttranscriptional Upregulation of RNA Splicing Pathway in CLL**

To study the spliceosome mutation-independent splicing defects in CLL, we analyzed RNA-seq data of B cells derived from 6 healthy donors and 36 CLL patients (13 with and 23

<sup>1</sup>Department of Systems Biology, Beckman Research Institute, City of Hope National Comprehensive Cancer Center, Monrovia, California.

<sup>2</sup>Department of Statistics, University of California, Los Angeles, California.

<sup>3</sup>Department of Computational Medicine, University of California, Los Angeles, California.

<sup>4</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts.

<sup>5</sup>The Hematopoietic Tissue Biorepository, City of Hope National Comprehensive Cancer Center, Duarte, California.

<sup>6</sup>Department of Hematological Malignancies Translational Sciences, Beckman Research Institute, City of Hope Comprehensive Cancer Center, Duarte, California.

<sup>7</sup>Department of Molecular Imaging and Therapy, Diabetes and Metabolism Research Institute, Beckman Research Institute, City of Hope, Duarte, California.

<sup>8</sup>Department of Hematology & Hematopoietic Cell Transplantation, City of Hope Comprehensive Cancer Center, Duarte, California.

<sup>9</sup>Department of Pathology, City of Hope National Comprehensive Cancer Center, Duarte, California.

<sup>10</sup>Toni Stephenson Lymphoma Center, Beckman Research Institute, City of Hope Comprehensive Cancer Center, Duarte, California.

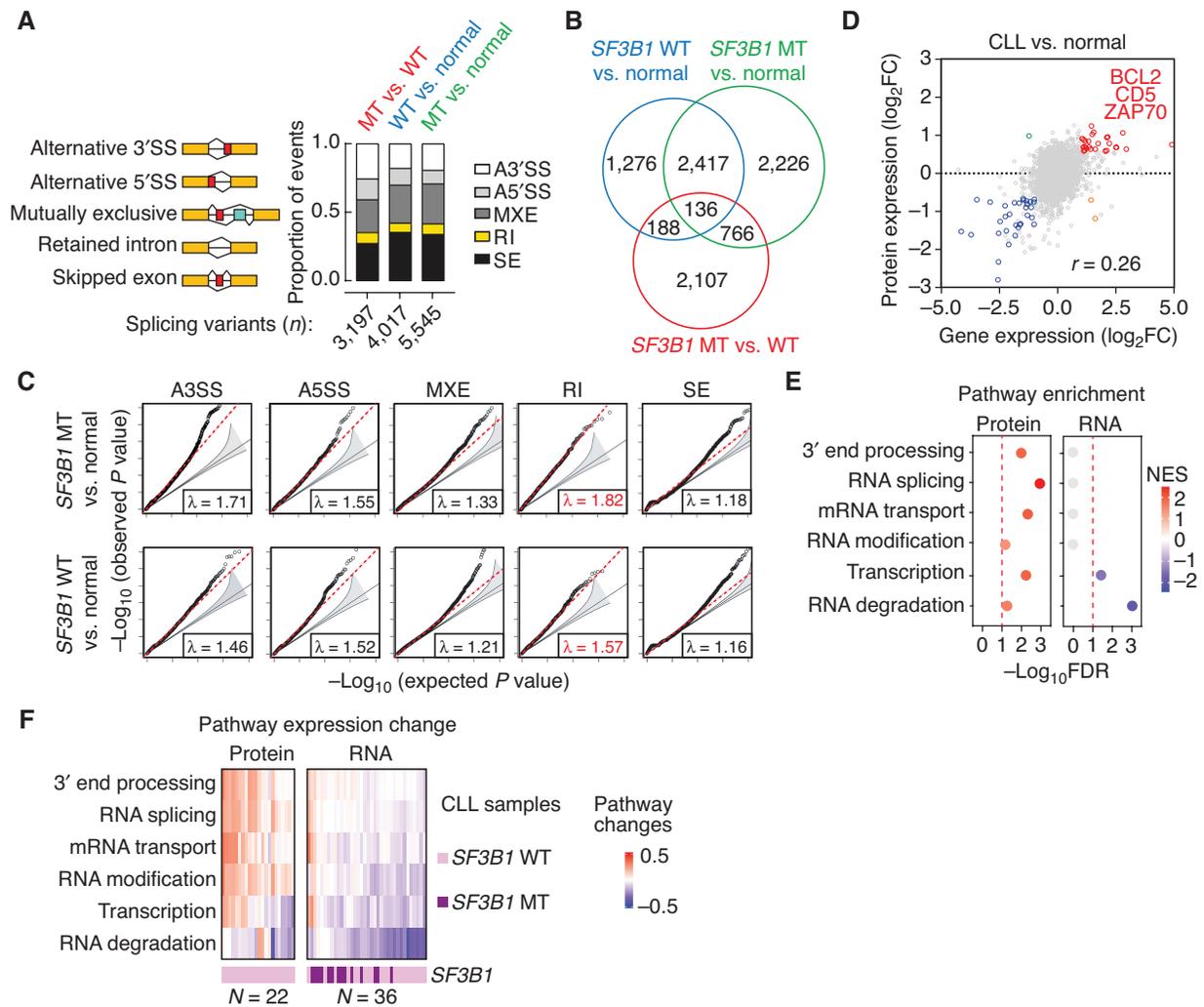
Y. Wu and M. Jin contributed equally to this article.

**Corresponding Author:** Lili Wang, Systems Biology, City of Hope National Cancer Center, 1218 S Fifth Avenue, Monrovia, CA 91016. Phone: 626-218-7016; E-mail: lilwang@coh.org

Blood Cancer Discov 2023;4:228–45

doi: 10.1158/2643-3230.BCD-22-0156

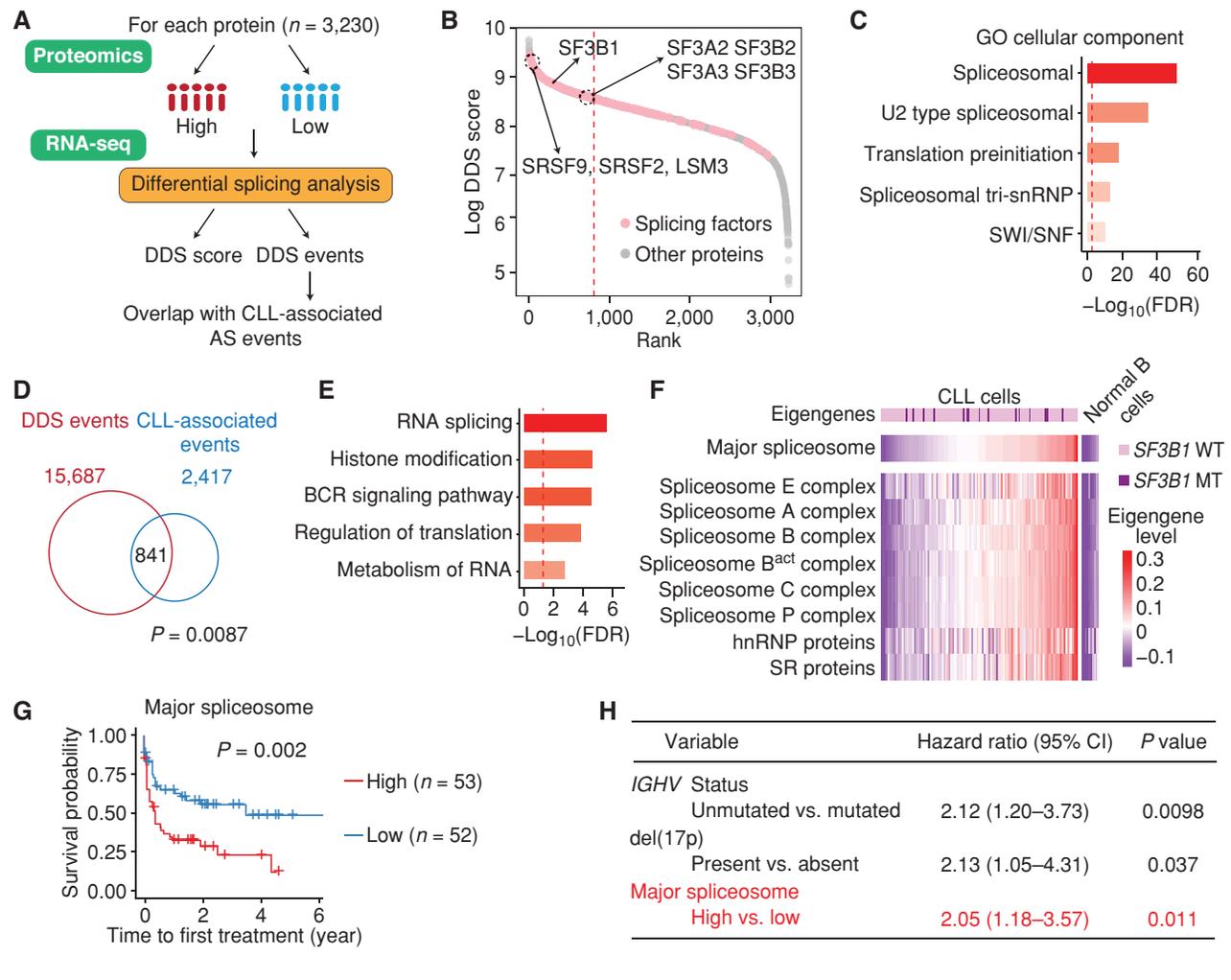
©2023 American Association for Cancer Research



**Figure 1.** Integrated transcriptomic and proteomic analysis identifies posttranscriptional upregulation of splicing factors in CLL. **A**, Left: schema of five alternative RNA splicing categories. Right: the proportion of five alternative splicing categories within the significant alternative splicing events in each comparison among normal B-cell samples (normal,  $n = 6$ ) as well as CLL samples either with [Mutant (MT),  $n = 13$ ] or without [Wild-type (WT),  $n = 23$ ]. **B**, Overlapping of significant alternative splicing events among three comparisons. Proportion test: (blue, green) vs. (red, blue),  $P < 0.0001$ ; (blue, green) vs. (red, green),  $P < 0.0001$ . **C**, Q-Q plots of observed P values against the expected P values under uniform distribution of five alternative splicing categories. Red lines indicate the least-squares linear fit to the lower 95th percentile of points with slope  $\lambda$ . Gray-shaded areas represent 95% confidence intervals for the expected distribution. **D**, Correlative scatter plot of gene and protein expression changes between 22 SF3B1 wild-type CLL and 10 normal B cell samples based on the transcriptomic and proteomic data. Color-coded dots indicate genes with significant changes in CLL samples at both RNA and protein levels (RNA level:  $\log_2$  fold change (FC)  $> 1$ , false discovery rate (FDR)  $< 0.05$ ; protein level: FC  $> 1.5$ , FDR  $< 0.05$ ). **E**, Gene set enrichment analysis of differentially expressed genes and proteins between CLL and normal B cells in RNA metabolism-related pathways. Normalized enrichment score (NES) is color coded. Red dashed line as a cutoff for significance, FDR  $< 0.1$ , permutation test. **F**, Median gene and protein expression changes related to RNA metabolism pathways between CLL and normal B cells samples.

without SF3B1 mutations; Supplementary Table S1) using a bioinformatic pipeline that integrates transcript isoform quantification algorithms StringTie (33), LeafCutter (34), and rMATS (35) to maximally improve the power of detection of alternative splicing events, including *de novo* junctions. When comparing CLL with normal B cells, we identified that 5,545 and 4,017 alternative splicing events were associated with or without SF3B1 mutations, respectively, with 2,553 splicing events shared (CLL-associated events; Fig. 1A and B). When comparing CLL cells with or without SF3B1 mutations, 3,197 splicing events were differentially spliced, with 136 of these among the CLL-associated events (Fig. 1B).

The proportion of shared splice variants between normal and CLL with or without SF3B1 mutations (Fig. 1B, blue and green) is greater than that of SF3B1 mutation-specific events shared with either group (Fig. 1B, red, proportion test,  $P < 0.0001$ ). Of note, splice variants associated with CLL had pervasive changes in intron retention (Fig. 1C), irrespective of SF3B1 mutations. As previously reported (17), SF3B1-mutant samples had cryptic 3' splicing site selection through aberrant branch point usage (Supplementary Fig. S1A and S1B). Taken together, our results suggest that, alongside spliceosomal mutations, other shared regulators impact the alternative RNA splicing in CLL.



**Figure 2.** Widespread upregulation of splicing factors contributes to RNA splicing dysregulation and poor clinical outcome in CLL. **A**, Schema of splicing factor dosage-dependent alternative splicing (DDS) analysis using 19 *SF3B1* wild-type (WT) CLL samples with both RNA-seq and proteomics data. DDS defined as differential splicing events number normalized to expression difference between highly expressed group vs. lowly expressed group. **B**, Ranking of proteomics detected proteins according to DDS score. Splicing-related proteins are pink color coded and the upper quartile marked as a red dashed line. **C**, Metascape analysis of proteins with upper quartile DDS scores, red dashed line: significance cutoff,  $\text{FDR} < 0.05$ , hypergeometric test. **D**, Venn plot between DDS events derived from the top 10 splicing factors and CLL-associated alternative splicing events. The top 10 splicing factors are SMNDC1, PRPF31, SRSF9, DDX46, PPIL1, USP39, LSM3, DDX41, HNRNPK, and CRNKL1. Permutation test,  $P = 0.0087$ . **E**, Metascape analysis of 841 splice variants shared between the top 10 splicing factors associated DDS events and CLL-associated events. **F**, Heat map of the eigengene level for each spliceosome complex, SR proteins, and hnRNP proteins across 10 normal B cells, 15 *SF3B1* MT CLL, and 98 *SF3B1* WT CLL samples in the integrated data set. Overall major spliceosome complex was summarized in the top row. **G**, Kaplan-Meier plot of TTF with the abundance of spliceosome complex, Log-rank test. Groups are divided based on median expression. **H**, Cox multivariable regression model of TTF for major spliceosome complex, *IGHV* mutation status, and *del(17p)*.

To discover regulators for RNA splicing, we performed transcriptomic and quantitative tandem mass tag (TMT) proteomic analyses using B cells derived from healthy donors ( $n = 10$ ) and untreated CLL patients ( $n = 22$ ; Fig. 1D; Supplementary Fig. S2A–S2C). 3,230 proteins were consistently detected across all samples. 2,912 genes and 328 proteins were differentially expressed between normal and CLL B cells based on RNA-seq and proteomics data, respectively (Supplementary Table S2). There is a general correlation between mRNA and protein expression, and as expected, CLL tumor-specific genes CD5 and ZAP70 were upregulated at both the RNA and protein levels (Fig. 1D). Interestingly, gene set enrichment analysis (GSEA) revealed that the RNA splicing process was highly upregulated at the protein level but not at

the RNA level (Fig. 1E and F). This observation was further validated in two independent proteomic, as well as transcriptomic, data sets (refs. 17, 36–38; Supplementary Fig. S2D and S2E), highlighting that the RNA splicing pathway is likely upregulated posttranscriptionally in CLL.

### Splicing Factor Upregulation Results in Splicing Dysregulation and Is an Independent Risk Factor in CLL

To determine whether the abundance of splicing factors impacts RNA splicing, we designed a dosage-dependent splicing (DDS) score system to assess protein dosage-dependent RNA splicing using matched proteomic and RNA-seq data derived from 19 CLL samples (Fig. 2A). Specifically, for each

Downloaded from <http://aacrjournals.org/bloodcancerdiscovery/article-pdf/4/3/228/3327865/228.pdf> by Harvard Medical School user on 05 May 2023

protein, we first identified two subsets of CLL samples, high (top 5 samples) and low (bottom 5 samples), based on its protein expression. Then, we systematically quantified alternative splicing events associated with the protein expression using matched RNA-seq data (Fig. 2A). Last, we assigned to each protein a DDS score, which is the number of splicing events the protein affected normalized to the protein-level difference between the high and low groups, and ranked all proteins by DDS score (Fig. 2B). A high DDS score indicates more splicing events corresponding to protein-level change. Proteins within the upper quartile of the DDS score were significantly enriched in the spliceosome complex, especially U2-type and tri-snRNP spliceosome (Fig. 2C). Upon overlapping DDS events derived from the top 10 splicing factors with 2,417 CLL-associated and *SF3B1* mutation-independent splicing events (Fig. 1B; Supplementary Fig. S2F), we observed more than 35% overlap (Fig 2D, permutation test,  $P = 0.0087$ ), with overlapped events significantly enriched for CLL-related pathways (Fig. 2E). These results suggest that upregulation of splicing factors indeed contributes to RNA splicing dysregulation in CLL by mechanisms yet to be elucidated.

To explore the impacts of splicing factor abundance on CLL pathogenesis, we examined the associations of spliceosome complexes and RNA-binding proteins [serine and arginine-rich (SR) and heterogeneous nuclear ribonucleoprotein (hnRNP) proteins] with clinical outcomes through a combined proteomic data set, which includes 123 samples (10 normal B cell and 113 CLL samples, 2,703 proteins) with 91 samples from a public data set (36). Because splicing factor abundance is highly correlated, we utilized “eigengenes” (39) to summarize the expression dynamics within spliceosome complexes (Fig. 2F). Eigengene is defined as the first principal component of the expression matrix, maximally allowing the diversity among different splicing factors and making it possible to calculate the associations with other features. The abundance of overall spliceosome complex was predictive of a shorter time to first treatment (TTFT) and retained significance when *IGHV* mutation status and 17p deletion were added (Fig. 2G and H). In particular, the abundance of spliceosome complexes A, B, B<sup>act</sup>, C, P, as well as hnRNP proteins was all significantly associated with TTFT (Supplementary Fig. S3A). Moreover, a higher abundance of hnRNP proteins and SR proteins was also associated with inferior overall survival (OS; Supplementary Fig. S3B). Similar results were observed when only considering *SF3B1* wild-type (WT) CLL samples (Supplementary Fig. S3C and S3D). These significant associations highlight that the abundance of splicing factors affects the progression of CLL.

### METTL3 Potentially Regulates Splicing Factors in CLL

To decipher the mechanism underlying the widespread upregulation of splicing factors, we performed a correlational analysis between splicing factors and nonsplicing factors based on protein expression. As for each nonsplicing factor, we calculated the Pearson correlation coefficient for every detected splicing factor and then ranked these nonsplicing factors according to the number of correlation edges. METTL3, as one of the proteins with the most correlated edges with the detected splicing factors (111/143), had a

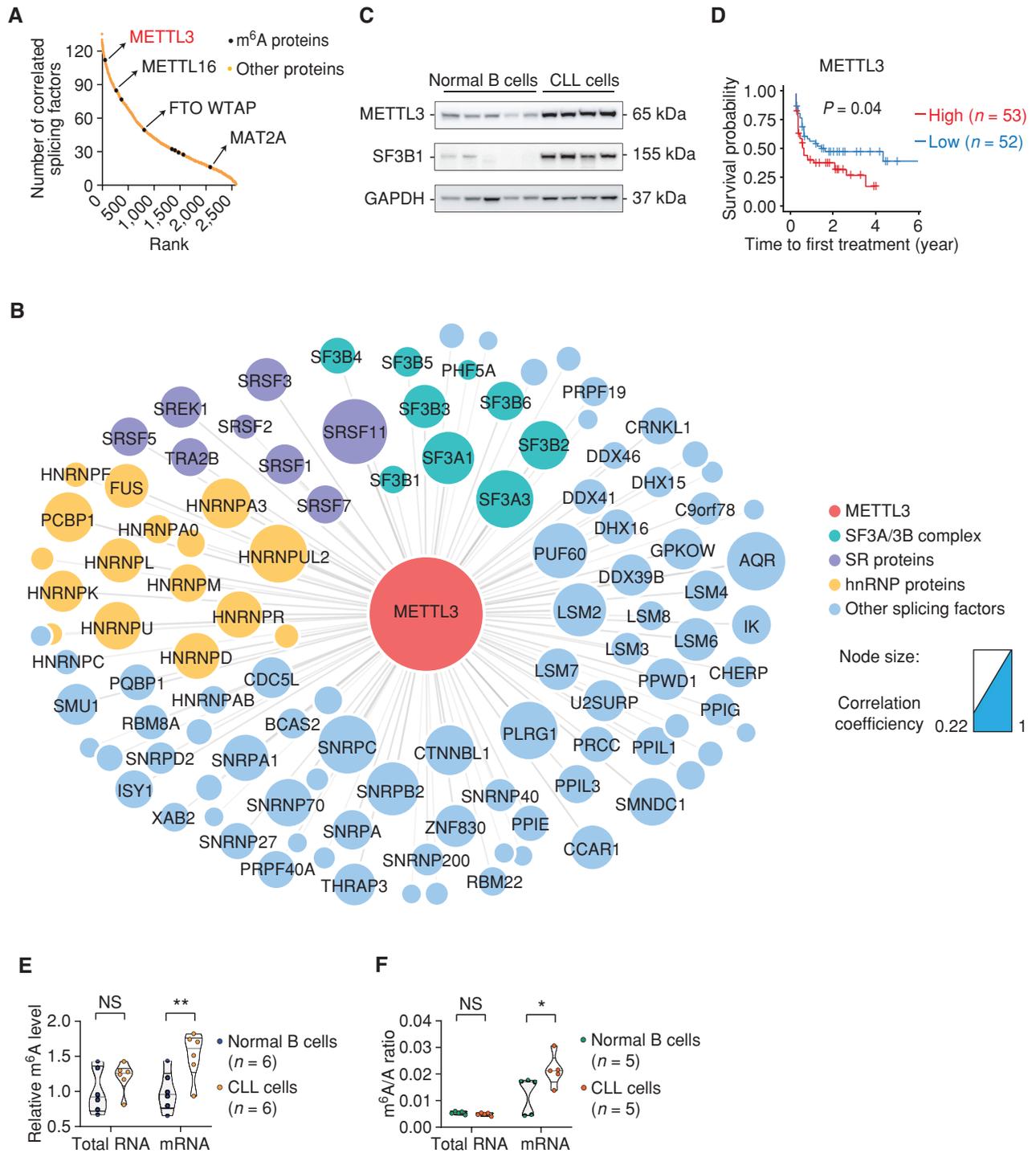
77.6% significant positive correlation (Fig. 3A). Importantly, these splicing factors were involved in splicing complexes with clinical associations (Fig. 3B). The other known RNA methyltransferase, METTL16, also showed a 58.7% positive correlation with splicing factors. These results imply that m<sup>6</sup>A writer proteins may serve as the posttranscriptional regulators for splicing factor abundance in CLL.

To explore the role of *METTL3* and *METTL16* in the regulation of splicing factor abundance, we first validated our proteomics results as well as computational analysis by quantifying protein expression of METTL3 and METTL16 as well as splicing factors (SF3B1, SF3A1, SF3A2, and SF3A3) using independent cohorts of samples with immunoblot (Fig. 3C; Supplementary Fig. S4A). Indeed, CLL cells exhibited upregulation of METTL3 and METTL16, which were positively correlated with spliceosome component expression. *METTL3* and *METTL16* appear to be regulated transcriptionally as they were consistently upregulated at both RNA and protein levels in CLL (Supplementary Fig. S4B). However, we observed no apparent RNA expression changes of splicing factors (*SF3A1*, *SF3A2*, and *SF3B1*), corroborated with results from RNA-seq analysis.

To assess the impacts of *METTL3* and *METTL16* on RNA splicing dysregulation in CLL, we identified 2,010 and 1,927 METTL3- and METTL16-associated DDS events using matched CLL RNA-seq and proteomics data. It appears that these DDS events are enriched for different cellular processes (Supplementary Fig. S4C). Consistent with this, only 56 events were shared, suggesting that *METTL3* and *METTL16* affect RNA splicing on different substrates (Supplementary Fig. S4D). Through the integration of CLL proteomic data with clinical annotation, we found that higher METTL3 but not METTL16 protein expression was associated with a shorter time to first therapy (log-rank test,  $P = 0.04$ ; Fig. 3D; Supplementary Fig. S4E), suggesting METTL3 plays a role in driving aggressive CLL.

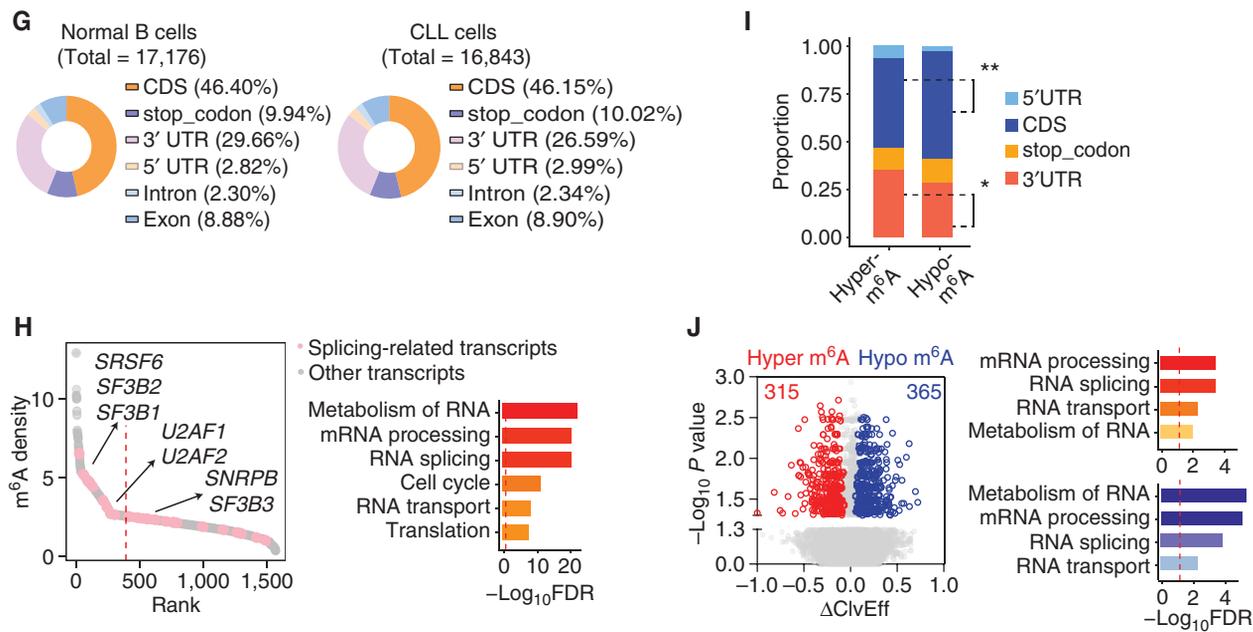
### CLL Cells Exhibit Increased m<sup>6</sup>A Enriched for Transcripts Associated with mRNA Processing

Increased protein abundance of METTL3 in CLL motivated us to examine global m<sup>6</sup>A levels in these cells. Using liquid chromatography mass spectrometry (LC-MS) and m<sup>6</sup>A dot-blot assays, we detected an increase of m<sup>6</sup>A level (1.5–2-fold) in the mRNA derived from CLL cells compared with that of normal B cells (Fig. 3E and F; Supplementary Fig. S4F). There is no difference of m<sup>6</sup>A when analyzing total cellular RNA, confirming that increased m<sup>6</sup>A modification is specific to mRNA. Leveraging an individual-nucleotide-resolution m<sup>6</sup>A modification sequencing technology (MAZTER-seq; refs. 40, 41), we next mapped m<sup>6</sup>A sites on mRNA derived from normal ( $n = 5$ ) and CLL ( $n = 9$ ) B cells (Fig. 3G). The m<sup>6</sup>A writer complex preferentially deposits m<sup>6</sup>A on the consensus sequence, DRA\*CH (D = G/A/U, R = G/A, A\* = m<sup>6</sup>A, H = U/A/C) (25), and the m<sup>6</sup>A modification is enriched around the stop codon. As MAZTER-seq is biased toward ACA sites, we detected GGACA as the most enriched motif within DRACA, accounting for 50% of all the identified m<sup>6</sup>A sites (Supplementary Fig. S4G). Interestingly, transcripts related to RNA metabolism, mRNA processing, and RNA splicing appear to have a higher m<sup>6</sup>A density in the B-cell context (Fig. 3H; Supplementary Table S3), suggesting that these pathways are subjected to m<sup>6</sup>A-dependent modulation.



**Figure 3.** Upregulation of splicing factors is associated with elevated METTL3 expression. **A**, Ranking of nonsplicing factors based on the number of significantly correlated edges (Pearson correlation, FDR <0.05) with splicing factors according to CLL proteomic data. m<sup>6</sup>A-related proteins are black-coded with representative genes labeled. **B**, Protein coexpression network between METTL3 and significantly correlated splicing factors, FDR <0.05. SF3A/3B complex, SR, hnRNP, and other spliceosomal proteins are color coded. The square legend represents the node size which is Pearson correlation coefficient. The size of METTL3 represents 1. **C**, Validation of proteomic data using an independent cohort of normal and CLL B cells by immunoblotting. **D**, Kaplan-Meier plot of TTF1 with METTL3 protein expression. Groups are divided based on the median expression. P = 0.04, log-rank test. **E**, m<sup>6</sup>A level on total RNA and mRNA measured by m<sup>6</sup>A dot-blot assay. \*\*, P = 0.0074, one-tailed Student t test. **F**, m<sup>6</sup>A/A ratio on total RNA and mRNA measured by LC-MS. \*, P = 0.028, one-tailed Student t test. (continued on next page)

Downloaded from <http://aacrjournals.org/bloodcancerdiscovery/article-pdf/4/3/228/3327865/228.pdf> by Harvard Medical School user on 05 May 2023



**Figure 3. (Continued)** **G**, MAZTER-seq was performed using 5 normal B-cell, 2 *SF3B1* MT, and 7 *SF3B1* WT CLL samples. Pie chart shows the percentages of m<sup>6</sup>A sites located within different regions. **H**, Left: ranking of transcripts based on the m<sup>6</sup>A density in all CLL and normal B cell samples, with the upper quartile marked as a red dashed line. Genes involved in RNA splicing are color coded with representative genes labeled. Right: metascape analysis of transcripts within upper quartile; red dashed line: significance cutoff, FDR < 0.05, hypergeometric test. **I**, Transcriptome-wide distribution pattern of m<sup>6</sup>A hypomethylated and hypermethylated sites between CLL and normal B cells. \*,  $P = 0.016$ , \*\*,  $P = 0.0017$ , Binomial test. **J**, Left: volcano plot of differential m<sup>6</sup>A modification sites on mRNA between CLL and normal B cell samples ( $|\Delta\text{ClvEff}| > 0.1$ ,  $P < 0.05$ , Wilcoxon rank sum test, within DRACA motif). Red and blue indicate significant hyper and hypomethylated sites, respectively. Right: metascape analysis of transcripts with significant hyper (top) and hypo (bottom) m<sup>6</sup>A in CLL cells; red dashed line: significance cutoff, FDR < 0.05, hypergeometric test.

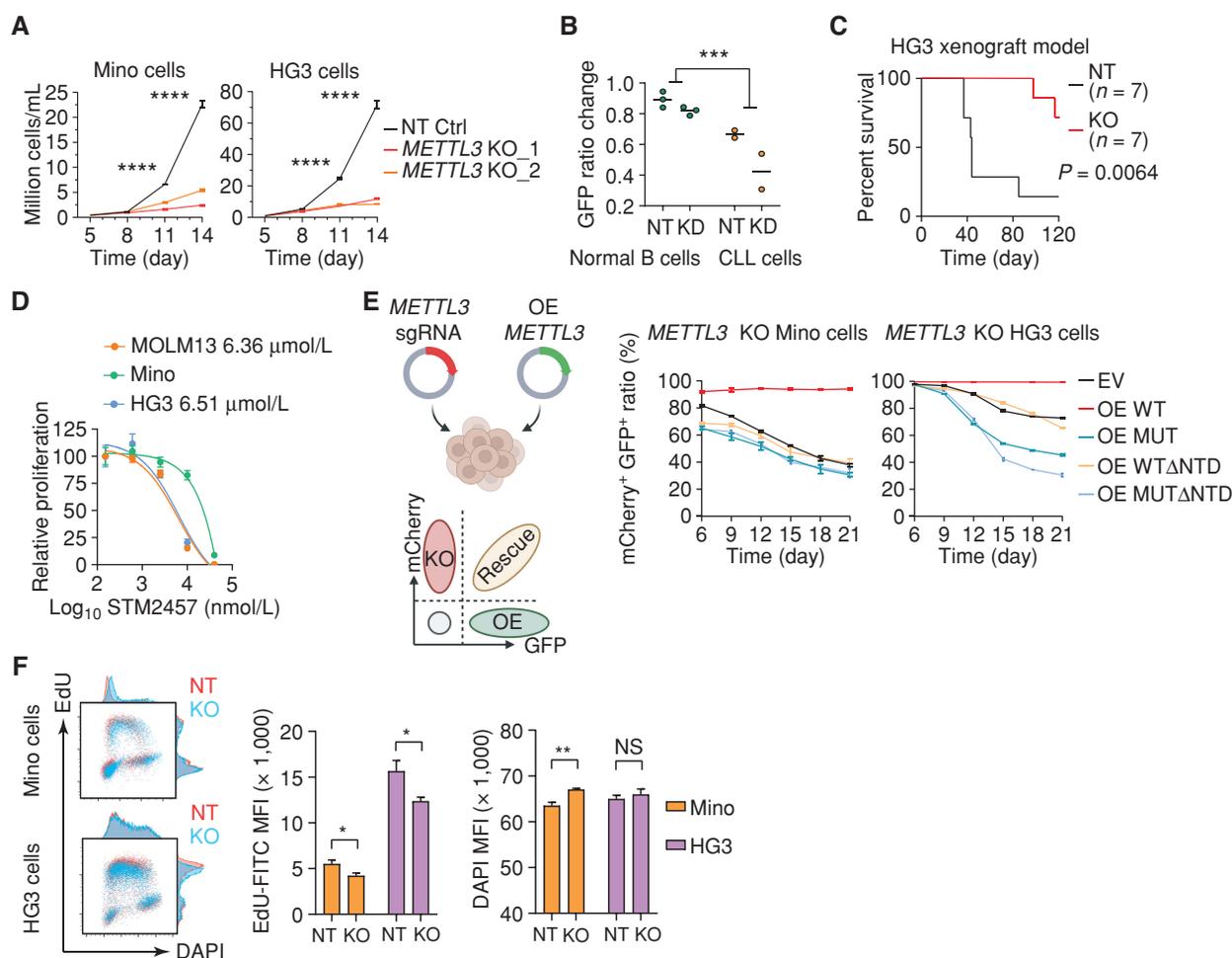
Although CLL and normal B cells shared a similar m<sup>6</sup>A distribution, we observed an increased m<sup>6</sup>A level (hyper-m<sup>6</sup>A) in the 3'UTR regions of CLL cells, concomitant with a decreased m<sup>6</sup>A level (hypo-m<sup>6</sup>A) in coding sequence (CDS) regions (Fig. 3I, Binomial test, CDS:  $P = 0.0017$ ; 3'UTR:  $P = 0.016$ ). These location-dependent m<sup>6</sup>A changes might be caused by the upregulation of both m<sup>6</sup>A writer and eraser proteins in CLL cells (Supplementary Fig. S4A). 680 sites with differential m<sup>6</sup>A modification were identified between normal and CLL B cells with dysregulated transcripts enriched in RNA metabolism-related pathways, including RNA splicing (Fig. 3J; Supplementary Fig. S4H). These results support that m<sup>6</sup>A modification preferentially regulates splicing-related transcripts in CLL in a location-dependent manner.

### METTL3 Is Essential for Cell Proliferation in CLL Cells

*METTL3* has been functionally demonstrated to promote or suppress tumor growth in different cancer types (32, 42). To fully understand the function of *METTL3* in CLL, we examined its impact on cell growth using cell lines and primary B cells both *in vitro* and *in vivo*. Knockout (KO) of *METTL3* with CRISPR/Cas9 technology resulted in substantial inhibition of cell growth in all the mature B cell lines we tested, with Mino and HG3 cells most sensitive to *METTL3* depletion (Fig. 4A; Supplementary Fig. S5A). We depleted *METTL3* by shRNAs in primary normal and CLL B cells and then cocultured them with CD40L-overexpressing stromal cells. Depletion of *METTL3* led to a greater growth disadvantage in CLL cells than

in normal B cells (Fig. 4B; Supplementary Fig. S5B,  $P = 0.007$ ). Our *in vivo* mouse xenograft of HG3 cells demonstrated a similar trend, with KO of *METTL3* inhibiting disease progression and prolonging survival (Fig. 4C,  $P = 0.0064$ ). Treatment with the *METTL3* inhibitor STM2457 also resulted in growth defects in HG3 cells at a level similar to that of MOLM13 cells (Fig. 4D; Supplementary Fig. S5C and S5D), an STM2457-sensitive acute myeloid leukemia (AML) cell line (43). To determine whether *METTL3* is a potential therapeutic target for CLL, we established xenograft mice using HG3 cells with doxycycline-inducible *METTL3* KO and stable luciferase expression. Doxycycline diet led to a significant delay in tumor growth measured by bioluminescence imaging (Supplementary Fig. S5E). Taken together, the evidence supports that *METTL3* is essential to cell growth in mature B-cell leukemia and lymphoma and is a potential therapeutic target for CLL treatment. To expand our understanding of the requirement of m<sup>6</sup>A and RNA-methylation complexes for cell growth in hematologic malignancies, we utilized DepMap data (<https://depmap.org/portal/>; ref. 44) to analyze the dependency of genes related to m<sup>6</sup>A modification in cell lines derived from different hematologic malignancies including AML, acute lymphocytic leukemia, CLL, diffuse large B-cell lymphoma, and multiple myeloma. It appeared that all these cell lines had higher dependency on m<sup>6</sup>A writers (Supplementary Fig. S5F) than erasers or readers for cell proliferation, highlighting the essentiality of m<sup>6</sup>A writers across hematologic malignancies.

To determine which domain of *METTL3* is required for cell growth, we ectopically expressed full-length WT, or



**Figure 4.** KO or pharmacologic inhibition of *METTL3* affects cell proliferation and cell cycle. **A**, Proliferation of Mino and HG3 cells upon *METTL3* KO in technical triplicates. Data are mean  $\pm$  SD. \*\*\*\*,  $P < 0.0001$ , two-tailed Student *t* test. **B**, Normal and CLL B cells were nucleofected with plasmids expressing shRNA-GFP targeting *METTL3* or scramble (NT), and further cultured for two days on LL8 cells. Competitive cell growth was monitored by analyzing the GFP ratio change (day 2 vs. day 0).  $P = 0.007$ , two-way ANOVA. **C**, Kaplan-Meier plot of NSG mice intravenously engrafted with control and *METTL3* KO HG3 cells.  $P = 0.0064$ , log-rank test. **D**, STM2457 dose-response curves for a panel of leukemia cell lines in technical triplicates. Data are mean  $\pm$  SD. IC<sub>50</sub> for each cell line is labeled in the legend. **E**, Competitive growth assay in *METTL3* KO Mino and HG3 cells (with mCherry expression) overexpressed with different *METTL3* constructs (with GFP expression). Cell growth was monitored by analyzing the ratio of mCherry and GFP double-positive cells in technical triplicates using flow cytometry. Data are mean  $\pm$  SD. **F**, Left: cell cycle measured by EdU and DAPI staining in HG3 and Mino cells with or without *METTL3*. Right: mean fluorescence intensity (MFI) of EdU incorporation in cells at S phase (left) and DAPI staining (right) in technical triplicates. Data are mean  $\pm$  SD, two-tailed Student *t* test.

methyltransferase domain catalytic mutant (MUT), or N-terminal domain (NTD, aa1-200)-deleted *METTL3* with or without catalytic domain mutation (WT  $\Delta$ NTD, MUT  $\Delta$ NTD) in Mino and HG3 cells with KO of *METTL3* (Fig. 4E). sgRNAs and *METTL3* expression constructs were tagged with mCherry and GFP, respectively. Using growth competition assay to track the ratio of double-positive cells over time, we observed that WT *METTL3* completely rescued the growth defect in *METTL3* KO cells. However, *METTL3* MUT, WT  $\Delta$ NTD, or MUT  $\Delta$ NTD all lacked the potential to rescue the phenotype. Although *METTL3* has been reported as an oncogene to promote the initiation and development of a variety of cancers (32, 42, 45), we failed to detect significantly enhanced cell proliferation upon overexpression of WT *METTL3* in both Mino and HG3 cells (Supplementary Fig. S5G and S5H). Of note, overexpression of either catalytic mutant *METTL3*

(MT) or WT  $\Delta$ NTD *METTL3* led to more drastic growth disadvantage in HG3 cells (Supplementary Fig. S5G) possibly due to competitively binding with another m<sup>6</sup>A writer protein *METTL14*, which interfered with its functions. These results highlight that both the catalytic domain and NTD of *METTL3* play indispensable roles in cell growth.

### Loss of *METTL3* Affects the Translation of Splicing Factors

To dissect the underlying molecular mechanism by which *METTL3* affects cell growth, we examined if KO of *METTL3* affects B-cell receptor signaling, cell cycle, and apoptosis. Loss of *METTL3* subtly increased early apoptosis (Supplementary Fig. S6A). Upon anti-IgM stimulation in JeKo-1 cells, AKT pathway activation was slightly reduced in STM2457-treated cells but ERK activation had no change (Supplementary Fig. S6B).

Moreover, we observed a small but significant reduction of 5-ethynyl-2'-deoxyuridine (EdU) incorporation during the S phase and higher overall DNA content, suggesting compromised DNA synthesis and mitotic stress (Fig. 4F). Likewise, treatment with STM2457 also led to consistently reduced incorporation of EdU with subtle impact on the cell cycle (Supplementary Fig. S6C and S6D). Taken together, these changes suggest that other mechanisms beyond BCR signaling, cell cycle, or apoptosis may contribute to growth defects observed in *METTL3* KO cells.

To define whether *METTL3* KO affects translation, we examined O-propargyl-puromycin (OPP) incorporation into newly synthesized proteins in cells with or without *METTL3* and found loss of *METTL3* resulted in a decrease in OPP incorporation (Fig. 5A; Supplementary Fig. S6E). We further performed an integrated ribosome profiling (Ribo-seq) and RNA-seq analysis on HG3 cells with or without *METTL3* KO to determine the transcripts for which *METTL3* preferentially affects translation. Consistent with the OPP assay, loss of *METTL3* decreased overall translation efficiency (TE; Fig. 5B), defined as the number of ribosome-protected fragments (RPF) divided by the mRNA expression. Analysis of the top 1,000 transcripts with decreased TE revealed that *METTL3* KO preferentially affected the transcripts involved in the metabolism of RNA, cell-cycle, RNA splicing, and mRNA processing (Fig. 5B), in which splicing factors (*SF3B1*, *SF3A1*, *SF3A2*, and *SF3A3*) were confirmed by immunoblot (Fig. 5C). *METTL14* expression was also decreased upon KO of *METTL3*. In contrast, we detected no gene-expression changes at the RNA level, highlighting the importance of posttranscriptional regulation of splicing factors by *METTL3* in mature B cells (Fig. 5D).

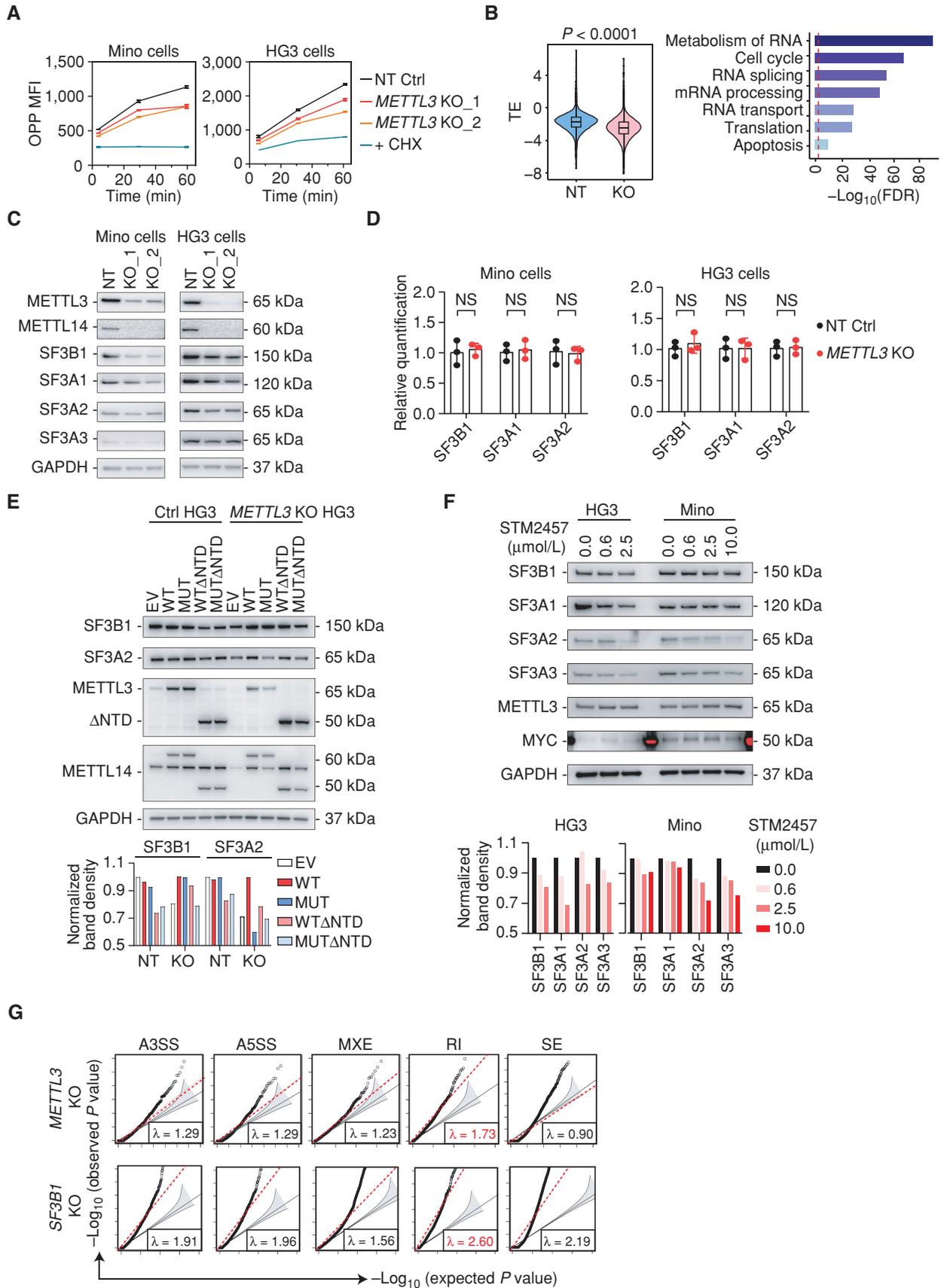
We next determined whether the overexpression of *METTL3* can rescue splicing factor expression. In response to overexpression of WT, but not catalytic mutant *METTL3*, splicing factor expression and *METTL14* abundance were restored in *METTL3* KO cells (Fig. 5E), indicating that protein expression of splicing factors is methyltransferase-dependent. Consistent with this observation, treatment with STM2457 also decreased the abundance of splicing factors (Fig. 5F; Supplementary Fig. S6F), suggesting that the NTD is functionally dependent on the catalytic domain of *METTL3* in CLL. Furthermore, we found overexpression of *SF3B1* in *METTL3* KO cells partially rescued the growth defect in HG3 cells, highlighting splicing factors as the targets for *METTL3* in CLL (Supplementary Fig. S6G). Apart from *METTL3*, *YTHDF* paralogs are reported as m<sup>6</sup>A reader proteins enhancing the translation of m<sup>6</sup>A installed transcripts (26, 46). Although KO of *YTHDF2* led to growth defects, none of *YTHDF* paralogs affected splicing factor expression (Supplementary Fig. S6H–S6J). Taken together, our results suggest that *METTL3* likely influences the translation of splicing factors by using both catalytic domain and NTD.

To determine the relationship between *METTL3*-regulated splicing factor expression and RNA splicing dysregulation in CLL, we measured the direct impact of splicing factor protein expression change on RNA splicing. We chose the splicing factor *SF3B1* to examine this relationship as it is one of the most consistently upregulated splicing factors in CLL. Splice variants associated with the depletion of either *SF3B1* or *METTL3* displayed similar splicing pattern change, with intron retention as the most affected splicing type (Fig. 5G; Supplementary Fig. S6K), suggesting *SF3B1* protein change is involved in the RNA splicing dysregulation in CLL. To further examine the contribution of *SF3B1* and *METTL3* protein expression to RNA splicing changes observed in CLL, we first identified “HG3-associated aberrant splicing events” (similar to CLL-associated events) by comparing RNA-seq data derived from HG3 cells and normal B cells. We then determined whether *SF3B1* KO or *METTL3* KO could rescue HG3-associated aberrant splicing events. 66.9% and 65.2% of HG3-associated aberrant splicing events showed  $|\Delta\text{PSI}|$  more than 0.1 upon *SF3B1* KO or *METTL3* KO, respectively (Supplementary Fig. S6L). In contrast, only 12.1% and 11.2% of events that are not HG3-associated aberrant splicing events had such changes (Supplementary Fig. S6L). Moreover, 85% of *SF3B1* KO- and *METTL3* KO-rescued events were overlapped (Supplementary Fig. S6M). These results indicated that *METTL3*-regulated splicing factor expression strongly influences RNA splicing in CLL.

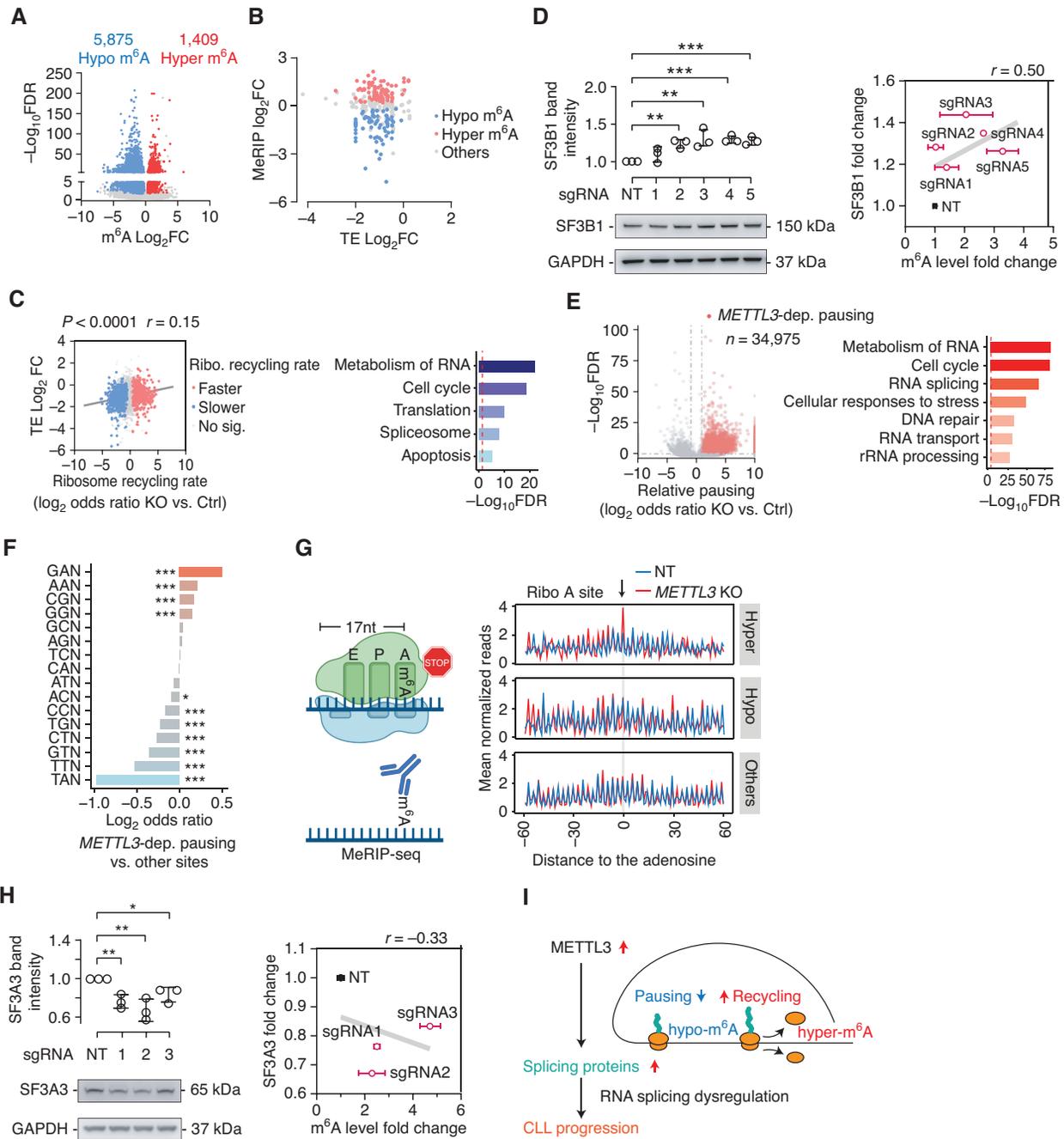
### KO of *METTL3* Regulates Abundance of Splicing Factors via Modulating m<sup>6</sup>A at Stop Codons and CDS Regions

To dissect the mechanism by which *METTL3* influences the abundance of splicing factors via m<sup>6</sup>A modification, we performed an integrative methylated RNA immunoprecipitation sequencing (MeRIP-seq), which identifies transcriptome-wide m<sup>6</sup>A sites, Ribo-seq, and RNA-seq analysis in cells with and without *METTL3* KO. We first mapped out *METTL3*-responding m<sup>6</sup>A sites upon *METTL3* KO. 5,875 hypomethylated and 1,409 hypermethylated sites were detected, with GGACU as the most enriched motif (Fig. 6A; Supplementary Fig. S7A), validating the role of *METTL3* as an RNA methyltransferase. Differential m<sup>6</sup>A sites were primarily found in both exons and 3'UTRs, with the highest enrichment of m<sup>6</sup>A residues located near the stop codon (Supplementary Fig. S7A). Gene-expression changes negatively correlated with that of the m<sup>6</sup>A level, with upregulated genes enriched in inflammatory response and TNF $\alpha$  signaling pathways (Supplementary Fig. S7B). We further integrated MeRIP-seq and Ribo-seq data to determine the proteins specifically affected by *METTL3*-associated m<sup>6</sup>A changes. Interestingly, our results revealed that genes with reduced TE displayed not only

**Figure 5.** KO or pharmacologic inhibition of *METTL3* affects protein expression of splicing factors. **A**, OPP Run-On assay on Mino and HG3 cells either with or without *METTL3*. OPP incorporation was assessed at 5, 30, and 60 minutes after OPP exposure by flow cytometry. Cycloheximide (CHX) treated cells as positive control. **B**, Left: violin plot of TE in control and *METTL3* KO HG3 cells.  $P < 0.0001$ , Wilcoxon signed-rank test. Right: metascape analysis of the top 1,000 transcripts with downregulated TE; red dashed line: significance cutoff,  $\text{FDR} < 0.05$ , hypergeometric test. **C**, Immunoblot of splicing factors expression in control and *METTL3* KO Mino and HG3 cells. **D**, qPCR results of *SF3B1*, *SF3A1*, and *SF3A2* expression in control and *METTL3* KO Mino and HG3 cells. Two-tailed Student *t* test. **E**, Splicing factors expression in control and *METTL3* KO HG3 cells that overexpressed with different *METTL3* variants detected by immunoblotting. **F**, Splicing factors expression in Mino and HG3 cells treated with different dosage of STM2457 for 7 days detected by immunoblotting. **G**, Major changed alternative splicing type corresponds to either *METTL3* or *SF3B1* KO in HG3 cells. Q-Q plots of observed *P* values against the expected *P* values under the uniform distribution of five alternative splicing categories. Red lines indicate the least-squares linear fit to the lower 95th percentile of points with slope  $\lambda$ . Gray-shaded areas represent 95% confidence intervals for the expected distribution.



Downloaded from <http://aacrjournals.org/bloodcancerdiscovery/article-pdf/4/3/228/3327865/228>; pdf by Harvard Medical School user on 05 May 2023



**Figure 6.** *METTL3* regulates splicing factor abundance through  $m^6A$ -mediated translational control. **A**, Volcano plot of differential  $m^6A$  modification sites upon *METTL3* KO in HG3 cells. **B**, Scatter plot of splicing factors between  $m^6A$  changes and TE changes in HG3 cells with and without *METTL3*. **C**, Left: correlation scatter plot of changes between TE and ribosome recycling rate upon KO of *METTL3*,  $P < 0.0001$ ,  $r = 0.15$ , Pearson correlation. Color-coded dots indicate transcripts with significant changes in ribosome recycling rate.  $|\text{Log}_2 \text{ odds ratio}| > 1$ ,  $\text{FDR} < 0.05$ , Fisher exact test. Gray line: linear regression. Right: metascape analysis of transcripts with a significant decrease of ribosome recycling rate. Red dashed line: significance cutoff,  $\text{FDR} < 0.05$ , hypergeometric test. **D**, Left: immunoblot of SF3B1 in dCasRx-*METTL3* HG3 cells transduced with correspond sgRNA. Quantification results are from three biological replicates. Data are mean  $\pm$  SD. \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ , two-tailed Student t test. Right: correlation scatter plot of  $m^6A$  level fold change in technical triplicates as mean  $\pm$  SD, and average fold change of SF3B1 protein expression. Gray line: linear regression. **E**, Relative ribosome pausing in CDS regions upon *METTL3* KO. The colored points indicate *METTL3*-dependent pausing sites ( $\text{Log}_2 \text{ odds ratio} > 1$ ,  $\text{FDR} < 0.05$ ), with all other translational positions in gray, Fisher exact test. Metascape analysis of transcripts with significantly increased ribosome pausing upon *METTL3* KO. Red dashed line: significance cutoff,  $\text{FDR} < 0.05$ , hypergeometric test. **F**, A site codon associated with *METTL3*-dependent vs. other pausing sites (32,473 vs. 1,763,686 sites), two-sided Fisher exact tests. **G**, Aggregation plots of the mean ribosome densities (at ribosomal A sites) along mRNA aligned to the  $m^6A$  sites. Blue and pink denote HG3 cells with or without *METTL3*, respectively. **H**, Left: immunoblot of SF3A3 in dCasRx-*METTL3* HG3 cells transduced with correspond sgRNA. Quantification results are from three biological replicates. Data are mean  $\pm$  SD. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ , two-tailed Student t test. Right: correlation scatter plot of  $m^6A$  level fold change in technical triplicates as mean  $\pm$  SD, and average fold change of SF3A3 protein expression. Gray line: linear regression. **I**, Working model of *METTL3*-mediated  $m^6A$  modification controls splicing factor abundance to impact the progression of CLL.

Downloaded from <http://aacrjournals.org/bloodcancerdiscovery/article-pdf/14/3/228/3327865/228> by Harvard Medical School user on 05 May 2023

hypomethylated but also hypermethylated sites (Supplementary Fig. S7C), possibly as direct or indirect targets of *METTL3*. We then explicitly examined m<sup>6</sup>A status on the mRNA of splicing factors ( $n = 189$ ). 104 sites in 67 splicing factors out of 122 splicing factors with detectable m<sup>6</sup>A sites were found with dysregulated m<sup>6</sup>A status (Fig. 6B). Transcripts with hypomethylated sites were enriched at stop codon regions for most splicing factors, including *SF3A2*, *SRSF1*, *SRSF5*, *SRSF6*, *SF3A3*, and *SF3B1*, indicating these sites as direct target sites of *METTL3* (Fig. 6B; Supplementary Fig. S7D). We further verified that *METTL3* KO-induced methylation reduction in these splicing factors is not due to decreased RNA expression level as detected by MeRIP-PCR and qPCR (Supplementary Fig. S7E).

### KO of *METTL3* Influences Ribosome Recycling to Regulate Protein Translation

Although *METTL3* has been reported to promote protein synthesis by binding eIF3h and forming an mRNA loop via m<sup>6</sup>A modification at stop codon regions (47, 48), the mechanism by which *METTL3* KO influences translation via m<sup>6</sup>A modification is not completely elucidated. It is known that mRNA looping can facilitate faster ribosome recycling (49), which in turn could promote translation. The time of the ribosome at the stop codon regions reflects the ribosome recycling ratio. To determine whether *METTL3* KO reduces translation by delaying ribosome recycling via downregulation of m<sup>6</sup>A methylation at the stop codon region, we analyzed the ribosome recycling rate by assessing the ribosomal coverage of P (peptidyl)-sites at CDS regions relative to that of stop codon regions. The change of TE was significantly correlated with that of ribosome recycling rate upon *METTL3* KO (Fig. 6C,  $r = 0.15$ ), and transcripts with decreased ribosome recycling rate were enriched for the RNA splicing pathway (Fig. 6C), suggesting that *METTL3* regulates the splicing factor expression via ribosome recycling through m<sup>6</sup>A modification at or near the stop codon region.

To validate the functional association between m<sup>6</sup>A modification and splicing factor abundance, we used an m<sup>6</sup>A editing platform CRISPR-dCas13Rx-METTL3 to install m<sup>6</sup>A in a site-specific manner. Specifically, we fused dCasRx with the methyltransferase domain of *METTL3* along with nuclear localization signal (NLS), and established stable HG3 cells with overexpression of CRISPR-dCas13Rx-METTL3 (Supplementary Fig. S8A and S8B). m<sup>6</sup>A at the hypomethylated site (m<sup>6</sup>A<sub>3972</sub>) of *SF3B1* was selected for validation as *SF3B1* is one of the most consistently upregulated proteins in CLL (Supplementary Table S4,  $q = 0.0033$ ). Upon introduction of m<sup>6</sup>A at the 3972 site, we observed a 1.4-fold increase of *SF3B1* protein in HG3 cells, with its RNA level unchanged (Fig. 6D; Supplementary Fig. S8C), confirming *METTL3* directly regulates *SF3B1* expression via m<sup>6</sup>A. Collectively, these results suggest that *METTL3* controls splicing factor translation via slowing ribosome recycling at stop codon regions in an m<sup>6</sup>A-dependent manner.

### KO of *METTL3* Affects Ribosome Pausing to Regulate Protein Translation

In addition to direct targets (Fig. 6B, hypomethylated transcripts), we also found hypermethylated transcripts associated with *METTL3* KO, which may be indirect targets of *METTL3*. These sites were more localized at CDS and 5' UTR region of mRNA (Supplementary Fig. S7A). m<sup>6</sup>A at the CDS

region has been reported to impact the kinetics of tRNA entering into the ribosomal aminoacyl (A) sites and leading to ribosome pausing, which in turn affects ribosome elongation and reduces translational output (29, 50). We therefore sought to determine whether *METTL3* KO-associated hypermethylated m<sup>6</sup>A sites at CDS regions can reduce TE via increased ribosomal pausing. We devised a ribosome pausing score (51) for each codon site in the CDS regions to assess the impact of m<sup>6</sup>A on translation elongation. Pause scores at each codon position in a transcript were calculated by dividing the number of ribosome-protected reads at that position by the average number of reads of that transcript. Within 1,798,661 ribosome-covered sites from our data, 34,975 sites (1.94%) displayed significantly increased ribosome pausing upon *METTL3* KO compared with control cells, which is substantially higher than 6,788 sites with decreased pausing (Fig. 6E). Increased pausing sites were highly enriched in genes involved in RNA splicing (Fig. 6E). Of note, the sequence specificity of pausing sites was mostly enriched for GAN in the ribosomal A sites and well matched with the m<sup>6</sup>A motif (Fig. 6F; Supplementary Fig. S8D), strongly implicating ribosome pausing is tightly linked with m<sup>6</sup>A modification.

To determine whether m<sup>6</sup>A modification in these transcripts results in ribosome pausing, we performed an integrative MeRIP-seq and Ribo-seq analysis and found an elevated ribosome density on hypermethylated adenosines at the ribosomal A sites (Fig. 6G), corroborating that hypermethylation influences the decoding process. In supporting this observation, we further experimentally validated that increased m<sup>6</sup>A modification at the endogenous CDS region of *SF3A3* decreased its protein expression with no impact on the overall mRNA abundance using dCasRx-METTL3 (Fig. 6H; Supplementary Fig. S8E). This result led us to consider the possible underlying mechanisms that account for m<sup>6</sup>A methylation upon *METTL3* KO. Systematic examination of m<sup>6</sup>A writers and erasers revealed that *METTL3* KO resulted in an increase in *METTL16* and a decrease in *ALKBH5* protein expression (Supplementary Fig. S8F), indicating that a complex m<sup>6</sup>A circuitry precisely controls the RNA splicing network by controlling splicing factor protein expression.

## DISCUSSION

The discovery of recurrent somatic mutations in the core spliceosomal components and their significant associations with clinical outcomes emphasized the importance of genome-wide RNA splicing dysregulation and oncogenesis (13–15, 52–54). Hence, understanding the underlying mechanism of RNA splicing dysregulation, especially in cancer types without genetic lesions in spliceosomal components, becomes a high priority in cancer biology. Here, we report that RNA spliceosome abundance is an independent risk factor for poor prognosis in CLL, which contributes to splicing dysregulation and disease progression. Moreover, we discover that *METTL3* regulates RNA splicing dysregulation through m<sup>6</sup>A-mediated translational control of splicing factors (Fig. 6I). Our results have implications in cancer biology, disease progression, and treatment.

Our studies highlight that the novel regulatory axis between *METTL3* and splicing factors is likely to be a shared mechanism in cancer. Disruption of the regulatory axis led to cell growth defects *in vitro* and affected leukemia progression *in vivo* in

CLL. Moreover, inhibition of METTL3 with STM2457 in multiple cancer types all resulted in cell death along with reduced expression of splicing factors, highly suggestive of a common regulatory axis controlling cancer cell growth. Given *METTL3* is significantly upregulated in CLL and has been reported to act as an oncogene in multiple cancer types (23, 42, 55), we speculate that METTL3 may contribute to the onset and progression of CLL via regulating splicing networks. Previous studies in AML have indicated that MYC is one of the METTL3 target genes contributing to oncogenesis (23); however, our results showed that no change of MYC was detected upon inhibition or KO of METTL3 in B-cell contexts, suggesting cellular context and differentiation stage are another layer of regulation of m<sup>6</sup>A-dependent mechanisms. With oral METTL3 inhibitors entering clinical trials, we observed the therapeutic effects of *METTL3* silencing in CLL xenograft mice, providing a proof-of-concept for targeting *METTL3* in aggressive CLL. Future experiments with CLL murine models are expected to comprehensively elucidate the role of METTL3 in a B-cell context as well as the therapeutic effect of STM2457 *in vivo*.

Our results also uncover novel mechanisms that METTL3 utilizes to control the translation of splicing factors in a site-dependent fashion, which includes promoting fast ribosome recycling via m<sup>6</sup>A at stop codon regions to enhance translation and increasing ribosome pausing via m<sup>6</sup>A at CDS regions to perturb ribosomal elongation along transcripts and slow mRNA translation. This is different from previous studies, which reported that METTL3 enhances translation either by forming mRNA looping through binding with eIF3 or by inducing m<sup>6</sup>A on a subset of transcripts associated with chromatin and relieving ribosome stalling (32, 47). All these observations support that METTL3 uses multiple mechanisms to ensure proper protein regulation, implicating a critical role of METTL3 in cancer biology. This led us to propose a working model for METTL3 in the pathogenesis of CLL. Transcriptional upregulation of *METTL3* in CLL causes an elevated protein expression of METTL3, which promotes m<sup>6</sup>A at stop codon regions to enhance translation of splicing factors whereas it also modulates m<sup>6</sup>A at CDS regions to remove translational machinery roadblocks to ensure upregulation of splicing factors. The increased abundance of splicing complex in turn generates aberrant RNA splicing and contributes to disease onset and progression (Fig. 6I).

Altogether, our study highlights the critical role of RNA splicing dysregulation in CLL that can be induced by a splicing factor mutation or an aberrant expression of an upstream regulator like m<sup>6</sup>A writer *METTL3*, which regulates the translation of splicing factor-encoding mRNAs via m<sup>6</sup>A methylation-induced ribosome pausing and recycling, leading to splicing dysregulation and progression of the disease.

## METHODS

### Human Samples

Heparinized blood samples were obtained from healthy donors and patients enrolled on clinical protocols with informed consent, approved by the Human Subjects Protection Committee of City of Hope (IRB#18067 and IRB#06229) or Dana-Farber Cancer Institute. All studies were performed following written informed consent from patients and in accordance with the Declaration of Helsinki. Peripheral blood mononuclear cells were isolated by density gradient centrifugation using Ficoll-Paque Medium (GE Healthcare). Normal B

cells were isolated by immuno-magnetic negative selection with Pan B-cell isolation kit (MiltenyBiotec). All samples were cryopreserved with fetal bovine serum (FBS) 10% DMSO and stored in vapor-phase liquid nitrogen until the time of analysis. In total, 40 CLL samples were used from MAZTER-seq, TMT proteomics, and RNA-seq. The samples were selected based on *SF3B1* mutation status and cytogenetic abnormalities to represent the spectrum of CLL.

### Cell Lines and Reagents

Leukemia cell lines NALM6 (CRL-3273, ATCC), Mino (CRL-3000, ATCC), JeKo-1 (CRL-3006, ATCC), HG3 (provided by Dr. Richard Rosenquist, Uppsala University), MEC1 (ACC497, DSMZ), and MOLM13 (provided by Dr. Jianjun Chen, City of Hope) were cultured in RPMI-1640 (Invitrogen) supplemented with 10% FBS and 1% penicillin/streptomycin. Lenti-X 293T cells (TaKaRa) and LL8 cells with CD40 ligand overexpression (provided by Dr. John Chan, City of Hope) were cultured in DMEM with 10% FBS and 1% penicillin/streptomycin. When cocultured with primary B cells, LL8 cells were pretreated with 5 µg/mL Mitomycin C (Millipore Sigma) for 3 hours followed by medium change. We received these cell lines in 2018, and our recent STR profiling was conducted in 2022. All the cell lines were used within 25 passages. All cell lines tested negative for *Mycoplasma*.

Antibodies used in this study include anti-METTL3 (#A301-568A, Bethyl Laboratories; #ab195352, Abcam), anti-METTL14 (#HPA038002, Millipore Sigma), anti-METTL16 (#HPA020352, Millipore Sigma), anti-FTO (#ab126605, Abcam), anti-ALKBH5 (#ab195377, Abcam), anti-SF3B1 (#14434S, Cell Signaling Technology, #PAS-19679, Invitrogen), anti-SF3A1 from Reed Laboratory, anti-SF3A2 (#sc-390444, Santa Cruz Biotechnology), anti-SF3A3 (#sc-393673, Santa Cruz Biotechnology), anti-MYC (#D84C12, Cell Signaling Technology), anti-YTHDF1 (#57539S, Cell Signaling Technology), anti-YTHDF2 (#ab220163, Abcam), anti-YTHDF3 (#24206S, Cell Signaling Technology), anti-Annexin V (#640906, BioLegend), anti-GAPDH (#sc-365062, Santa Cruz Biotechnology), and goat F(ab')<sub>2</sub> anti-human IgM-UNLB (#2022-01, SouthernBiotech). Secondary antibody: Goat anti-Rabbit IgG secondary antibody, HRP (#65-6120, Invitrogen), and goat anti-mouse IgG secondary antibody, HRP (#65-6520, Invitrogen). Horseradish peroxidase activity was revealed using Clarity or Clarity Max ECL Western Blotting Substrates (#1705061 or #1705062, Bio-Rad). STM2457 (#HY-134836) was purchased from MedChemExpress.

### Plasmid Constructions

Catalytic wild-type and mutant *METTL3* sequence was amplified from pCSC-METTL3 (WT/MUT)-IRES-GFP plasmids (provided by Dr. Yanhong Shi, City of Hope; ref. 45) using the following primers: forward 5'-CTAGACTAGTCCACCATGTCCGACACGTGGAGCTCT ATC-3' and reverse 5'-ATAAGAATGCGGCCGCCTATAAAATCTT A GGTTTAGAGATGATACCATCTGGG-3', and then cloned into pCRII-Blunt-TOPO plasmid (Kit #450245, Thermo Fisher Scientific).

*METTL3* sgRNA1 and sgRNA2 targeted sequences were silently mutated into 5'-AAGCGACCTCGCACTCACGT-3' and 5'-GAAC TAATC GAAGTTAAAAG-3' by using Q5 Site-Directed Mutagenesis Kit (Kit #E0554S, New England Biolabs). This kit was also used to generate NTD- (aa1-187) deleted *METTL3*. Full-length and NTD-deleted catalytic WT and mutant *METTL3* were amplified by PCR and cloned into pLVX-EF1α-IRES-zsGreen1 plasmid (TaKaRa). Nontargeting shRNA and shRNAs targeting *METTL3* were amplified from pHIV7-mhC/shMETTL3-GFP plasmids (provided by Dr. Yanhong Shi, City of Hope), and further cloned into pMAX-GFP plasmid (Lonza). Cas9-sgRNA targeting to *METTL3*, *YTHDF1*, *YTHDF2*, *YTHDF3*, and *SF3B1* were subcloned into either pLKO.5-mCherry or tet-pLKOsgRNA-puro plasmid (#104321, addgene). All shRNA and sgRNA sequences were shown in Supplementary Table S5. *SF3B1* with N-terminal 3 × Flag tag was cloned into pLVXEF1α-IRES-zsGreen1 plasmid (TaKaRa) between SpeI and BglIII. dCasRx, deactivated RfxCas13d, along with the 5'NLS

was amplified from the pXR002 plasmid (#109050, addgene) using Q5 polymerase (# M0494S, New England Biolabs) and cloned into pLVX-EF1 $\alpha$ -IRES-zsGreen1 plasmid (TaKaRa) between SpeI and NotI. A FragmentGene (GeneWiz) containing a 3'NLSXTEN-METTL3 (aa273–580)-HA sequence served as a template for amplifying and joining the METTL3 methyltransferase domain to dCasRx via NotI and BamHI, generating pLVX-dCasRx-NLS-METTL3IRES-ZsGreen1. dCasRx-METTL3 was generated by cloning an HA epitope tag (Y-P-Y-D-V-P-D-Y-A) along with methyltransferase domain of METTL3 onto the C terminus of dCasRx epitranscriptomic editors.

### Lentivirus Production and Transduction

For lentivirus production, Lenti-X 293T cells were plated on 6-well plates at  $3 \times 10^5$  per well in complete DMEM with 10% serum and allowed to adhere for 16 hours. Transfections were performed following the manufacturer's instruction with 3  $\mu$ g of pLKO.5-sgRNA, 0.4  $\mu$ g pMD2.G, 1.5  $\mu$ g pSPAX2, with PEI-MAX 40K (Polysciences). The lentivirus particles were collected at 48 and 96 hours after transfection and filtered through a 0.45- $\mu$ m Nalgene syringe filter SFCA (Whatman), concentrated with ultracentrifuge at 4°C,  $10,000 \times g$  for 2 hours, and resuspended in 300  $\mu$ L medium. Spin infection was used in all our leukemia and lymphoma cell lines with lentivirus. Specifically, 0.5 million cells were plated in 1 mL of complete media supplemented with 8  $\mu$ g/mL polybrene and concentrated virus, and then the plate was spun at 37°C for 1.5 hours at  $950 \times g$ . The medium was changed 8 hours after the spin transduction.

### Xenograft Mice Models

For control and *METTL3* KO HG3 xenograft mice, we first transduced nontarget control or *METTL3* sgRNA into HG3 cells with Cas9 expression. Then, 6- to 8-week-old NSG mice were injected with  $3 \times 10^5$  HG3 cells by intravenous injection. For doxycycline-inducible control and *METTL3* KO HG3 cells xenograft mice, we generated doxycycline-inducible nontarget control or *METTL3* KO cells and then overexpress luciferase. Then, we engrafted 6- to 8-week-old NSG mice by subcutaneous injection of  $4.5 \times 10^5$  cells. Doxycycline diet was given 1 week after the engraftment. Whole-body bioluminescence imaging was performed with Lago X (Spectral Instruments Imaging) at 10 minutes after D-luciferin (Goldbio) intraperitoneal cavity injection. All the protocols were approved by the Institutional Animal Care and Use Committee (IACUC) of City of Hope.

### Cell Proliferation Assay

To study the effects of *METTL3* on cell proliferation, cells were seeded in a 24-well plate at a concentration of 20,000 to 40,000 cells per well in triplicate, and cell proliferation was assessed by absolute quantification by coupling flow cytometry and cell counting with hemocytometer. Briefly, at each time point, 150  $\mu$ L cells were collected and mixed with 50  $\mu$ L of precounted CFP-positive cells ( $5 \times 10^5$ /mL). Cell proliferation was assessed by the following two steps: Calculate the ratio of mCherry<sup>+</sup> cells (for KO cells with sgRNAs) or mCherry<sup>+</sup>GFP<sup>+</sup> cells (for *METTL3* reconstituted cells) as well as the ratio of CFP<sup>+</sup> cells in the mixed cells. Quantify the cell density based on the counting of CFP<sup>+</sup> cells using hemocytometer.

### Cell-Cycle and OPP Run-On Assays

For the cell-cycle assay, cells were starved for 24 hours in a medium without FBS. The next day, the cells were replenished with a medium containing 10% FBS and cultured for 1 hour. EdU (Invitrogen) was then added into the medium at a concentration of 10  $\mu$ M/L for 1 hour before being collected for staining. For the OPP Run-On assay, cells were pretreated with cycloheximide (Thermo Fisher Scientific) at a concentration of 50  $\mu$ g/mL for 12 hours as a positive control. OPP (Click Chemistry Tools) was added to the medium at a concentration of 20  $\mu$ M/L for 0.5 and 1 hour and then collected for staining. After either EdU or OPP

treatment, the Click-&-Go Plus 488 Imaging Kit (#1314, Click Chemistry Tools) was used to assess the dye incorporation efficiency based on the manufacturer's recommended procedures. Then, the cells were stained with DAPI. All the samples were analyzed on an LSRFortessa cell analyzer. Data were analyzed with FlowJo software (Tree Star).

### Electroporation of Plasmids into Normal and CLL B Cells

$1 \times 10^6$  cells were collected and washed with PBS before electroporation. Cell pellet was resuspended in 20  $\mu$ L electroporation buffer (Celextrix) with 4  $\mu$ L (10  $\mu$ g) pMAX-shRNA plasmid. The cells were then electroporated (Celextrix electroporator LE+) with the condition of 870 V, 20 ms, and 1 pulse. The electroporated cells were cocultured with LL8 cells in RPMI-1640 medium supplemented with 20% FBS, antibiotics, and 2 ng/mL recombinant human IL4 cytokines (R&D Systems) for monitoring cell survival.

### Drug and Proliferation Assays

All suspension cells were plated in 96-well plates in triplicate at 5,000 to 10,000 cells/well in 200  $\mu$ L medium and treated for 72 hours either with 0.5% dimethyl sulfoxide (DMSO) or different concentrations of STM2457 (0.04–40  $\mu$ M/L). After 72 hours, 100  $\mu$ L aliquots of cells were collected for the day 3 survival assay. The remaining cells were continuously cultured using a fresh medium with the addition of STM2457. The ATPase-based CellTiter-Glo 2.0 Cell Viability Assay (#G9242, Promega) was used to measure cell survival on cells collected on days 3 and 7. Relative cell survival was calculated by normalizing ATP amount in the STM2457- versus DMSO-treated cells.

### Immunofluorescence Staining

HEK293T expressing dCasRx-METTL3 were seeded at  $5 \times 10^5$ /mL and cultured on coverslips (#CLS1760-015, Chemglass) in 24-well plates. After 48 hours, the culture medium was discarded, and the coverslips were washed once with PBS gently. Cells were fixed with 4% paraformaldehyde for 15 minutes at room temperature, washed three times with PBS, and then permeabilized with PBS containing 0.2% Triton-X100 (PBST) for 15 minutes at room temperature. Cells were blocked in blocking buffer (10% BSA in PBST) for 30 minutes and stained with anti-HA antibody (Antibody #2367, Cell Signaling Technology) in blocking buffer overnight at 4°C. Cells were then washed three times with PBST and stained with anti-mouse secondary antibodies conjugated with Alexa Fluor 647 (Antibody #A10037, Thermo Fisher Scientific) in blocking buffer for 1 hour at RT. Coverslips were then washed three times in PBST and mounted on mounting media (F6057-20ML, Sigma-Aldrich).

Images were acquired using a confocal laser scanning microscope (FV3000-IX83, Olympus).

### m<sup>6</sup>A LC-MS, Dot Blot, and SELECT Assays

Total RNA was extracted using TRIzol, and polyA RNA was further enriched with PolyAtract mRNA isolation System IV (Kit #Z5310, Promega) in accordance with the manufacturer's instructions. For LC-MS, polyA RNA was digested by nuclease P1 (N8630, Sigma-Aldrich) for 1 hour at 42°C. Subsequently, 1 Unit alkaline phosphatase (P5931, Sigma-Aldrich) and NH<sub>4</sub>HCO<sub>3</sub> (100 mmol/L) were added and incubated for another 1 hour at 37°C. Enzymes were removed by Ultrafiltration. Samples were loaded onto the C18 column and eluted by gradient methanol. The Ultimate 3000 system coupled with a TSQ Quantiva mass spectrometer (Thermo Fisher Scientific) was applied to quantify the levels of Adenosine and m<sup>6</sup>A. For the m<sup>6</sup>A dot blot, total RNA and polyA RNA were diluted in RNA-binding buffer and denatured at 65°C for 5 minutes. Then, one volume of 20 $\times$  SSC buffer was added into the RNA samples before dotted onto Amersham Hybond-N+ membrane (#45-000-850, GE Healthcare) with Bio-Dot Apparatus (Bio-Rad).

The RNA samples were cross-linked onto the membrane via UV irradiation with the UV crosslinker (Thermo Fisher Scientific). The

membrane was stained with 0.02% methylene blue (MB) as a loading control. After UV cross-linking and MB staining, the membrane was washed with PBST, blocked with 5% nonfat dry milk for 1 hour at room temperature, and incubated with the anti-m<sup>6</sup>A antibody at 4°C overnight. The membrane was then incubated with secondary HRP-conjugated goat anti-rabbit IgG antibody and developed with Clarity Max ECL Western Blotting Substrates. For the single-base elongation- and ligation-based PCR amplification method (SELECT) to determine the m<sup>6</sup>A level, the protocol was followed as previously published (56).

### TMT Proteomics Sample Preparation, LC-MS, and Data Analysis

CLL and normal B cells were lysed with TEAB buffer supplemented with protease inhibitor and PMSF. Lysates (300 µg) were precipitated and digested to obtain peptides. TMT 10-plex labeling was performed according to the manual from Thermo Fisher Scientific, and peptides were fractionated via BPRP HPLC. An 1100 pump (Agilent) equipped with a degasser, and a photodiode array (PDA) detector (Thermo Fisher Scientific) was used. Peptides were subjected to a linear gradient from 3% to 25% acetonitrile in 0.125% formic acid using an Agilent 300 Extend-C18 column (Agilent) and were fractionated into a total of 96 fractions. Mass spectrometry was performed using an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) coupled to a Proxeon EASY-nLC 1000 liquid chromatography (LC) pump (Thermo Fisher Scientific). Peptides were detected (MS1) and quantified (MS3) in the Orbitrap. MS2 spectra were searched using the SEQUEST algorithm against a Uniprot composite database derived from the mouse proteome containing its reversed complement and known contaminants. Peptide spectral matches were filtered to a 1% false discovery rate (FDR) using the target-decoy strategy combined with linear discriminant analysis. The detected proteins were filtered to  $\alpha = 200$  and an isolation specificity of 0.5. Statistical analysis of the proteome was performed based on the normalized intensities of the TMT-reporter ions.

The peptide and protein abundance from TMT proteomics data was log<sub>2</sub>-transformed. Samples with Gaussian distribution of log<sub>2</sub>-transformed protein abundance were proceeded (Supplementary Fig. S2A). Additionally, bimodality coefficient and Hartigan's dip test were used to test the bimodal and significant skew (tailing) distribution (Supplementary Fig. S2B and S2C). Outlier samples were measured by pair-wise Pearson correlation between samples. "Noisy" proteins were removed if they were detected on less than 3 technical replicates. Mean protein intensity was calculated among technical replicates. Proteins detected in all samples were retained for downstream analysis.

For differential protein expression analysis between CLL and normal B cells, we used a previously established method with minor modification (57). Size factor according to total loading for each sample first was calculated to normalize the total amount of detected peptides. An internal control sample was introduced to all three batches to normalize the batch effect. At last, log<sub>2</sub>-transformed protein intensities were normalized by quantile normalization for all samples. Differentially expressed proteins were identified using LIMMA linear model methodology (58). Proteins with more than 1.5-fold change and FDR less than 0.05 were used as significance cutoff. Differential protein expression between CLL and normal B cells was shown in Supplementary Table S2.

To summarize the protein expression of splicing factors, we first grouped splicing factors according to the HUGO group 1518 (59). In addition, we also included hnRNP and SR proteins from the Spliceosome database (60). In total, this list contains 189 proteins. Eigengenes were calculated for each group of proteins using R package WGCNA (39). 153 out of 189 splicing factors were detected in our proteomics cohort, whereas 143 of them were consistently detected in the combined cohort. Correlation was calculated by using the Pearson correlation method and *P* values for multiple comparisons were adjusted using the Benjamini-Hochberg correction. A correlation edge with FDR less than 0.05 was defined as a significant correlation edge.

Positive correlation edge file and node information file were uploaded to the network visualization software Cytoscape for further visualization (61). Proteins involved in hnRNPs, SR proteins, SF3A/3B complex, and belonging to major spliceosome complex are color coded.

### RNA-seq Sample Preparation and Data Analysis

Libraries for RNA-seq were constructed using the Stranded Total RNA Prep with Ribo-Zero Plus Kit (Illumina) and sequenced on the NovaSeq S4 platform using a paired-end 150-bp mode. Reads were aligned to the human reference genome (GRCh38/hg38) using STAR with default parameters (62). Differential expression of mRNA analysis was performed using the DESeq2 R package (63). Genes with more than 2-fold change and FDR less than 0.05 were used as significance cutoff.

For RNA splicing analysis, we established a pipeline that integrates StringTie (33), LeafCutter (34), and rMATS (35) to maximally improve the power of detection of alternative splicing events. In brief, *de novo* transcripts were assembled using StringTie with default parameters. LeafCutter was used to detect novel exon boundaries. With the isoform annotation file downloaded from GENCODE (release 26), all isoform information was merged to generate a comprehensive isoform annotation file using custom R script as a reference file for rMATS. Percent spliced-in (PSI) value was calculated using rMATS for each splicing event. For differential splicing analysis of cell lines, we adopted the differential splicing analysis statistical model from rMATS and used absolute delta PSI more than 0.1 and FDR less than 0.05 as the significance cutoff. For differential splicing analysis in the patient cohort, we logistically transformed PSI values followed by a Student *t* test to obtain the significance. The significantly dysregulated splicing events were identified as absolute delta PSI more than 0.1 and *P* value less than 0.05 as significance cutoff. The significantly changed splicing type was identified by using a Q-Q plot. Expected *P* values were calculated by using a uniform distribution. The least-square linear fit to the lower 95th percentile of points with slope  $\lambda$  was calculated. Higher  $\lambda$  represents more dysregulation in the specific splicing type.

For splicing factor dosage-dependent alternative splicing (DDS) analysis, we utilized 19 samples with matched RNA-seq and proteomics data from our cohort. As for each protein, we performed RNA splicing analysis between five CLL samples with highest abundance (high group) and five with lowest abundance (low group). DDS score was calculated as the number of significant alternative splicing events normalized to the protein abundance difference between these two groups of samples. To calculate the *P* value of overlap between dosage-dependent alternative splicing events with CLL-related splicing events, we randomly selected 10 proteins 10,000 times. *P* value was defined as the probability of higher or the same overlap.

### m<sup>6</sup>A-Sensitive RNase Sequencing Sample Preparation and Data Analysis

Libraries were constructed according to the published protocol (40, 41) and sequenced on HiSeq 2500 platform using a paired-end 100-bp mode. Reads were aligned to the human reference genome (GRCh38/hg38) using STAR with default parameters (62). The cut efficiency matrix was calculated using mazter\_mine with minor modification (40). Only ACA sites in line with m<sup>6</sup>A motif DRACH were included for differential m<sup>6</sup>A analysis. Delta cut efficiency was calculated as the median cut efficiency of CLL samples subtracting with that of normal B cell samples. To reduce noise, only ACA sites detected in at least 8 samples were included. The statistical significance was calculated using the Wilcoxon rank sum test. Candidate m<sup>6</sup>A sites were defined as median cut efficiency less than or equal to 0.9 for CLL patients or normal B cells. m<sup>6</sup>A density was calculated using candidate m<sup>6</sup>A sites number divided by transcript length (Fig. 3H). m<sup>6</sup>A sites with absolute delta cut efficiency greater than 0.1 and *P* value less than 0.05 were used as significance cutoff.

### MeRIP Sequencing Sample Preparation and Data Analysis

Methylated RNA immunoprecipitation was processed by using the EpiMark N<sup>6</sup>-Methyladenosine Enrichment Kit (#E1610S, NEB). Purified RNA fragments were used for library construction using Stranded Total RNA Prep with Ribo-Zero Plus Kit (#20040529, Illumina) and sequenced with Illumina HiSeq 2500 platform using a paired-end 100-bp mode. Reads were aligned to the human reference genome (GRCh38/hg38) using STAR with default parameters (62). m<sup>6</sup>A peak calling, quantification, annotation, and differential m<sup>6</sup>A site analyses were performed using the R package exomePeak (64). *De novo* sequence motif discovery from m<sup>6</sup>A peaks was performed using HOMER with the command line “findMotifsGenome.pl diff\_peak.sort.bed hg38 motif -rna -size 200 -len 5,6,7 -noknown -p 16 -cache 6000” (65). Visualization files were generated using igvtools with parameters “-z 5 -w 10 -e 0.” Integrative Genomics Viewer (IGV) was used to visualize the distributions of the m<sup>6</sup>A peaks.

For m<sup>6</sup>A site prediction, m<sup>6</sup>A sites were first narrowed down by a sliding window of 30 nucleotides in each m<sup>6</sup>A peak region. m<sup>6</sup>A enrichment fold change and *P* values relative to RNA-seq reads were calculated using the statistical model from exomePeak (64). Peak-over-median score (POM) was calculated for each nucleotide in the highest enrichment sliding window. The nucleotide with the highest POM was defined as “peak summit.” All sequences following the DRACH motif inside each m<sup>6</sup>A peak region were marked. Sites in line with the DRACH motif and closest to the peak summit were defined as the m<sup>6</sup>A sites.

### Ribosome Profiling Sample Preparation and Data Analysis

Ribosome profiling was carried out following protocol from the Illumina TruSeq Ribo Profiler (Mammalian) kit (Illumina). The library was sequenced with the Illumina HiSeq 2500 platform using a single-end 50-bp mode. Reference sequences of rRNAs were downloaded from Ensembl (version 95; ref. 66). Reads aligned to rRNA sequences were removed from further analysis using the alignment tool Bowtie (67). Remaining reads were aligned to the human reference genome (GRCh38/hg38) using HISAT2 with default parameters (68). Multiple alignment reads were filtered out for downstream analysis. Quality control including reads length distribution, reading frames quantification, and 3 nucleotide periodicity were analyzed using the R package riboWaltz (69).

For the calculation of TE, the RPFs were counted in CDS regions with read length between 28 and 34 nt and not aligned to the first 15 and last 5 codons (70, 71). RPFs and RNA-seq reads for each transcript were normalized by the trimmed mean of *M* values (TMM) from edgeR (72, 73). TE was calculated as a ratio between normalized RPFs and normalized RNA-seq reads. The Fisher exact test was used to test the significance of the independent distribution of RPFs relative to RNA-seq reads. *P* values for multiple comparisons were adjusted using the Benjamini–Hochberg correction. Genes with more than 2-fold change and FDR less than 0.05 were used as significance cutoff.

For the ribosome recycling analysis, the exact P (peptidyl)-site of each RPFs reads was corrected by using R package riboWaltz according to the offset distance. Ribosome recycling rate was defined as Psite coverage of CDS regions (300-nt-long region upstream of stop codon region) divided to stop codon regions (45 nt upstream and 15 nt downstream of stop codons). Reads coverage was normalized to Reads Per Kilobase of transcript, per million mapped reads (RPKM). Differential ribosome recycling at stop codon regions was examined using the Fisher exact test to check if the RPF changes at stop codon regions are independent of those at upstream CDS regions. The odds ratio was calculated according to the recycling rate in *METTL3* KO cells compared with that in control cells. Genes with  $|\text{Log}_2 \text{ odds ratio}| \geq 0.5$  and *P* value less than 0.05 were used as significance cutoff.

For the ribosome pausing analysis, the pause score at each codon position was calculated based on a previous study (51). Ribosome A (aminoacyl) site information was extracted from riboWaltz output. To identify *METTL3*-dependent pausing sites, for each position of the

transcript, 2 × 2 contingency tables were created to perform a two-tailed Fisher exact test to compare the ratio of RPFs in the WT and KO fractions at a given position to the ratio at all other positions in that transcript (that is, the summed reads in each fraction for the entire transcript minus the position of interest), Benjamini–Hochberg correction for multiple hypothesis testing was used. *METTL3*-dependent pause sites were defined as  $\text{Log}_2$  odds ratio higher than 1 and FDR less than 0.05. To examine the codon enrichment of *METTL3*-dependent pause sites to other sites, *METTL3*-dependent pause sites with pause score >10 in *METTL3* KO cells and at least 5 RPFs were used. The odds ratio was calculated based on the codon frequency of *METTL3*-dependent pause sites to other sites. Two-tailed Fisher exact test followed by Benjamini–Hochberg correction to stratify the significance.

For the analysis of m<sup>6</sup>A-related ribosome pausing, the A-site coverage ± 60 nt around the m<sup>6</sup>A sites (described above) was calculated, and the regions with less than 20 reads coverage were excluded. The A-site coverage in each transcript at this range was normalized to mean A-site coverage of this region. The mean of normalized A-site coverage at each nucleotide was calculated and plotted out.

### Pathway Enrichment Analysis

Enrichment pathways were identified using two methods: GSEA based on databases from the Molecular Signature Database (MSigDB) including KEGG pathway, HALLMARK, and Reactome database. Gene sets with FDR <0.1 were considered significantly enriched pathways (74–76). A web-based pathway enrichment tool Metascape (77). Gene sets with FDR <0.05 were considered significantly enriched pathways. The pathway enrichment table related to each figure was shown in Supplementary Table S5.

### Statistical Methods

Survival curves were estimated using the Kaplan–Meier method, and the log-rank test was used to assess statistical significance for TTFT and OS by combining our cohort (*n* = 22) with the publicly available cohort (*n* = 91). Univariate Cox regression analysis was used to assess the prognostic impact of spliceosome complexes. Multivariate Cox regression analysis was used to assess the independent prognostic impact from 17p deletion status, *IGHV* mutational status, and eigengene of spliceosome complex for outcomes in the CLL cohort. Other statistical analyses were performed using one/two-tailed Student *t* test, two-way ANOVA, Kolmogorov–Smirnov test, Fisher exact test, and Mann–Whitney *U* test, all included in the figure legends.

### Data Availability

RNA-seq, MeRIP-seq, MAZTER-seq, and ribosome profiling raw and processed data have been deposited in the Gene-Expression Omnibus (GSE223731) and dbGAP (phs003191).

### Authors' Disclosures

G. Marcucci reports being a speaker for AbbVie; two clinical trials sponsored by Merck. A.V. Danilov reports grants and personal fees from Astra-Zeneca, Nurix Therapeutics, BeiGene, AbbVie, MEI, Incyte, and BMS, personal fees from Pharmacyclics and MingSight, grants from Bayer Oncology, Cyclacel, Tg Therapeutics, and SecuraBio outside the submitted work. J.R. Brown reports personal fees from AbbVie, Acerta/Astra-Zeneca, grants and personal fees from BeiGene, personal fees from Bristol Myers Squibb/Juno/Celgene, Catapult, Genentech/Roche, Grifols Worldwide Operations, Hutchmed, grants and personal fees from iOnctura, Janssen, Kite, grants and personal fees from Loxo/Lilly, MEI Pharma, personal fees from Morphosys AG, Novartis, Numab Therapeutics, Pfizer, Pharmacyclics, Rigel, grants from Gilead, Verastem/SecuraBio, and Tg Therapeutics outside the submitted work. T. Siddiqi reports Astra-Zeneca, BeiGene, BMS, Kite Pharma, AbbVie, and Celgene outside the submitted work. No disclosures were reported by the other authors.

## Authors' Contributions

**Y. Wu:** Data curation, validation, investigation, writing—original draft, writing—review and editing. **M. Jin:** Data curation, investigation, writing—original draft, writing—review and editing. **M. Fernandez:** Investigation. **K.L. Hart:** Investigation. **A. Liao:** Investigation. **X. Ge:** Methodology. **S.M. Fernandes:** Resources. **T. McDonald:** Resources. **Z. Chen:** Methodology. **D. Roth:** Methodology. **L.Y. Ghoda:** Resources. **G. Marcucci:** Resources. **M. Kalkum:** Methodology. **R.K. Pillai:** Resources. **A.V. Danilov:** Resources. **J.J. Li:** Methodology. **J. Chen:** Resources. **J.R. Brown:** Resources. **S.T. Rosen:** Resources. **T. Siddiqi:** Resources. **L. Wang:** Conceptualization, resources, data curation, supervision, funding acquisition, writing—original draft, project administration, writing—review and editing.

## Acknowledgments

The authors would like to thank Drs. Renjang Lin and Nora Heisterkamp (City of Hope) for critical review of our manuscript. We appreciated the kind help from Jennifer Zankich Dettor, Elena Pulkkinen, Ashish Streatfield, and Kelly Synold from City of Hope Hematopoietic Tissue Biorepository for extracting and compiling clinical data for annotation of the specimens. This work was supported by a Startup package from City of Hope (L. Wang), grants from the NIH (NCI) R01CA240910 and R01CA21623 (to L. Wang). Research work related to sample collection in the Hematopoietic Tissue Biorepository Core supported by the NIH (P30CA033572).

The publication costs of this article were defrayed in part by the payment of publication fees. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

## Note

Supplementary data for this article are available at Blood Cancer Discovery Online (<https://bloodcancerdiscov.aacrjournals.org/>).

Received September 23, 2022; revised January 30, 2023; accepted March 10, 2023; published first April 17, 2023.

## REFERENCES

- Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009;136:701–18.
- Moore MJ, Proudfoot NJ. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* 2009;136:688–700.
- Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 2008;40:1413–5.
- Harper SJ, Bates DO. VEGF-A splicing: the key to anti-angiogenic therapeutics? *Nat Rev Cancer* 2008;8:880–7.
- Venables JP. Unbalanced alternative splicing and its significance in cancer. *Bioessays* 2006;28:378–86.
- Amin EM, Oltean S, Hua J, Gammons MV, Hamdollah-Zadeh M, Welsh GI, et al. WT1 mutants reveal SRPK1 to be a downstream angiogenesis target by altering VEGF splicing. *Cancer Cell* 2011;20:768–80.
- Kim E, Ilagan JO, Liang Y, Daubner GM, Lee SC, Ramakrishnan A, et al. SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* 2015;27:617–30.
- Obeng EA, Chappell RJ, Seiler M, Chen MC, Campagna DR, Schmidt PJ, et al. Physiologic expression of Sf3b1(K700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. *Cancer Cell* 2016;30:404–17.
- Shirai CL, Ley JN, White BS, Kim S, Tibbitts J, Shao J, et al. Mutant U2AF1 expression alters hematopoiesis and Pre-mRNA splicing in vivo. *Cancer Cell* 2015;27:631–43.
- Yin S, Gambe RG, Sun J, Martinez AZ, Cartun ZJ, Regis FFD, et al. A murine model of chronic lymphocytic leukemia based on B cell-restricted expression of Sf3b1 mutation and atm deletion. *Cancer Cell* 2019;35:283–96.
- Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med* 2005;352:804–15.
- Ten Hacken E, Valentin R, Regis FFD, Sun J, Yin S, Werner L, et al. Splicing modulation sensitizes chronic lymphocytic leukemia cells to venetoclax by remodeling mitochondrial apoptotic dependencies. *JCI Insight* 2018;3:e121438.
- Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* 2011;365:2497–506.
- Rossi D, Bruscazzin A, Spina V, Rasi S, Khiabani H, Messina M, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* 2011;118:6904–8.
- Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* 2011;44:47–52.
- Shuai S, Suzuki H, Diaz-Navarro A, Nadeu F, Kumar SA, Gutierrez-Fernandez A, et al. The U1 spliceosomal RNA is recurrently mutated in multiple cancers. *Nature* 2019;574:712–6.
- Wang L, Brooks AN, Fan J, Wan Y, Gambe R, Li S, et al. Transcriptomic characterization of SF3B1 mutation reveals its pleiotropic effects in chronic lymphocytic leukemia. *Cancer Cell* 2016;30:750–63.
- Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature* 2015;526:525–30.
- Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* 2013;49:18–29.
- Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* 2014;10:93–95.
- Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res* 2014;24:177–89.
- Patil DP, Chen CK, Pickering BF, Chow A, Jackson C, Guttman M, et al. m(6A) RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 2016;537:369–73.
- Vu LP, Pickering BF, Cheng Y, Zaccara S, Nguyen D, Minuesa G, et al. The N(6)-methyladenosine (m(6A))-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat Med* 2017;23:1369–76.
- Roundtree IA, Luo GZ, Zhang Z, Wang X, Zhou T, Cui Y, et al. YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife* 2017;6:e31311.
- Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m(6A) RNA methylation. *Nat Rev Genet* 2014;15:293–306.
- Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* 2015;161:1388–99.
- Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m(6A) reader YTHDC1 regulates mRNA splicing. *Mol Cell* 2016;61:507–19.
- Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, et al. Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol* 2018;20:285–95.
- Mao Y, Dong L, Liu XM, Guo J, Ma H, Shen B, et al. m(6A) in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun* 2019;10:5332.
- Rossello-Tortella M, Ferrer G, Esteller M. Epitranscriptomics in hematopoiesis and hematologic malignancies. *Blood Cancer Discov* 2020;1:26–31.
- Bueno-Costa A, Pineyro D, Garcia-Prieto CA, Ortiz-Barahona V, Martinez-Verbo L, Webster NA, et al. Remodeling of the m(6A) RNA landscape in the conversion of acute lymphoblastic leukemia cells to macrophages. *Leukemia* 2022;36:2121–4.
- Barbieri I, Tzelepis K, Pandolfini L, Shi J, Millan-Zambrano G, Robson SC, et al. Promoter-bound METTL3 maintains myeloid leukaemia by m(6A)-dependent translation control. *Nature* 2017;552:126–31.

33. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015;33:290–5.
34. Li YL, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, et al. Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet* 2018;50:151–8.
35. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A* 2014;111:E5593–5601.
36. Meier-Abt F, Lu J, Cannizzaro E, Pohly MF, Kummer S, Pfammatter S, et al. The protein landscape of chronic lymphocytic leukemia. *Blood* 2021;138:2514–25.
37. Johnston HE, Carter MJ, Larrayoz M, Clarke J, Garbis SD, Oscier D, et al. Proteomics profiling of CLL versus healthy B-cells identifies putative therapeutic targets and a subtype-independent signature of spliceosome dysregulation. *Mol Cell Proteomics* 2018;17:776–91.
38. Chuang HY, Rassenti L, Salcedo M, Licon K, Kohlmann A, Haferlach T, et al. Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression. *Blood* 2012;120:2639–49.
39. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf* 2008;9:559.
40. Garcia-Campos MA, Edelleit S, Toth U, Safra M, Shachar R, Viukov S, et al. Deciphering the “m(6A code)” via antibody-independent quantitative profiling. *Cell* 2019;178:731–47.
41. Zhang Z, Chen LQ, Zhao YL, Yang CG, Roundtree IA, Zhang Z, et al. Single-base mapping of m(6A) by an antibody-independent method. *Sci Adv* 2019;5:eaax0250.
42. Zeng C, Huang W, Li Y, Weng H. Roles of METTL3 in cancer: mechanisms and therapeutic targeting. *J Hematol Oncol* 2020;13:117.
43. Yankova E, Blackaby W, Albertella M, Rak J, De Braekeleer E, Tsagkogeorga G, et al. Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature* 2021;593:597–601.
44. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. *Cell* 2017;170:564–76.
45. Cui Q, Shi H, Ye P, Li L, Qu Q, Sun G, et al. m(6A) RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. *Cell Rep* 2017;18:2622–34.
46. Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. *Cell Res* 2017;27:315–28.
47. Choe J, Lin S, Zhang W, Liu Q, Wang L, Ramirez-Moya J, et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* 2018;561:556–60.
48. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m(6A) methyltransferase METTL3 promotes translation in human cancer cells. *Mol Cell* 2016;62:335–45.
49. Imataka H, Gradi A, Sonenberg N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J* 1998;17:7480–9.
50. Jeong KW, Indrisiunaite G, Prabhakar A, Puglisi JD, Ehrenberg M. N<sup>6</sup>-Methyladenosines in mRNAs reduce the accuracy of codon reading by transfer RNAs and peptide release factors. *Nucleic Acids Res* 2021;49:2684–99.
51. Stein KC, Morales-Polanco F, van der Lienden J, Rainbolt TK, Frydman J. Ageing exacerbates ribosome pausing to disrupt cotranslational proteostasis. *Nature* 2022;601:637–42.
52. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet* 2011;44:53–57.
53. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011;478:64–69.
54. Furney SJ, Pedersen M, Gentien D, Dumont AG, Rapinat A, Desjardins L, et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov* 2013;3:1122–9.
55. Liu J, Eckert MA, Harada BT, Liu SM, Lu Z, Yu K, et al. m(6A) mRNA methylation regulates AKT activity to promote the proliferation and tumorigenicity of endometrial cancer. *Nat Cell Biol* 2018;20:1074–83.
56. Xiao Y, Wang Y, Tang Q, Wei L, Zhang X, Jia G. An elongation- and ligation-based qPCR amplification method for the radiolabeling-free detection of locus-specific N(6)-methyladenosine modification. *Angew Chem Int Ed Engl* 2018;57:15995–6000.
57. Efstathiou G, Antonakis AN, Pavlopoulos GA, Theodosiou T, Divanach P, Trudgian DC, et al. ProteoSign: an end-user online differential proteomics statistical analysis platform. *Nucleic Acids Res* 2017;45:W300–6.
58. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
59. Tweedie S, Braschi B, Gray K, Jones TEM, Seal RL, Yates B, et al. GeneNames.org: the HGNC and VGNC resources in 2021. *Nucleic Acids Res* 2021;49:D939–46.
60. Cvitkovic I, Jurica MS. Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic Acids Res* 2013;41:D132–41.
61. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498–504.
62. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21.
63. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
64. Meng J, Lu Z, Liu H, Zhang L, Zhang S, Chen Y, et al. A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods* 2014;69:274–81.
65. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;38:576–89.
66. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, et al. The Ensembl genome database project. *Nucleic Acids Res* 2002;30:38–41.
67. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25.
68. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 2019;37:907–15.
69. Lauria F, Tebaldi T, Bernabo P, Groen EJM, Gillingwater TH, Viero G. riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* 2018;14:e1006169.
70. Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouar GJ, Jackson SE, et al. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep* 2014;8:1365–79.
71. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 2011;147:789–802.
72. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010;11:R25.
73. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–40.
74. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
75. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017;45:D353–61.
76. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417–25.
77. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level data sets. *Nat Commun* 2019;10:1523.