The eukaryotic translation initiation factor eIF4E reprogrammes the splicing machinery and drives alternative splicing

Mehdi Ghram[1,*], Gavin Morris[1,*], Biljana Culjkovic-Kraljacic[1,*], Patrick Gendron[2], Lucy Skrabanek[3], Maria Victoria Revuelta[4], Leandro Cerchietti[4], MonicaL Guzman[4], Katherine LB Borden[1,#].

* These authors contributed equally to this work

To whom correspondence should be addressed: Katherine.borden@umontreal.ca

[1] Institute for Research in Immunology and Cancer, Department of Pathology and Cell Biology, University of Montreal, Montreal, QC, H3C 3J7, Canada.

[2] Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, Montreal, Quebec H3C 3J7, Canada.

[3] Department of Physiology and Biophysics, Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY 10065, USA; Applied Bioinformatics Core, Weill Cornell Medicine, New York, NY 10065, USA.

[4] Division of Hematology/Oncology, Department of Medicine, Weill Cornell Medicine, Cornell University, New York, NY 10021, USA.

Abstract

Aberrant RNA splicing contributes to the pathogenesis of many malignancies including Acute Myeloid Leukemia (AML). While mutation is the best described mechanism underpinning aberrant splicing, recent studies show that predictions based on mutations alone likely underestimate the extent of this dysregulation[1] . Here, we show that elevation of the eukaryotic translation initiation factor eIF4E reprogrammes splicing of nearly a thousand RNAs in model cell lines. In AML patient specimens which did not harbour known splice factor mutations, ~4000 transcripts were differentially spliced based on eIF4E levels and this was associated with poor prognosis. Inhibition of eIF4E in cell lines reverted the eIF4E-dependent splice events examined. Splicing targets of eIF4E act in biological processes consistent with its role in malignancy. This altered splicing program likely arose from eIF4E-dependnet increases in the production of many components of the spliceosome including SF3B1 and U2AF1 which are frequently mutated in AML. Notably, eIF4E did not drive mutation of these factors, only their production. eIF4E also physically associated with many splice factors including SF3B1, U2AF1, and UsnRNAs. Importantly, many eIF4E-dependent splice events differed from those arising from SF3B1 mutation, and were more extensive highlighting that these splicing profiles arise from distinct mechanisms. In all, our studies provide a paradigm for how dysregulation of a single factor, eIF4E, can alter splicing.

eIF4E amplifies transcriptional signals for a subset of RNAs by promoting their nuclear export and increasing their translation efficiency[2-5]. More recent studies indicate that eIF4E also influences m^7G capping and cleavage and polyadenylation (CPA) of selected RNAs[4,6] suggesting it plays broader roles in mRNA maturation. RNA splicing usually occurs between the capping and CPA maturation steps[7], so we reasoned that eIF4E also played a role in splicing; and in this way, eIF4E would not only amplify transcriptional signals but also change the coding information within the RNA message. Previous studies revealed that eIF4E increased the protein levels of capping and CPA factors[4,6]. Thus, as a first step, we investigated if eIF4E could similarly influence the production of splice factors (SF) which comprise the spliceosome. The spliceosome is an assembly of ~150 SFs and 5 uridine-rich small nuclear UsnRNAs (U1, U2, U4, U5 and U6 snRNAs)[7]. Here, we monitored the impact of eIF4E on SFs derived from the major complexes and included those associated with cancer e.g. SF3B1 and U2AF1[8]. For this purpose, we employed U2OS cells since eIF4E plays well-established roles in capping, CPA, nuclear RNA export and translation in these cells[4,6,9-13]. We generated 3 stable 2FLAG-eIF4E or Vector control cell lines as previously descrcibed[4]. Notably, inspection of RNA-sequencing data revealed that neither the parent U2OS cells nor 2FLAG-eIF4E or Vector stable cell lines harbour SF mutations typically associated with cancer e.g. in SF3B1, U2AF1 or SRSF2[8] (GSE158728). Finally, in U2OS cells we observe endogenous eIF4E in the nucleus in bodies and diffusely throughout the nucleoplasm as well as in the cytoplasm consistent with previous studies (Fig 1A). Additionally, the extent of eIF4E elevation in these cells was physiologically relevant; 2FLAG-eIF4E U2OS cells had a ~3-fold elevation in eIF4E protein levels relative to Vector controls (Fig 1B). By comparison, AML specimens had up to ~8-fold elevation of eIF4E protein levels relative to healthy tissues[14,15]. To assess SF protein levels, we carried out western blot analyses of total cell lysates. SF3B1, U2AF1, U2AF2, PRPF6, PRPF8, PRPF19, PRPF31, SNRNP200 protein levels were elevated in 2FLAG-eIF4E cells relative to Vector controls (Fig 1A). Cyclin D1 and Mcl1 served as positive controls while β-actin provides a negative control as it is not an eIF4E target[4] (Fig 1A). We next examined whether depletion of endogenous eIF4E reduced levels of these SFs using CRISPR-eIF4E or CRISPR-Ctrl U2OS cell lines as described[4]. CRISPR-Ctrl cell lines were generated using guide RNAs to *Galaxidae* coral Azami-green transcripts[4]. CRISPR-eIF4E cell lines were heterozygous for eIF4E as its complete deletion was lethal. As expected, CRISPR-eIF4E cells had reduced levels of PRPF6, PRPF8, PRPF19, SF3B1, SNRNP200, and U2AF2 proteins relative to CRISPR-Ctrl

cells (Fig 1C). These findings indicated that endogenous eIF4E impacted production of the splicing machinery, and thus, eIF4E's effects on SFs were not limited to circumstances characterized by eIF4E elevation. In all, we demonstrate that SF production can be driven in an eIF4E-dependent manner. SFs targeted by eIF4E are found in a variety of spliceosome complexes (Fig1D), and included SFs that are often mutated in AML and other cancers e.g. SF3B1 and U2AF1 [1,8,16]. Thus, the impact of eIF4E does not appear restricted to any specific spliceosome complex, but rather, is positioned to influence several splicing steps. Further, SF3B1 and U2AF1 which are mutated in many cancers with dysregulated splicing, can also be dysregulated through increased production via eIF4E in the absence of mutation.

We assessed whether at least a subset of these splicing factors was directly regulated by eIF4E. eIF4E can drive production of proteins through increased cytoplasmic availability of transcripts via increased nuclear export activity and/or by elevated numbers of ribosomes per transcript thereby increasing translation efficiency. We observed many RNA processing factors were nuclear export targets and thus focussed on this here[4,6]. First, we carried out eIF4E RNA immunoprecipitations (RIPs) from nuclear lysates since these interactions predict sensitivity to eIF4E-dependent RNA export; by contrast, all examined RNAs bind to eIF4E in the cytoplasm, whether they are translation targets of eIF4E or not[4,6,12,13]. Using nuclear lysates from U2OS cells, we found that endogenous nuclear eIF4E RIPs with *SF3B1, SNRNP200, PRPF6, PRPF8, PRPF31, U2AF1, U2AF2* (Fig 2A) were enriched by ~2-4 fold versus input. Negative control RNAs *ACTB*, *GAPDH*, *POLR2A* RNAs and *18*S rRNA were not found in the RIPs while positive controls *MCL1*, and *MYC* were present consistent with previous studies[4,6,12,13]. Given this physical interaction, we examined whether these RNAs were targets of eIF4E-dependent nuclear export by monitoring whether their RNA export was increased in stable 2FLAG-eIF4E cells relative to Vector controls, using the three different cell lines as above. For RNA export assays, cells were fractionated into nuclear and cytoplasmic compartments and RNAs quantified in each fraction using RT-qPCR. As a proof-of-principle, we monitored *SF3B1, U2AF1, U2AF3, SNRNP200, PRPF6, PRPF8* and *PRPF31* which had ~2 fold increased in cytoplasmic/nuclear ratios upon eIF4E overexpression indicating their nuclear export was increased upon eIF4E overexpression (Supp Fig 2A). *MYC* and *MCL1* serve as a positive control and were elevated while *GAPDH* and *POLR2A* were negative controls and unchanged as expected (normalized to *ACTB*). Fraction quality was assessed by semi-

quantitative PCR monitoring U6snRNA and $tRNA^{Met}$ for controls for the nuclear and cytoplasmic fractions respectively (Supp Fig 2B)[11].Typical of eIF4E targets, total levels of the corresponding RNAs were not altered by eIF4E overexpression as observed by RT-qPCR (Supp Fig 2C). In all, eIF4E physically interacts with, and promotes the nuclear export of, many SF-encoding RNAs providing a mechanism for eIF4E-mediated elevation of these factors.

Given the impact of eIF4E on the production of SFs in U2OS cells, we investigated whether eIF4E drove their production in an High-eIF4E cancer opting to study High-eIF4E AML in which eIF4E has been targeted with ribavirin in early phase clinical trials leading to objective responses including remissions[14,15]. In these High-eIF4E patients, eIF4E is typically elevated ~3-8 fold and predominantly localized to the nucleus and characterized by elevated eIF4E nuclear RNA export activity[14,15,17] (e.g. Supp Fig1A). Further supporting this choice of indication, AML is often associated with dysregulated splicing, which is typically attributed to mutation in specific SFs[1,8,16,18,19]. We used two experimental systems to explore the impact of eIF4E on SFs in AML: the NOMO-1 AML cell line, and primary specimens from AML patients. Given the close link between dysregulated splicing and mutations in *U2AF*, *SF3B1* or *SRSF2* in AML[1,8,16,18], we inspected our RNA-Seq data to ensure that NOMO-1 (Supp Fig 1C) and the primary AML specimens (GSE67040) examined did not harbour these mutations. In NOMO-1 cells, eIF4E levels are very similar to those in healthy volunteers or Normal-eIF4E AML specimens (Supp Fig 1B); thus, these cells provided a unique opportunity to engineer High-eIF4E AML cells, and thus, to genetically dissect the role of eIF4E in an AML context and to compare to our U2OS system above. eIF4E or Vector NOMO-1 cell lines were produced as pools. NOMO-1 eIF4E cells produced eIF4E protein levels and localization similar to High-eIF4E AML specimens (Fig 1E, Supp Fig 1A bottom panel). We note that eIF4E overexpression in NOMO-1 cells did not lead to mutations found in AML in the above SFs. We observed elevation of PRPF6, PRPF8, PRP31, SF3B1, SNRNP200, U2AF1, and U2AF2 proteins in eIF4E NOMO-1 relative to Vector NOMO-1 cells (Fig 1E and 1F). Further, positive controls Myc, Mcl1, and CCND1 were elevated while the negative control β-actin was unchanged. To assess the impact of eIF4E inhibition, we monitored the impact of the eIF4E-inhibitor ribavirin, which has been used to target eIF4E in early phase clinical trials[14,15,20,21]. We observed reduced SF levels upon eIF4E inhibition with clinically achievable doses of ribavirin in both conditions relative to vehicle treated controls consistent with

ribavirin targeting endogenous and overexpressed eIF4E (Fig 1F). Thus, overexpression of eIF4E in AML cells led to changes in the splicing machinery similar to that observed in U2OS cell lines suggesting that eIF4E can dysregulate SF production in multiple contexts. Further, its effects on SF production can be reversed by a clinically used eIF4E inhibitor (Fig 1F).

Next, we examined whether eIF4E status correlated with levels of splice factors in primary AML specimens. We isolated de-identified AML patient blasts using FACS[14,15,22,23] and compared these with bone marrow mononuclear cells or CD34+ cells from healthy volunteers. Analysis of extracted proteins revealed that High-eIF4E AML specimens had elevated levels of PRPF6, PRPF8, PRPF19, SF3B1, U2AF1, U2AF2 and SNRNP200 relative to healthy volunteers or Normal-eIF4E AMLs (Fig 1G). Longer exposures of the western blots revealed that all proteins were present in the normal specimens, the exposure shown highlights the relative differences between High-eIF4E AML and the others. Thus, elevation of these SFs was not a general feature of AML patient specimens but correlated with eIF4E levels consistent with our studies in U2OS and NOMO-1 cells. While eIF4E is correlated with elevated levels of these factors, it is likely that there are other mechanisms that can drive elevation of SFs, this would be an interesting avenue of future study.

Previously, studies showed that eIF4E interacted with RNA processing machinery involved in capping, CPA, export and translation[4,6,9,10,13,24]. Thus, we examined whether eIF4E similarly interacted with SFs; providing an additional means by which eIF4E could impact splicing[4,6,10]. In the nuclear eIF4E RIPs described above, we observed that endogenous eIF4E immunoprecipitated with U1, U2, U4, U5 and U6 snRNAs relative to inputs with a ~3-10 fold enrichment in nuclear lysates from U2OS cells (Fig 2A), as well as in NOMO-1 cells (supp Fig 2D). Endogenous eIF4E did not associate with negative controls (e.g. *GAPDH, ACTB etc*) indicating these RIPs were specific (Fig 2A & Supp Fig 2D). Additionally, levels of the UsnRNAs were not altered upon eIF4E overexpression in U2OS (Supp Fig 2E). Given UsnRNAs play both structural and catalytic roles in the spliceosome[7], we investigated whether eIF4E physically associated with protein components of the spliceosome. We found that eIF4E also immunoprecipitated with protein components of the spliceosome including PRPF6, PRPF8, PRPF19, PRPF31, SF3B1, U2AF1, U2AF2 and SNRNP200 proteins in U2OS (Fig 2B) and NOMO-1 cells (Fig 2C). Notably, these

factors were also elevated upon eIF4E overexpression (Fig 1) and except for PRPF19, their corresponding RNAs bound to eIF4E in nuclear RIPs (Fig 2A and Supp Fig 2D). In all, eIF4E binds to several SFs and UsnRNAs of the major spliceosome, and elevated SF protein levels in both AML and U2OS cell contexts. Thus, eIF4E is positioned to impact on splicing in different cellular contexts, suggesting this could be a wide-ranging property of eIF4E.

Given these findings, we monitiored the impact of eIF4E on splicing. RNAs from total cell lysates were isolated from the three 2FLAG-eIF4E and Vector U2OS cell lines and subjected to RNA-Seq. Differences in splicing profiles were quantified using replicate Multivariate Analysis of Transcript Splicing (rMATS)[25]. rMATs calculates the "inclusion level differences" for splicing events such as exon skipping (ES), inclusion of mutually exclusive exons (MXE), intron retention (IR) and alternative splice site usage (Fig 3A). For example, it calculates the extent an exon is included, i.e. not skipped, in Vector versus 2FLAG-eIF4E cells. In this case, positive values indicate an exon is more included in Vector, and negative values that this exon is more included in eIF4E-overexpressing cells. Here, a value of +1 for SE events indicated that the relevant exon is 100% included in Vector cells and 0% in 2FLAG-eIF4E cells. Many of the same events (129; Supp Table1) were also observed using EBSeq, which maps exons; but in contrast to rMATS, EBSeq does not provide splice site event information. Thus, we focussed on rMATs. Finally, as stated above, we note that eIF4E-overexpression did not induce mutation in *SF3B1*, *U2AF1* or *SF3B1*, as observed by our RNA-Seq data (GSE158728). We note that RNAs were ribo-depleted to remove ribosomal RNA. Ribodepletion has the advantage over polyA selection that RNAs with short polyA tails or with no polyA tails can be captured in contrast to polyA selection strategies. Given eIF4E has impacts on production of CPA factors[6] and impacts on APA for hundreds of transcripts (our unpublished observation), we deemed this a necessary consideration.

Using rMATs, we observed that ~890 splicing events for ~760 transcripts were altered (FDR-adjusted p-value <0.15; Inclusion differences of >0.05 or <-0.05) or using more stringent cutoffs (p-value <0.1, Inclusion level differences of >0.1 or <-0.1) we observed 555 splicing events for 493 transcripts (Fig 3B, Supp Table2). In the RNA-Seq experiments, a total of 5738 annotated transcripts were detected with >10 TPM indicating that ~15% of annotated transcripts were differentially spliced in an eIF4E-dependent manner. SE was the most frequent event (~55%)

followed by MXE events (~30%) (Supp Table2; Fig 3B). More rarely, Intron Retention (IR) and altered 3'/5' splice sites usage were impacted. The majority of transcripts affected were coding RNAs, but some non-coding RNAs were also impacted (Supp Table 2). Hierarchical clustering analyses based on the inclusion levels across replicates revealed that all types of splicing events (e.g. SE, MXE, IR etc) segregated solely based on eIF4E levels (Fig 3C, Supp Fig 3A). Further, the clustering indicated replicates within each group were highly similar. To assess if eIF4E overexpression was associated with repression or promotion of splicing, we analyzed inclusion level differences of the spliced events. Overall, these were roughly equally positive or negative (45% versus 50 % respectively) suggesting a splicing reprogramming of a specific subset of RNAs. In all, these findings support a role for eIF4E as a mediator of alternative splicing (AS) for a subgroup of transcripts; importantly, our studies indicate that eIF4E did not elicit global changes to splicing when overexpressed in U2OS cell lines. Notably, eIF4E overexpression did not lead to a global reprogramming of the transcriptome, with only 402 differentially expressed genes in Vector versus 2FLAG-eIF4E U2OS cells[4]. More importantly, only ~2% of the differentially expressed transcripts (13 targets) undergo AS upon eIF4E overexpression, suggesting that eIF4E-related AS is not correlated with RNA levels in these cells (Supp Fig 3B, Supp Table 3).

To understand the biological impact of eIF4E-dependent alterations to splicing, we carried out pathway and process enrichment analyses. Enrichment of GO terms for all significant events indicated that eIF4E influenced pathways that could support its oncogenic phenotype (Fig 3D, Supp Table 4). Top GO categories include Cell Cycle, Membrane trafficking, DNA repair and Microtubule cytoskeletal organization. Protein-protein interaction enrichment analysis are consistent with the GO terms with the top hits being Cell Cycle, RNA metabolism and Membrane Trafficking (Fig 3E, Supp Table 5). The enrichment of RNA metabolism is particularly interesting, and suggestive of a positive feedback loop resulting from eIF4E dysregulation and is consistent with its known multiple functions.

We next examined the relevance of eIF4E-dependent splicing to human diseases characterized by dysregulated eIF4E focussing on AML (Fig 4) given the extensive studies into dysregulated splicing there and because targeting eIF4E with ribavirin in AML provided clinical benefit including remissions in a subset of AML patients in early stage clinical trials[14,15]. As a first step,

we analysed RNA-Seq data from >450 de-identified AML patient specimens and 17 specimens of normal CD34+ cells derived from human cord blood (www.leucegene.ca). Dividing the entire leucegene cohort into two groups (above and below median RNA expression of eIF4E), we observed a substantial, and significant, reduction in surival for the High-eIF4E group (Supp Fig 4A). For further analysis, we pre-screened specimens by inspection of the RNA-Seq data to ensure they did not harbour the major mutations reported in AML *i.e*. those in *U2AF1*,*SRSF2* or *SF3B1*. We used the normal CD34+ cells to benchmark eIF4E levels in order to categorize AML specimens into High- or Normal-eIF4E groups (Fig 4A). For rMATS analysis, we selected 10 AML specimens with the highest eIF4E levels and 11 with normal eIF4E levels i.e. overlapping with or below levels in normal CD34+ cells (Fig 4B). Corresponding clinical data revealed that the 10 High-eIF4E AML patients had drastically reduced overall survival (median survival 136 days) when compared to the 11 patients with lowest eIF4E levels in AML (1396 days) (Fig 4C).

We compared the splicing profiles of these High-eIF4E and Normal-eIF4E specimens using rMATS. We observed differences in ~1600 splicing events impacting ~1500 RNAs (FDR-adjusted p-value <0.1, absolute value of inclusion level differences of >0.1) and if the threshhold is lowered, ~12000 events for ~4000 transripts (absolute inclusion value difference of >.05 and FDR-adjusted p-value<.15; Fig 4D, Supp Table 6). In terms of transcript levels, 8813 genes had an average TPM value >10. Thus alternative splicing of ~15-50% of detected transcripts corresponds with eIF4E levels, depending on the cutoffs used. Of these events, ~75% arose due to MXE, ~20% SE and the remainder were attribuable to retained introns or altered 3'/5' splice site usage (Fig 4D). As for U2OS, High-eIF4E levels correlated with equal number of promoted and repressed splicing events. The maximum magnitude of the average inclusion level differences across patients was ~0.4 (corresponding to ~40% change in population) versus ~0.8 (~80%) for U2OS cells (Fig 4D versus Fig 3B). The reduction in magnitude in primary specimens likely arises due to heterogeneity amongst patients' transcriptomes[1,26]. As in U2OS cells, we did not generally observe any correlation between AS reprogramming and changes in RNA levels, since only ~7% of the alternatively spliced targets were differentially expressed (Supp Fig 4C, Supp Table 7, 8).

To establish whether eIF4E levels were the factor most significantly influencing splicing, we conducted hierarchal clustering analysis. There was a very tight correlation between eIF4E status

in AML patients and MXE events. Specifically, 20/21 AML specimens clustered based on eIF4E levels when monitoring MXE events (Fig 4E). Similar analysis of RI events revealed groups clustered on eIF4E levels with the exception of two High-eIF4E patients that clustered with Normal-eIF4E patients (Supp Fig 4B). Finally, analyses of SE events indicated two distinct clusters of High-eIF4E AML patients, one cluster was more similar to Normal-eIF4E patients (Supp Fig 4B). Interestingly, we note that unlike the overexpression of eIF4E in U2OS cells where only 5 SFs were differentially spliced (Supp Table 2), eIF4E elevation in AML patients correlates with splicing reprogramming of a wide array of spliceosome components (~20 targets), including *PRPF3, PRPF8, PRPF4B, PRPF39, PRPF31, PRPF6, PRPF40A, SF1, SRSF10, SF3B3, SF3B2, SRSF5* and *U2AF1* (Supp Table 6). Thus, eIF4E appears to have a broader impact on the splicing of SFs in the context of AML. Similar to U2OS cells, we observed that the top GO enrichment terms for eIF4E-dependent MXE targets were highly similar to those in the U2OS datasets: Cell Cycle, DNA repair, Membrane trafficking and RNA metabolism (Supp Fig 4F, Supp Table 9). Protein interaction networks was enriched in Mitotic centrosomes, Membrane trafficking and Notch signalling(Fig 4G, Supp Table 10).

We reasoned that identification of splicing events conserved across disparate cell types would reveal pan-cancer core networks that underpin, at least in part, eIF4E's oncogenic effects (Fig 3 versus Fig 4). A comparison of all eIF4E-dependent splicing targets in U2OS and AML datasets evince a core set of ~450 common RNA targets (Fig 5A, Supp Table 11); absolute value of inclusion level difference>.05; FDR-adjusted p-value <.15). Indeed, 60% of targets identifed in the U2OS cells were in common with the AML targets. In terms of transcript expression, AML specimens and U2OS cells had 4705 transcripts in common (>10TPM) and thus about 10% of transcripts that were in common were also eIF4E-dependent splicing targets. Note that core targets were not necessarily characterized by the same splice event. GO term analysis of RNA targets revealed that the RNAs enriched between U2OS and AML were a general reflection of those identified in the analysis of either cell-type alone. In all, analyses identifed proteins involved in Cell Cycle, DNA Repair, Membrane Trafficking, Chromatin Modification and RNA metabolism for GO terms (Fig 5B, Supp Table 12) and DNA repair, adaptive immune system, MYC activation and others for protein-protein interaction networks (Fig 5C, Supp Table 13).

Next, we validated several of these core splicing targets in in U2OS and for AML used NOMO-1 cells (Fig 5D and 5E). We selected events of biological interest to eIF4E's established oncogenic activities focussing on transcripts from the different protein interaction networks but that were part of the core set of overlapping targets in AML and U2OS cells and further we assessed if ribavirin treatment could revert these events thereby converting to Vector. We also prioritized targets with no significant changes to their RNA levels as a function of eIF4E expression level. Using RT-qPCR, we validated splicing events identified from the rMATs analysis for the following targets: *IL1B*, *IL4R*, *IkBkB*, *MAPK8IP3* and *FAAP20* in U2OS cells and *IL1B*, *IL4R*, *IkBkB* and *FAAP20* in NOMO-1 cells. Interestingly, the event that we validated in U2OS cells for *IkBkB* was identified in the AML patients (skipping exon 2) rather than the one predicted by rMATs in the U2OS cells (skipping exon 21) confirming the known limitations of rMATS despite its high efficacy and reproducibility when compared to other event-based AS analysis programs[27]. We observed in both cell lines that ribavirin treatment reverted the splicing events in eIF4E overexpressing cells *i.e.* that the splice variants were more similar to Vector cells observed after treatment (Fig 5D and E).

Finally, we examined whether eIF4E-dependent splicing targets were characterized by any features that could provide mechanistic insights into RNA target selection. We monitored exon and intron length, GC content of introns, exon position within the RNA and splice site usage, each as a function of splicing event as well as of positive or negative inclusion level differences. We selected the strongest splice targets in order to increase the likelihood of identifying commonalities and thus, interrogated both AML and U2OS datasets using the threshold of >0.1 for absolute value of inclusion level differences and FDR-adjusted p-value <0.1. We observed that longer introns were associated with eIF4E-dependent MXE and SE splicing events (Supp Fig 5C, Supp Fig 6C). Introns were up to ~9000-15000 bp in average length for eIF4E-dependent events in U2OS cells and ~6000-9500 bp in AML specimens relative to average intron length for all introns which was ~4000-5000 bp (Supp Table 14). We observed a trend towards higher GC content for High-eIF4E AML specimens for A3SS, A5SS, and SE events (Supp Fig 6E) and for A5SS and MXE events in 2FLAG-eIF4E relative to Vector controls (Supp Fig 5E). We did not observe significant alterations in splice site usage in either cell context (Supp Fig 5D, Supp Fig 6D); consistent with the low number of these events identified by rMATs (Supp Table 2, Supp Table 6). We noted no significant eIF4E-dependent differences for exon length or position (Supp Fig 5A, Supp Fig 5B

Supp Fig 6A, Supp Fig 6B). In all, exons involved in eIF4E-dependent splicing events tend to be bracketed by longer than average introns which likely contain relevant secondary structure elements that act as USER codes for eIF4E-dependent splicing. This will be interesting possibility to investigate in future.

Conclusions

We discovered an unanticipated role for eIF4E in splicing and further, provide a novel paradigm for dysregulation of splicing. We observe that eIF4E reprogrammes splicing of ~15% of transcripts in U2OS cells and ~15-50% of transcripts in High-eIF4E AML. Consistent with these observations, eIF4E drives wide-scale, simultaneous dysregulation in the production of the splicing machinery, and physically interacts with the spliceosome. eIF4E's alterations do not induce well known mutations in SFs or generally alter RNA levels. In this way, classical genomic and transcriptomic approaches would be blind to eIF4E's impacts on splicing.

The physical association of eIF4E with splicing machinery and with nuclear RNAs as seen by RIP suggests that eIF4E could play a direct role in splicing. Given its role in m^7G cap-binding, the simplest possibility appears that eIF4E recruits specific capped RNAs to the spliceosome. This is a significant shift from the current paradigm whereby only the cap-binding complex (CBC) chaperones RNAs to the spliceosome. Recent studies suggest that there is more diversity in this process. For instance, splicing can also occur post-transcriptionally[28], and the role of CBC there has been inferred but remains untested. Our studies suggest the possibility that eIF4E escorts some RNAs to the splicing machinery in a manner similar to CBC, whether these events are co- and/or post-transcriptional is not yet known. Importantly, the interaction of eIF4E and CBC with the m^7G cap on any given RNA is mutually exclusive. Thus, RNAs could only be escorted by one of the cap-binding proteins at a time. In addition to this direct role, eIF4E's impact on the production of SFs suggest that eIF4E could also modify the composition and thus activity of the splicing machinery, potentially producing specialized spliceosomes in order to favour certain splicing outcomes. In all, eIF4E-dependent splicing defines a non-canonical model of splicing, and moreover, positions eIF4E as a factor that can impact all the major levels of RNA maturation.

Given eIF4E only impacts on a subset of transcripts, selection of target RNAs must depend on features in addition to their m^7G cap. eIF4E appears to favour splicing changes to RNAs with longer introns, a feature that is associated with slow and/or post-transcriptional splicing[29]. Further, specificity is not based *a priori* on sensitivity to eIF4E regulation at other levels of RNA processing. For instance, *CCND1*, *MCL1* and *MYC* are found in eIF4E nuclear RIPs and are targets of eIF4E-dependent capping and nuclear export but were not observed as eIF4E-dependent splicing targets according to the rMATs analysis (Supp Table 2, Supp Table 6)[4,12,13,30]. A comparison of data sets reveals that six targets of eIF4E-dependent capping[4] are also splicing targets (*MDM2*, *AGRN*, *TUBB6*, *CATSPER1*, *CTNNB1* and *ERCC*2; Supp Fig 7B, Supp Table 15) indicating that co-regulation at multiple levels can occur. In all, it appears that there are cis-acting elements in the RNA targets that sensitize RNAs specifically to eIF4E-dependent splicing and not universally to all eIF4E-related processes. Identification of these eIF4E-dependent splicing elements will be an important future direction.

Previous studies suggest that mutations in *SF3B1*, *U2AF1* and *SRSF2* do not occur simultaneously in the same AML cell likely because they are synthetically lethal[8]. Strikingly, we observed that eIF4E functionally dysregulates these factors simultaneously through their elevation. Given the pervasive effects of eIF4E, it is extremely unlikely that its splicing signature would be attributable to elevation and/or interaction with any single SF. Consistent with this notion, in secondary AML (sAML) specimens without SF mutations, PRPF6 and SF3B1 proteins were found to be elevated relative to healthy volunteer specimens[19]. Only 363/4670 of eIF4E targets (363/963 for secondary AML targets) were in common between High-eIF4E AML and sAML specimens (Supp Fig 7C, Supp Table 16). Thus, the elevation of PRPF6 and SF3B1 alone did not recapitulate the eIF4E-dependent splicing profile. Moreover, splicing patterns arising from eIF4E overexpression were not simply a recapitulation of those observed in AML cells with SF mutations. For instance, a comparison of splicing events arising in High-eIF4E AML cells relative to AML cells with *SF3B1* hotspot mutation ($SF3B1^{K700}$ and other SF3B1 mutations were also included)[26] showed only a modest overlap (47/4670) of eIF4E targets but included a substantial fraction of SF3B1 targets (47/83 of SF3B1) (Supp Fig 7C, Supp Table 17). Recent findings indicate that Myc hyperactivation stimulates splicing factor SF3A3 translation through eIF machinery, leading to a metabolic reprogramming that amplifies Myc's oncogenic potential[31]. A comparison of splicing

targets arising due to Myc dysregulation with eIF4E targets reveal a common 146/764 of targets for eIF4E and 146/1442 of Myc targets (Supp Fig 7D, Supp Table 18). In all, these analyses suggest that eIF4E dysregulation of splicing is not readily recapitulated by any single downstream target of eIF4E. Our studies with eIF4E also provide at least one possible mechanism for simultaneous dysregulation of multiple splicing factors which has been reported in some solid tumours. For example, PRPF8, SRSF1, SRSF2, and U2AF2 protein levels are upregulated, rather than mutated, in some solid tumours[32]. Consistent with our findings, modulation of the core components of the splicing machinery do not elicit global effects on splicing but rather target specific pre-mRNAs[33].

Recent studies demonstrated that classical mutation analysis of SFs can underestimate their impact on driving post-transcriptional programmes that can support cancer[1]. Thus, we assessed whether eIF4E might impact splicing of SFs themselves, which would be an additional mechanism by which eIF4E could influence splicing. Manual inspection of rMATs data revealed that 5 SFs in U2OS and ~20 SFs in AML undergo eIF4E-dependant AS. Interestingly, *SF3B1* is a splice target of eIF4E (Supp table 2). Moreover, SF3B1 is found in nuclear eIF4E RIPs (Fig 2A) and is also an eIF4E-dependent RNA export target (Supp Fig 2A) suggesting this factor is regulated by eIF4E at multiple levels. In future, it will be exciting to determine whether these splicing events also contribute to the splicing signature of eIF4E

In all, the collective alterations to the SF content in combination with the interaction of some of these factors with eIF4E likely combinatorially drive the splicing patterns observed. Targeted RNAs act in processes that could underpin, at least in part, the oncogenic activities of eIF4E, and likely represent an eIF4E-dependent splicing RNA regulon. In summary, our study reveals a novel means to reprogramme transcriptional signals and allows eIF4E to not only translate, but also to re-write, the transcript message.

Acknowledgements.

We are grateful for assistance from Genomics Platform at IRIC. Primary human specimens were collected and analyzed by the Banque de cellules leucémiques du Québec (BCLQ), supported by the Cancer Research Network of the Fonds de Recherche du Québec - Santé (FRQS). KLBB

acknowledges funding from LLS, NIH, CIHR, Canada Research Chair in Molecular Biology of the Cell Nucleus; and MG from The Cole Foundation and Baumgartner fellowships.

Materials and Methods

Plasmids, antibodies and Reagents

pcDNA-2Flag-eIF4E wild-type and MSCV-pgk-GFP-eIF4E constructs were previously described[13,17,30]. Antibodies for immunoblotting: Mouse monoclonal anti-eIF4E (BD Biosciences), mouse monoclonal anti-β-actin (Sigma Aldrich), rabbit polyclonal anti-Mcl-I (S-19) (Santa Cruz), mouse monoclonal anti-HSP90α/β (F-8) (Santa Cruz), rabbit polyclonal anti-Myc (ab32072 Abcam), rabbit polyclonal antiCyclinD1 (ab134175 Abcam), rabbit polyclonal SF3B1 (Cell Signaling), rabbit polyclonal anti-PRP8 (Bethyl), rabbit polyclonal anti-PRP6 (Bethyl), rabbit polyclonal anti-SnRNP200 (Bethyl), rabbit polyclonal anti-U2AF1 (Bethyl), rabbit polyclonal anti-PRP31 (Bethyl), mouse monoclonal anti-PRP19 (Santa Cruz) and mouse monoclonal anti-U2AF2 (Santa Cruz).

Cell Culture

U2OS cells were obtained from ATCC, and maintained at 37°C and 5% CO2 in Dulbecco's modified Eagle's medium (DMEM) (ThermoFisher Scientific) supplemented with 10% fetal bovine serum (FBS) (ThermoFisher Scientific) and 1% penicillin-streptomycin (ThermoFisher Scientific). NOMO-1 cells (obtained from DSMZ) were maintained at 37°C and 5% CO2 in Roswell Park Memorial Institute (RPMI) 1640 medium (ThermoFisher Scientific) supplemented with 10% FBS and 1% penicillin-streptomycin. Vector and 2FLAG-eIF4E wildtype U2OS cell lines were generated as described previously[13,30] , and maintained as U2OS cells with addition of G-418 (1 mg/mL, Wisent Bioproducts). MSCV-pgk-GFP-eIF4E wildtype were used for retroviral transduction of NOMO-1 cells as in (Topisirovic et al 2003)[17]. Internal ribosomal entry site (IRES) is placed between GFP and eIF4E and thus eIF4E is not produced as a fusion protein. Transduced cells were isolated using FACSAria cell sorter (BD Biosciences). Cell lines were authenticated using STR profiling (Wyndham Forensic Group). Cultured cells were routinely checked to ensure that there was no mycoplasma contamination by PCR (Sung et al, 2006)[34]. U2OS cells were treated

with 20mM ribavirin (Kemprotec, UK) for 72h, and NOMO-1 cells were treated with 10μM ribavirin for 72h.

Primary specimens

For all patients in this study, written informed consent was obtained in accordance with the Declaration of Helsinki. All specimens were deidentified. Patient specimens for analysis of splicing factors levels were obtained from the Banque de Cellules Leucémiques du Québec with institutional ethics approval. Samples for WB in Fig 1G (left panel) and AML-High in Supplementary Fig 1B were samples obtained from AML patients prior to treatment with ribavirin combinations in clinical trial (ClinicalTrials.gov registry is NCT02073838). This study received institutional review board and Health Canada approval. Blasts were isolated using flow cytometry as described (Assouline et al. 2009)[14]. Normal specimens were obtained from ATCC. Samples used for rMATS analysis are listed in Supp table 19

Immunofluorescence and Laser-Scanning Confocal Microscopy

U2OS cells were grown on 4 well glass slides (Millicell EZ SLIDE 4 well glass, Millipore Sigma PEZGS0416). After washing three times in 1× PBS (pH 7.4), cells were fixed in pre-chilled Methanol at -200C for 10 min, and air dried at room temperature for 30min. NOMO-1 cells were harvested and washed three times in 1xPBS in eppendorf tubes, with centrifugation of 5 minutes at 1200 rpm. Washed pellets were resuspended in 1xPBS at 50000 cells/10μl, and 20μl cell suspensions were spotted on glass slides. After drying for 30 minutes at room temperature, NOMO-1 cells were fixed in pre-chilled Methanol at -200C for 10 min, and air dried at room temperature for 30min. After drying, slides were blocked for 1 h in Blocking solution (10% FBS and 0.1% Tween 20 in PBS), and incubated with eIF4E-FITC (BD Biosciences) directly conjugrated antibody diluted in Blocking solution (1:50) overnight at 4 0C. After washing three times in PBS, cells mounted in antifade mounting medium with DAPI (Vector Laboratories). The cells were washed four times with PBS and mounted in mounting media with DAPI (Vector Laboratories, H-2000). Analysis was carried out using a laser-scanning confocal microscope (LSM700 META; Carl Zeiss, Inc.), exciting 405 nm and 488 nm with 63x oil objective and numerical aperture of 1.4. Channels were detected separately, with no cross talk observed.

Confocal micrographs represent single sections through the plane of the cell. Images were obtained from ZEN software (Carl Zeiss, Inc.) and displayed using Adobe Photoshop CS6 (Adobe).

Cellular Fractionation and RNA export assay

About 5×10^7 U2OS cells were collected and washed twice in ice-cold PBS (300 x g for 3–5 min) and then resuspended with slow pipetting in 0.5 mL of lysis buffer B (10 mM Tris (pH 8.4), 140 mM NaCl, 1.5 mM MgCl2, 0.25% Nonidet P-40, 1 mM DTT, 100 U/mL RNase inhibitors). The lysate was centrifuged at 1000 x g for 3 min at 40C, and supernatant (cytoplasmic fraction) was transferred into a fresh microtube. The pellet (nuclear fraction) was resuspended in 1 Volume of lysis buffer B and transferred to a round-bottomed polypropylene tube, and 1/10 volume of detergent stock (3.3% sodium deoxycholate, 6.6% Tween 40 in DEPC H2O) was added with slow vortexing (to prevent the nuclei from clumping) and incubated on ice for 5 min. The suspension was transferred to a microtube and centrifuged at 1000 x g for 3 min at 40C. Supernatant (post-nuclear fraction) was added to cytoplasmic fraction. RNA was extracted from the different fractions by adding TRIzol reagent (ThermoFisher Scientific) and isolated using Direct-zol RNA Mini-prep Kit (Zymo Research).

Co-Immuniprecipitation and RIP

Nuclei isolated using the Cellular Fractionation protocol were rinsed 2x with 1xPBS and fixed with 1% PFA for 10min at RT with rotation, quenched 5min with 0.15M Glycine (RT with rotation), then washed 3 times with 1xPBS and lysed in 0.5ml NT-2 buffer by 3 times 6 seconds bursts (with 30 second pause between each burst) using microtip at 25% power (Sonic Dismembrator Model 500, Fisher, Max Output 400W). NT-2 buffer: 150mM NaCl, 50mM Tris-HCl (pH 7.4), 2.5mM MgCl2, 0.05% NonidentP-40, 8 supplemented with 1mM DTT, 1x protease inhibitors without EDTA, 200U/ml RNaseOut. Nuclear lysates were centrifuged at 10 000 x g for 10min, and supernatants were transferred into fresh tubes. After adjusting the concentration to be no more than 1mg/ml, nuclear extracts were pre-cleared with 50 µL protein G conjugated superparamagnetic beads (Dynabeads Protein G, ThermoFisher Scientific) for 30 min at 40C. Pre-cleared lysates (1mg) were incubated with 10µg of anti-eIF4E antibody(RN001P, MBL) or 10 µg of appropriate IgG as a control, and 0.5 mg/ml yeast tRNA (SigmaAldrich), overnight at 40C with

rotation. After ON incubation, 50µl of Dynabeads were added and incubated for additional 3h at 40C with rotation. Beads were washed once with NT-2 buffer supplemented with 1mg/mL heparin (Sigma-Aldrich) for 5min at 40C with rotation, and an additional six times with NT-2 buffer with 300mM NaCl. After washing, beads were resuspended in 2xLaemmli Buffer with β-mercaptoethanol and incubated for 5 min at 980C. Co-immunoprecipitated proteins were resolved on SDS-PAGE and visualized by Western blotting. To isolate RNAs from immunoprecipitated reactions, beads were resuspended in Elution Buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 12% (v/v) β-mercaptoethanol, and incubated for 5 min at 980C. RNA were isolated using TRIzol reagent (ThermoFisher) and Direct-zol RNA Micro-prep Kit (Zymo Research).

Reverse transcription and quantitative PCR

RNA samples were reversed transcribed using SuperScript VILO cDNA synthesis kit (for RIP experiments) (ThermoFisher Scientific) or MMLV reverse transcription kit and oligo-dT or Random hexamers (ThermoFisher Scientific). QPCR analyses were performed using SensiFast Sybr Lo-Rox Mix (Bioline, MA, U.S.A, Cat# BIO-94020) in Applied Biosystems Viia7 or QuantStudio7 thermal cyclers using the relative standard curve method (Applied Biosystems User Bulletin #2). All the primers are listed in Supp 20 and all conditions were described previously[30].

Western blot analysis was performed as described previously[23] . Blots were blocked in 5% milk in TBS–Tween 20. Primary antibodies were diluted in 5% milk.

Analysis of alternative splicing

Alternative splicing events were detected in RNA-seq data and the probability that the differences in isoform abundance were evaluated statistically with the help of rMATS version 4.1.1. Sequenced fragments were treated as non stranded to allow a uniform treatment of both stranded and non stranded samples and variable read lengths were allowed in the analysis. JC and JCEC quantifications have been combined for the complete analysis. Events with a false discovery rate below 0.1 or 0.15 and an absolute inclusion level higher than 0.1 or 0.05 were considered

significant and carried further in the analysis. Maxent was used to evaluate the strength of significant splice site events. MAXENT ranges considered for the analysis were : Strong: MAXENT score >= 7, Intermediate: MAXENT score >= 3 and < 7, Weak: MAXENT score < 3.

Data analysis and visualization. DESeq2 version 1.30.1[35] was used to normalize gene readcounts (non stranded) for the 21 AML samples. Significant differentially expressed genes (DEGs) are typically those with padj lower than 0.05 and an absolute fold change > 2. **Data visualization, graphics and plots** were made using R package ggplot2 and related packages digest, glue, grDevices, grid, gTable, isoband, MASS, mgcv, rlang, scales, stats, tibble, withr and dplyr. pheatmap" package was applied to construct heat maps and hierarchical clustering analyses. **Process and pathway enrichment analyses** were performed using METASCAPE[36]. Briefly, for each gene list, pathway and process enrichment analysis has been carried out with the following ontology sources: KEGG Pathway, GO Biological Processes, Reactome Gene Sets, Canonical Pathways, WikiPathways and PANTHER Pathway. All genes in the genome have been used as the enrichment background. Terms with a p-value < 0.01, a minimum count of 3, and an enrichment factor > 1.5 (the enrichment factor is the ratio between the observed counts and the counts expected by chance) are collected and grouped into clusters based on their membership similarities. More specifically, p-values are calculated based on the accumulative hypergeometric distribution, and q-values are calculated using the Benjamini-Hochberg procedure to account for multiple testings. **Protein Protein Interaction Analysis** was assisted by METASCAPE[36] and CYTOSCAPE[37]. For each target gene list, protein-protein interaction enrichment analysis has been carried out with the following databases: STRING, BioGrid, OmniPath, InWeb_IM. Only physical interactions in STRING (physical score > 0.132) and BioGrid are used. The resultant network contains the subset of proteins that form physical interactions with at least one other member in the list. If the network contains more than 5 proteins, the Molecular Complex Detection (MCODE) algorithm has been applied to identify densely connected network components. Pathway and process enrichment analysis has been applied to each MCODE component independently, and the best-scoring term by p-value have been retained as the functional description of the corresponding component. **Venn diagrams** were performed using Bioinformatics and evolutionary genomics webtool (http://bioinformatics.psb.ugent.be/webtools/Venn/). **All Figs and cartoons** were edited and laid out using the open-source Vector graphics editor Inkscape (https://inkscape.org/)

Figure Legends.

Fig1. eIF4E increases expression of the splicing machinery. A. Localization of eIF4E in U2OS cells. Confocal micrographs of cells stained with anti-eIF4E antibodies to detect endogenous eIF4E and DAPI as a nuclear marker. Single (eIF4E) and overlaid (eIF4E+DAPI) channels are shown. Micrographs are single sections through the plane of the cells with 63x magnification. B. Western blot (WB) analysis of splicing factors for Vector and 2FLAG-eIF4E U2OS cell lines. Myc, Mcl1, and CCND1 served as positive controls, and as a loading control. Both 2FlAG-eIF4E (2F4E) and endogenous eIF4E are shown. Each β-Actin blot corresponds to the above western blots. Experiments were carried out at least three independent times, and one representative experiment is shown. C. WB analysis of splicing factors as a function of eIF4E reduction using CRISPR-eIF4E and CRISPR-Ctrl cell lines. Myc, Mcl1 and CCND1 served as positive controls, and β-Actin was used as a loading control. Each β-Actin blot corresponds to the above western blots. Experiments were carried out at least three independent times, and one representative experiment is shown. D. eIF4E is positioned to influence multiple facets of spliceosome activity. Schematic representation summarizing the splicing factors targeted by eIF4E. Proteins are grouped based on their activity and/or association with a specific complex of the major spliceosome. eIF4E also physically interacts with all of the UsnRNAs shown (see Figure 2A and Supp Figure 2). E. WB analysis of splicing factors for Vector and eIF4E overexpressing NOMO-1 cell lines. Myc, and CCND1 served as positive controls, and β-Actin as a loading control. Each β-Actin blot corresponds to the above western blots. Experiments were carried out at least three independent times, and one representative experiment is shown. F. Western blot analysis of untreated (Ctrl) and Ribavirin treated (Rib) Vector and eIF4E NOMO-1 cell lines. Myc, and Mcl1 served as positive controls, and β-Actin as a loading control. Each β-Actin blot corresponds to the above WBs. Experiments were carried out at least three independent times, and one representative experiment is shown. G. WB analysis of Splicing Factors levels in primary AML samples with high (AML-H) or normal (AML-N) eIF4E levels, as well as bone marrow mononuclear cells from healthy volunteers (Norm for normal). β-Actinwas used as a loading control. (* = degraded sample).

Fig 2. eIF4E increases expression and interacts with components of the splicing machinery. A. The enrichment of mRNAs in RIPs of endogenous eIF4E versus input RNAs from the nuclear fractions of vector control U2OS cells monitored by RT-qPCR. Data were normalized to input samples and presented as a fold change. The mean, standard deviation and p-values were derived from five independent experiments (each carried out in triplicate). Myc, and Mcl1 are established eIF4E nuclear targets and served as positive controls, while ACTB, GAPDH, POLR2A and 18S rRNA were negative controls. B. eIF4E co-immunoprecipitated with SFs in the nuclear fractions of U2OS cells. Immunoprecipitations (IP) were carried out using U2OS nuclear lysates and anti-eIF4E antibody (eIF4E-IP) or appropriate IgG control (IgG-IP). IP samples along with Input (2%) and supernatants (Sn) after IPs were analyzed by WB using antibodies as indicated. C. eIF4E co-immunoprecipitated with SFs in the nuclear fractions of NOMO-1 cells. IPs were carried out using eIF4E overexpressing NOMO-1 nuclear lysates and anti-eIF4E antibody (eIF4E-IP) or appropriate IgG control (IgG-IP). IP samples were analyzed by WB as above and carried out in three biological replicates.

Fig 3. eIF4E overexpression reprogrammes splicing of a subset of transcripts in U2OS cells. A. Schematic representation of AS events characterized by rMATS. SE: skipped exon; MXE: mutually exclusive spliced exon; A5SS: alternative 5′ splicing site; A3SS: alternative 3′ splicing site; RI: retained intron. **B.** *Left.* Dot plot showing distribution of Inclusion Level Differences (IncLevDiff) values (IncLev (2FLAG-eIF4E) – IncLev (Vector)) for each splicing category. *Right.* Histogram showing the number of events for each splicing category. **C.** Splicing heatmap showing the values of IncLev for all the Skipping exon (SE) events between Vector (n =3, biological replicates) and 2FLAG-eIF4E (n = 3, biological replicates) U2OS cells. Other events are shown in Supp Fig3A. **D.** Pathway and process enrichment analysis with a bar graph showing top ranked enriched terms for alternatively spliced genes between Vector and 2FLAG-eIF4EU2OS cells .**E.** MCODE networks identified for alternatively spliced genes. Pathway and process enrichment analysis has been applied for each MCODE component and the best-scoring term by p-value is shown.

Fig 4. eIF4E levels correlate with a substantial alternative splicing reprogramming in AML.

A. Graph showing the distribution of eIF4E RNA levels in AML patients' specimens (blue, AML) in comparison with normal CD34+ cells (orange, hHSC). Data are derived from the Leucegene cohort. **B.** High-eIF4E (purple) and Normal-eIF4E (blue) classification of AML patients' specimens. eIF4E levels in CD34+ cells from cord blood derived from healthy volunteers is shown for comparison. These specimens were selected both on their eIF4E levels and the absence of SF mutations as observed from the Leucegene RNA-Seq dataset (Supp table 19). **C.** Survival analyses for of High-eIF4E patients and Normal-eIF4E patients for which specimens were studied (Supp table 19). The rest of the cohort is shown for comparison. Log Rank test was performed and shows a significant difference in survival between the High-eIF4E group and the rest of the cohort, as well as between High-eIF4E and Normal-eIF4E groups (p-value < 0.05). **D.** *Left.* Dot plot showing distribution of IncLevDiff values (IncLev (High-eIF4E) – IncLev (Normal-eIF4E)) for each splicing category. *Right.* Histogram showing the number of events for each splicing category. **E.** Splicing heatmap showing the values of IncLev for all the splicing events between Normal-eIF4E and High-eIF4EAML samples. **F.** Pathway and process enrichment analysis with a bar graph showing top ranked enriched terms for alternatively spliced genesbetweenNormal-eIF4E and High-eIF4E AML patients' samples. **G.** MCODE networks identified for MXE targets between Normal-eIF4Eand High-eIF4EAML patients' samples. Pathway and process enrichment analysis has been applied for each MCODE component and the best-scoring term by p-value.

Fig 5. Identification and characterization of a core alternative splicing network targeted by eIF4E in both U2OS and AML contexts. A. Venn diagram showing the overlap between genes affected by eIF4E-dependant alternative splicing in High-eIF4E AML patients' cells and 2FLAG-eIF4E U2OS cells. **B.** Pathway and process enrichment analyses with a bar graph showing top ranked enriched terms for alternatively spliced genes in common between and High-eIF4E AML patients' samples and 2FLAG-eIF4E U2OS cells. **C.**MCODE networks for identified alternative splicing targets common between High-eIF4E AML patients' samples and 2FLAG-eIF4E U2OS cells. Pathway and process enrichment analysis has been applied for each MCODE component and the best-scoring term by p-value. Common targets selected for subsequent validation by RT-qPCR are highlighted as larger circles (IkBkB, IL4R, FAAP20, IL1B). See main text for full description. **D. and E.** Validation of splicing targets identified by the rMATs in U2OS (D) and NOMO-1 (E) cell lines. RT-qPCR analysis using specific primers for each splicing event

normalized to the corresponding total levels of that given transcript (using primers specific to common regions without changes). Data were normalized to vector control to calculate fold change. The mean and standard deviation, as well as p-values were derived from three independent experiments. P-values were calculated between eIF4E overexpressing and eIF4E cells treated with Ribavirin, p-values for vector comparisons not shown.

References

1 Rivera, O. D. *et al.* Alternative splicing redefines landscape of commonly mutated genes in acute myeloid leukemia. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2014967118 (2021).

2 Borden, K. L. The eukaryotic translation initiation factor eIF4E wears a "cap" for many occasions. *Translation (Austin)* **4**, e1220899, doi:10.1080/21690731.2016.1220899 (2016).

3 Borden, K. L. B. The Nuclear Pore Complex and mRNA Export in Cancer. *Cancers (Basel)* **13**, doi:10.3390/cancers13010042 (2020).

4 Culjkovic-Kraljacic, B. *et al.* The eukaryotic translation initiation factor eIF4E elevates steady-state m(7)G capping of coding and noncoding transcripts. *Proc Natl Acad Sci U S A* **117**, 26773-26783, doi:10.1073/pnas.2002360117 (2020).

5 Culjkovic-Kraljacic, B. & Borden, K. L. B. The Impact of Post-transcriptional Control: Better Living Through RNA Regulons. *Front Genet* **9**, 512, doi:10.3389/fgene.2018.00512 (2018).

6 Davis, M. R., Delaleau, M. & Borden, K. L. B. Nuclear eIF4E Stimulates 3'-End Cleavage of Target RNAs. *Cell Rep* **27**, 1397-1408 e1394, doi:10.1016/j.celrep.2019.04.008 (2019).

7 Wahl, M. C., Will, C. L. & Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701-718, doi:10.1016/j.cell.2009.02.009 (2009).

8 Visconte, V., M, O. N. & H, J. R. Mutations in Splicing Factor Genes in Myeloid Malignancies: Significance and Impact on Clinical Features. *Cancers (Basel)* **11**, doi:10.3390/cancers11121844 (2019).

9 Topisirovic, I. *et al.* Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. *The EMBO journal* **28**, 1087-1098, doi:10.1038/emboj.2009.53 (2009).

10 Volpon, L. *et al.* A biochemical framework for eIF4E-dependent mRNA export and nuclear recycling of the export machinery. *RNA* **23**, 927-937, doi:10.1261/rna.060137.116 (2017).

11 Culjkovic-Kraljacic, B., Baguet, A., Volpon, L., Amri, A. & Borden, K. L. B. The oncogene eIF4E reprograms the nuclear pore complext to promote mRNA export and oncogenic transformation. *Cell Reports* **2**, 207-215 (2012).

12 Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M. & Borden, K. L. eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3'UTR. *J Cell Biol* **169**, 245-256 (2005).
13 Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M. & Borden, K. L. eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J Cell Biol* **175**, 415-426 (2006).
14 Assouline, S. *et al.* Molecular targeting of the oncogene eIF4E in acute myeloid leukemia (AML): a proof-of-principle clinical trial with ribavirin. *Blood* **114**, 257-260, doi:blood-2009-02-205153 [pii]
10.1182/blood-2009-02-205153 (2009).
15 Assouline, S. *et al.* A phase I trial of ribavirin and low-dose cytarabine for the treatment of relapsed and refractory acute myeloid leukemia with elevated eIF4E. *Haematologica* **100**, e7-9, doi:10.3324/haematol.2014.111245 (2015).
16 de Necochea-Campion, R., Shouse, G. P., Zhou, Q., Mirshahidi, S. & Chen, C. S. Aberrant splicing and drug resistance in AML. *J Hematol Oncol* **9**, 85, doi:10.1186/s13045-016-0315-9 (2016).
17 Topisirovic, I. *et al.* Aberrant eukaryotic translation initiation factor 4E-dependent mRNA transport impedes hematopoietic differentiation and contributes to leukemogenesis. *Molecular and cellular biology* **23**, 8992-9002 (2003).
18 Adamia, S. *et al.* A genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. *Clin Cancer Res* **20**, 1135-1145, doi:10.1158/1078-0432.CCR-13-0956 (2014).
19 Crews, L. A. *et al.* RNA Splicing Modulation Selectively Impairs Leukemia Stem Cell Maintenance in Secondary Human AML. *Cell Stem Cell* **19**, 599-612, doi:10.1016/j.stem.2016.08.003 (2016).
20 Dunn, L. A. *et al.* Phase I study of induction chemotherapy with afatinib, ribavirin, and weekly carboplatin and paclitaxel for stage IVA/IVB human papillomavirus-associated oropharyngeal squamous cell cancer. *Head Neck* **40**, 233-241, doi:10.1002/hed.24938 (2017).
21 Kosaka, T., Maeda, T, Shinojima, T, Nagata, H, Ryuchi, R, Oya M. a clinical study to evaluate the efficacy and safety of docetaxal with ribavirin in patients with progressive castration resistant prostate cancer who have previously received docetaxol alone. *Journal of Clinical Oncology*, https:/ascopubs.org/doi/abs/10.1200/JCO2017 .1235.1215_suppl.e14010 (2017).
22 Kraljacic, B. C., Arguello, M., Amri, A., Cormack, G. & Borden, K. Inhibition of eIF4E with ribavirin cooperates with common chemotherapies in primary acute myeloid leukemia specimens. *Leukemia* **25**, 1197-1200, doi:10.1038/leu.2011.57 (2011).
23 Zahreddine, H. A. *et al.* The eukaryotic translation initiation factor eIF4E harnesses hyaluronan production to drive its malignant activity. *Elife* **6**, doi:10.7554/eLife.29830 (2017).
24 Kapp, L. D. & Lorsch, J. R. The molecular mechanics of eukaryotic translation. *Annu Rev Biochem* **73**, 657-704, doi:10.1146/annurev.biochem.73.030403.080419 (2004).

25 Shen, S. *et al.* MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res* **40**, e61, doi:10.1093/nar/gkr1291 (2012).
26 Hershberger, C. E. *et al.* Complex landscape of alternative splicing in myeloid neoplasms. *Leukemia* **35**, 1108-1120, doi:10.1038/s41375-020-1002-y (2021).
27 Mehmood, A. *et al.* Systematic evaluation of differential splicing tools for RNA-seq studies. *Brief Bioinform* **21**, 2052-2065, doi:10.1093/bib/bbz126 (2020).
28 Girard, C. *et al.* Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat Commun* **3**, 994, doi:10.1038/ncomms1998 (2012).
29 Boutz, P. L., Bhutkar, A. & Sharp, P. A. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev* **29**, 63-80, doi:10.1101/gad.247361.114 (2015).
30 Culjkovic-Kraljacic, B. *et al.* Combinatorial targeting of nuclear export and translation of RNA inhibits aggressive B-cell lymphomas. *Blood* **127**, 858-868, doi:10.1182/blood-2015-05-645069 (2016).
31 Ciesla, M. *et al.* Oncogenic translation directs spliceosome dynamics revealing an integral role for SF3A3 in breast cancer. *Mol Cell* **81**, 1453-1468 e1412, doi:10.1016/j.molcel.2021.01.034 (2021).
32 Urbanski, L. M., Leclair, N. & Anczukow, O. Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. *Wiley Interdiscip Rev RNA* **9**, e1476, doi:10.1002/wrna.1476 (2018).
33 Papasaikas, P., Tejedor, J. R., Vigevani, L. & Valcarcel, J. Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol Cell* **57**, 7-22, doi:10.1016/j.molcel.2014.10.030 (2015).
34 Sung, H. *et al.* PCR-based detection of Mycoplasma species. *J Microbiol* **44**, 42-49 (2006).
35 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
36 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523, doi:10.1038/s41467-019-09234-6 (2019).
37 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

Fig 1

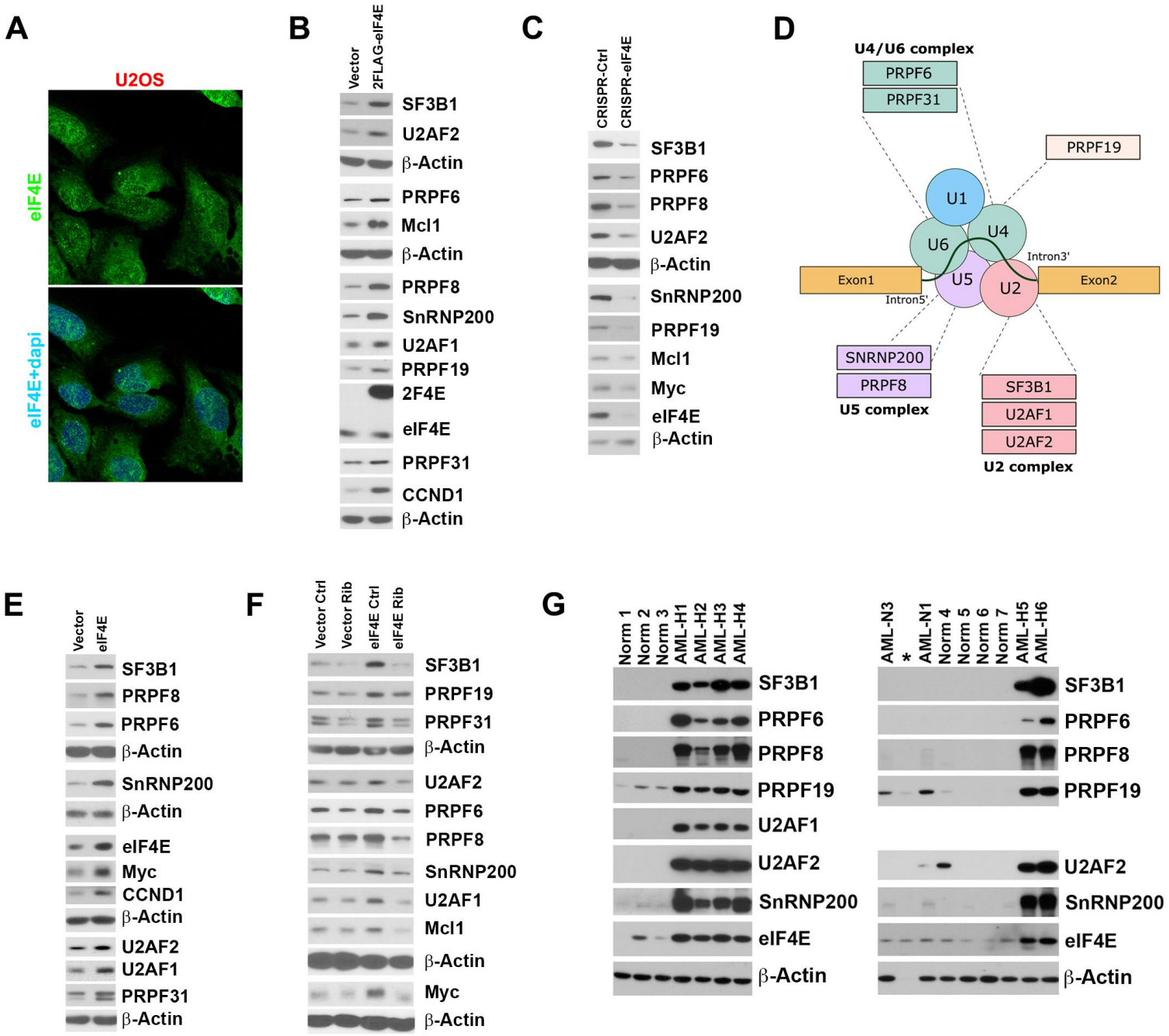

A
U2OS
eIF4E
eIF4E+dapi
B
Vector
2FLAG-eIF4E
SF3B1
U2AF2
β-Actin
PRPF6
Mcl1
β-Actin
PRPF8
SnRNP200
U2AF1
PRPF19
2F4E
eIF4E
PRPF31
CCND1
β-Actin
C
CRISPR-Ctrl
CRISPR-eIF4E
SF3B1
PRPF6
PRPF8
U2AF2
β-Actin
SnRNP200
PRPF19
Mcl1
Myc
eIF4E
β-Actin
D
U4/U6 complex
PRPF6
PRPF31
PRPF19
U1
U4
U6
U5
U2
Intron3'
Intron5'
Exon1
Exon2
SNRNP200
PRPF8
U5 complex
SF3B1
U2AF1
U2AF2
U2 complex
E
Vector
eIF4E
SF3B1
PRPF8
PRPF6
β-Actin
SnRNP200
β-Actin
eIF4E
Myc
CCND1
β-Actin
U2AF2
U2AF1
PRPF31
β-Actin
F
Vector Ctrl
Vector Rib
eIF4E Ctrl
eIF4E Rib
SF3B1
PRPF19
PRPF31
β-Actin
U2AF2
PRPF6
PRPF8
SnRNP200
U2AF1
Mcl1
β-Actin
Myc
β-Actin
G
Norm 1
Norm 2
Norm 3
AML-H1
AML-H2
AML-H3
AML-H4
SF3B1
PRPF6
PRPF8
PRPF19
U2AF1
U2AF2
SnRNP200
eIF4E
β-Actin
AML-N3
*
AML-N1
Norm 4
Norm 5
Norm 6
Norm 7
AML-H5
AML-H6
SF3B1
PRPF6
PRPF8
PRPF19
U2AF2
SnRNP200
eIF4E
β-Actin

Fig 2

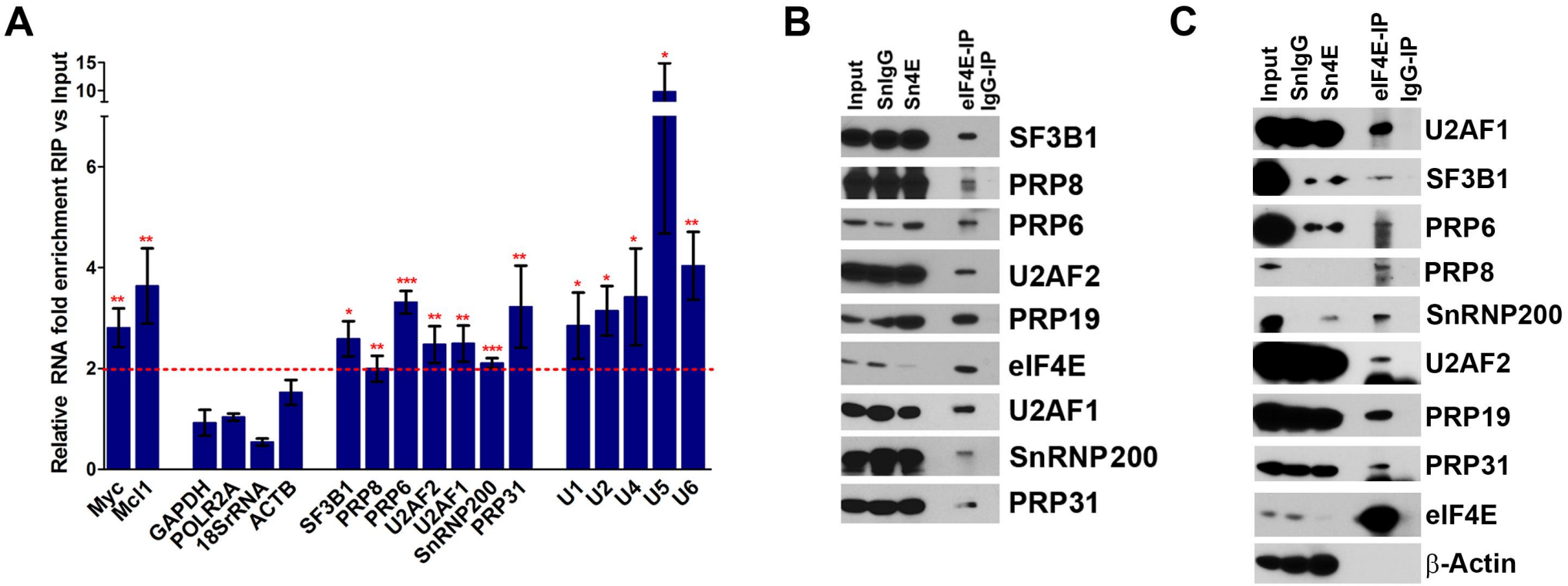
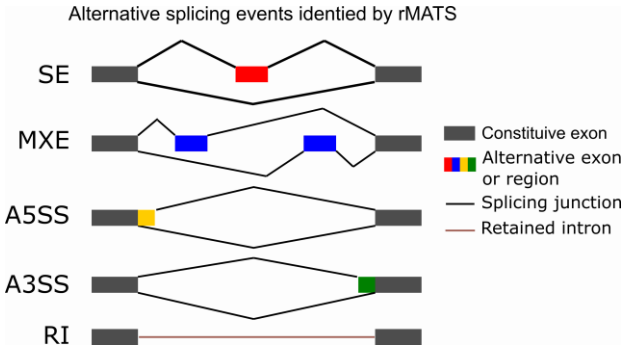
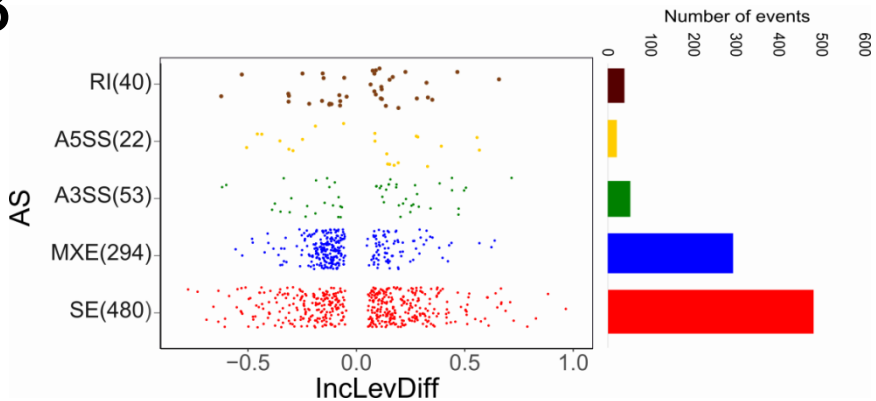

A
Relative RNA fold enrichment RIP vs Input
15
10
6
4
2
0
Myc
Mcl1
GAPDH
POLR2A
18SrRNA
ACTB
SF3B1
PRP8
PRP6
U2AF2
U2AF1
SnRNP200
PRP31
U1
U2
U4
U5
U6
B
Input
SnIgG
Sn4E
eIF4E-IP
IgG-IP
SF3B1
PRP8
PRP6
U2AF2
PRP19
eIF4E
U2AF1
SnRNP200
PRP31
C
Input
SnIgG
Sn4E
eIF4E-IP
IgG-IP
U2AF1
SF3B1
PRP6
PRP8
SnRNP200
U2AF2
PRP19
PRP31
eIF4E
β-Actin

Fig 3

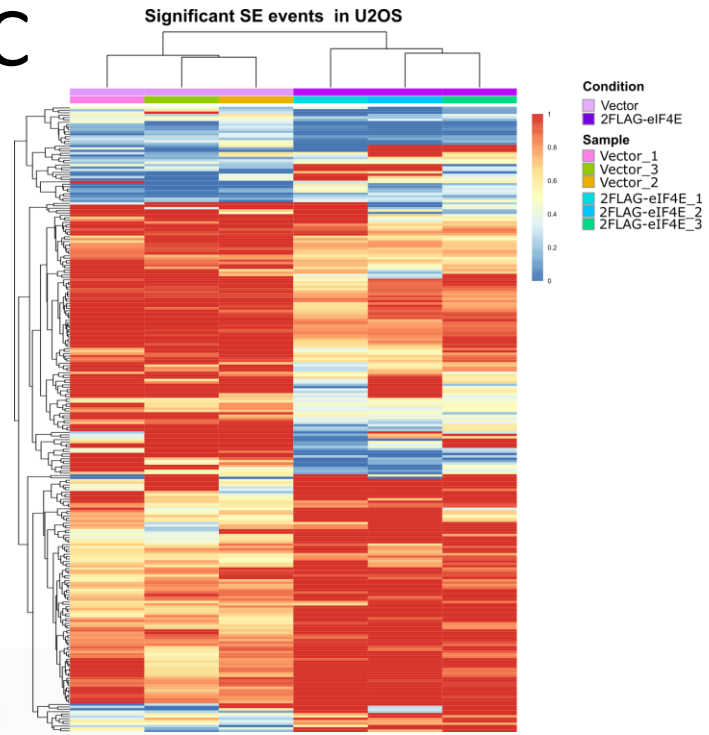
A


Alternative splicing events identied by rMATS
SE
MXE
A5SS
A3SS
RI
Constituive exon
Alternative exon or region
Splicing junction
Retained intron

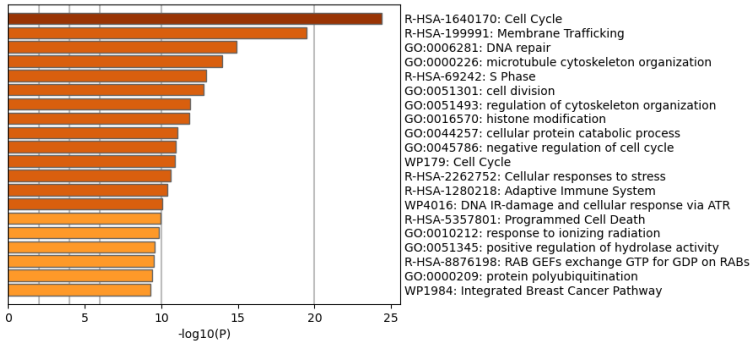
B


Number of events
0
100
200
300
400
500
600
RI(40)
A5SS(22)
A3SS(53)
MXE(294)
SE(480)
AS
−0.5
0.0
0.5
1.0
IncLevDiff

C


Significant SE events in U2OS
Condition
Vector
2FLAG-eIF4E
Sample
Vector_1
Vector_3
Vector_2
2FLAG-eIF4E_1
2FLAG-eIF4E_2
2FLAG-eIF4E_3

D


R-HSA-1640170: Cell Cycle
R-HSA-199991: Membrane Trafficking
GO:0006281: DNA repair
GO:0000226: microtubule cytoskeleton organization
R-HSA-69242: S Phase
GO:0051301: cell division
GO:0051493: regulation of cytoskeleton organization
GO:0016570: histone modification
GO:0044257: cellular protein catabolic process
GO:0045786: negative regulation of cell cycle
WP179: Cell Cycle
R-HSA-2262752: Cellular responses to stress
R-HSA-1280218: Adaptive Immune System
WP4016: DNA IR-damage and cellular response via ATR
R-HSA-5357801: Programmed Cell Death
GO:0010212: response to ionizing radiation
GO:0051345: positive regulation of hydrolase activity
R-HSA-8876198: RAB GEFs exchange GTP for GDP on RABs
GO:0000209: protein polyubiquitination
WP1984: Integrated Breast Cancer Pathway
0
5
10
15
20
25
-log10(P)

E

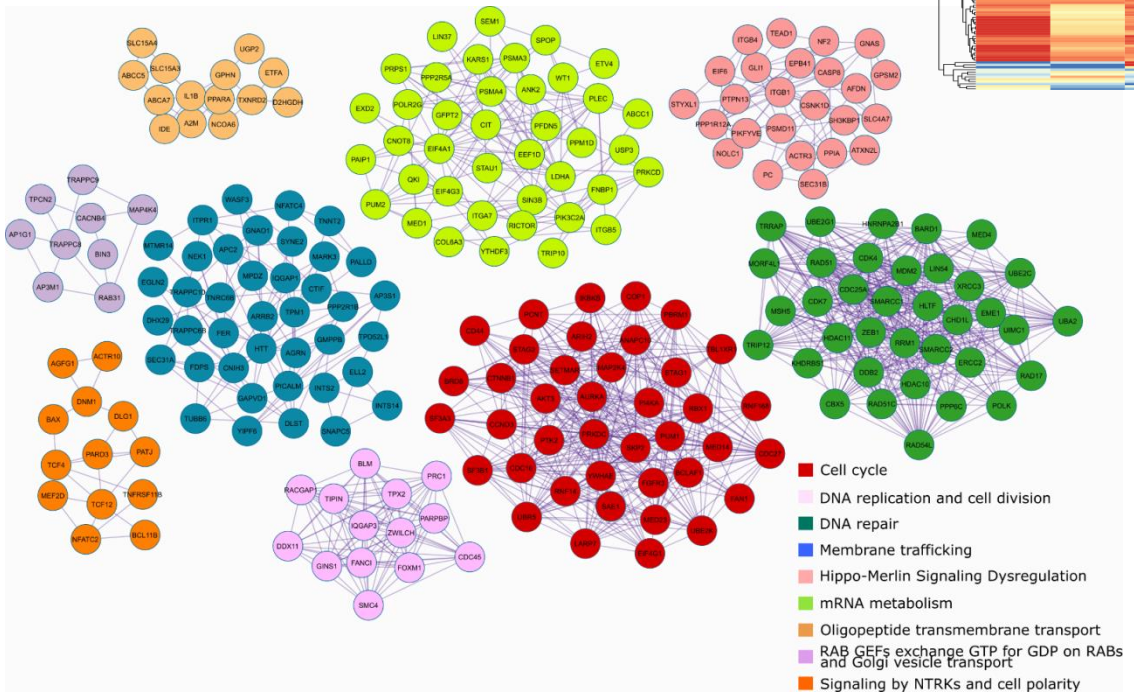

Cell cycle
DNA replication and cell division
DNA repair
Membrane trafficking
Hippo-Merlin Signaling Dysregulation
mRNA metabolism
Oligopeptide transmembrane transport
RAB GEFs exchange GTP for GDP on RABs and Golgi vesicle transport
Signaling by NTRKs and cell polarity

Fig 4

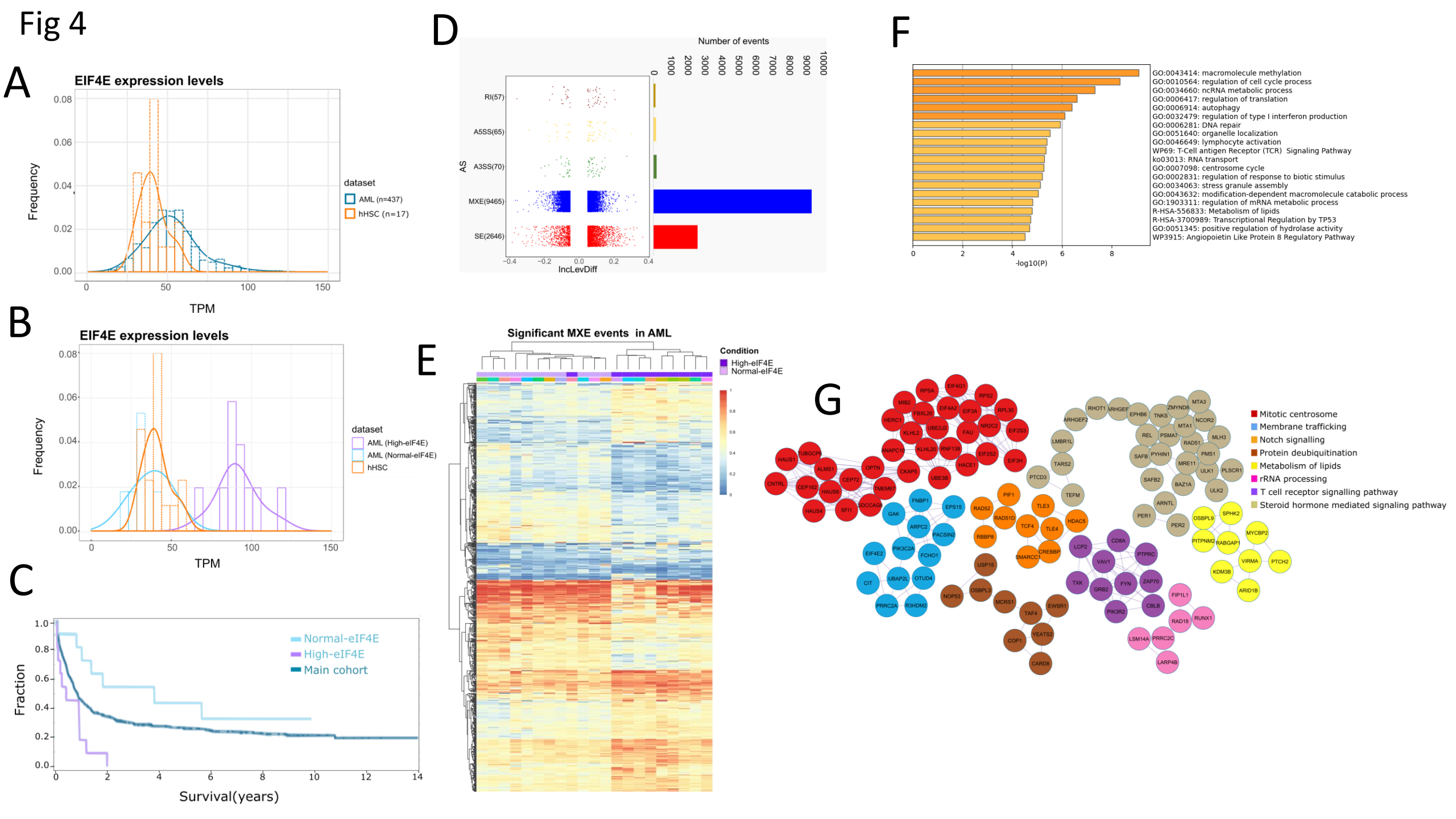
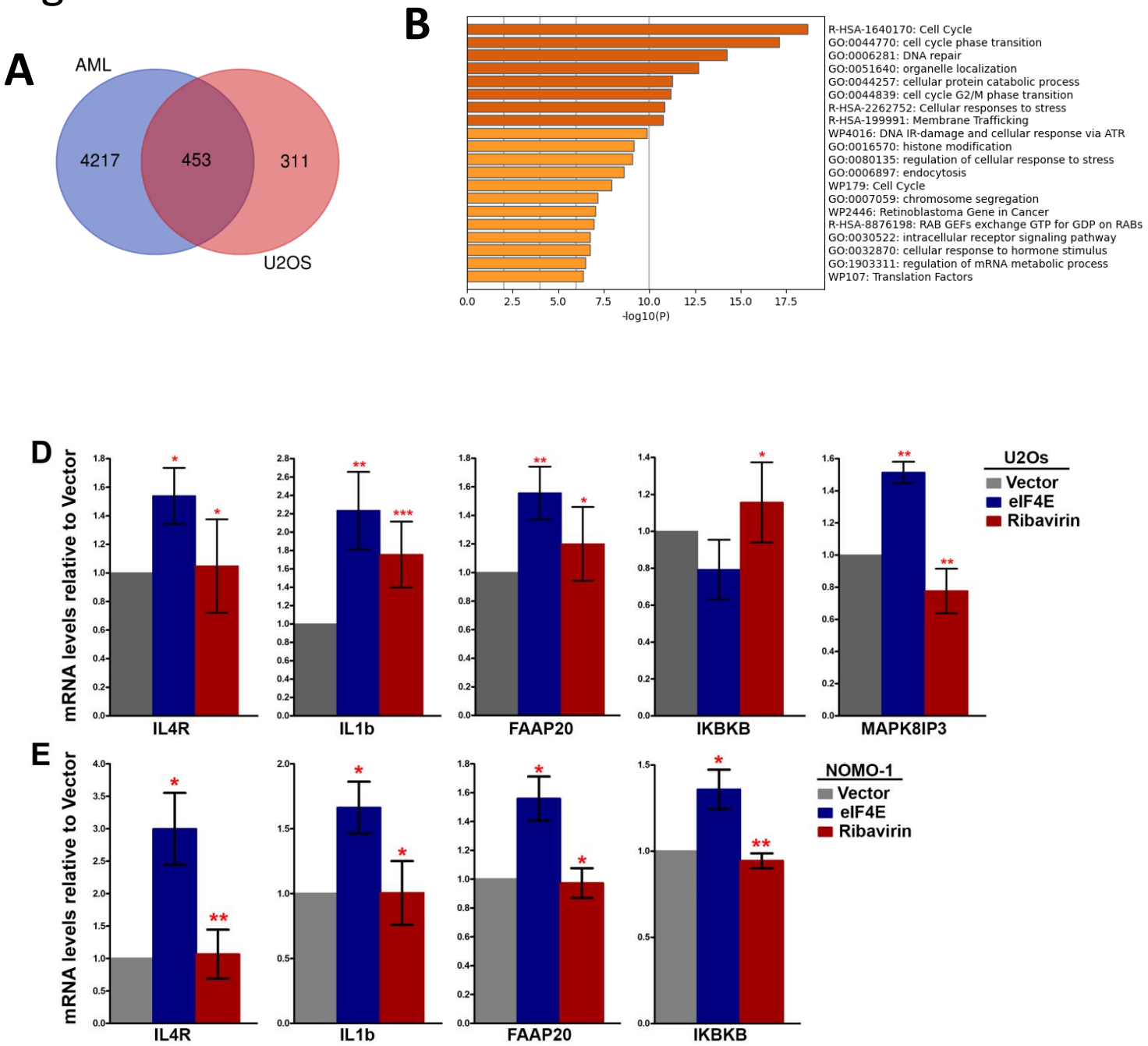

A
EIF4E expression levels
Frequency
TPM
dataset
AML (n=437)
hHSC (n=17)
B
EIF4E expression levels
Frequency
TPM
dataset
AML (High-eIF4E)
AML (Normal-eIF4E)
hHSC
C
Fraction
Survival(years)
Normal-eIF4E
High-eIF4E
Main cohort
D
Number of events
AS
RI(57)
A5SS(65)
A3SS(70)
MXE(9465)
SE(2646)
IncLevDiff
E
Significant MXE events in AML
Condition
High-eIF4E
Normal-eIF4E
F
GO:0043414: macromolecule methylation
GO:0010564: regulation of cell cycle process
GO:0034660: ncRNA metabolic process
GO:0006417: regulation of translation
GO:0006914: autophagy
GO:0032479: regulation of type I interferon production
GO:0006281: DNA repair
GO:0051640: organelle localization
GO:0046649: lymphocyte activation
WP69: T-Cell antigen Receptor (TCR) Signaling Pathway
ko03013: RNA transport
GO:0007098: centrosome cycle
GO:0002831: regulation of response to biotic stimulus
GO:0034063: stress granule assembly
GO:0043632: modification-dependent macromolecule catabolic process
GO:1903311: regulation of mRNA metabolic process
R-HSA-556833: Metabolism of lipids
R-HSA-3700989: Transcriptional Regulation by TP53
GO:0051345: positive regulation of hydrolase activity
WP3915: Angiopoietin Like Protein 8 Regulatory Pathway
-log10(P)
G
Mitotic centrosome
Membrane trafficking
Notch signalling
Protein deubiquitination
Metabolism of lipids
rRNA processing
T cell receptor signalling pathway
Steroid hormone mediated signaling pathway

Fig 5


A
AML
4217
453
311
U2OS
B
R-HSA-1640170: Cell Cycle
GO:0044770: cell cycle phase transition
GO:0006281: DNA repair
GO:0051640: organelle localization
GO:0044257: cellular protein catabolic process
GO:0044839: cell cycle G2/M phase transition
R-HSA-2262752: Cellular responses to stress
R-HSA-199991: Membrane Trafficking
WP4016: DNA IR-damage and cellular response via ATR
GO:0016570: histone modification
GO:0080135: regulation of cellular response to stress
GO:0006897: endocytosis
WP179: Cell Cycle
GO:0007059: chromosome segregation
WP2446: Retinoblastoma Gene in Cancer
R-HSA-8876198: RAB GEFs exchange GTP for GDP on RABs
GO:0030522: intracellular receptor signaling pathway
GO:0032870: cellular response to hormone stimulus
GO:1903311: regulation of mRNA metabolic process
WP107: Translation Factors
-log10(P)
C
DNA repair
Adaptive immune system
MYC activation pathway
Cytosolic transport
PPAR signaling pathway
Phosphatidylinositol signaling system
Ciliary basal body-plasma membrane docking
Class I mediated antigen processing and presentation
D
mRNA levels relative to Vector
IL4R
IL1b
FAAP20
IKBKB
MAPK8IP3
U2Os
Vector
eIF4E
Ribavirin
E
mRNA levels relative to Vector
IL4R
IL1b
FAAP20
IKBKB
NOMO-1
Vector
eIF4E
Ribavirin