

Isoform and pathway-specific regulation of post-transcriptional RNA processing in human cells

Ishwarya Venkata Narayanan^{1,7}, Karan Bedi^{1,2,3,7}, Brian Magnuson^{1,4}, Ariel McShane^{1,5}, Mario Ashaka¹, Michelle T. Paulsen¹, Thomas E. Wilson^{3,4} and Mats Ljungman^{1,3,6}

¹ Department of Radiation Oncology, University of Michigan, Ann Arbor, MI 48109, USA.

² Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA.

³ Rogel Cancer Center and Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI 48109, USA.

⁴ Department of Pathology and Department of Human Genetics, University of Michigan Medical School, University of Michigan, Ann Arbor, MI 48109, USA.

⁵ Program in Cellular and Molecular Biology, University of Michigan, Ann Arbor, MI 48109, USA

⁶ Department of Environmental Health Sciences, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA.

⁷These authors contributed equally to this work

Corresponding author: ljungman@umich.edu

ABSTRACT

Steady-state levels of RNA transcripts are determined by their rates of synthesis and degradation. In this study, we use nascent RNA Bru-seq and BruChase-seq to investigate RNA dynamics across 16 human cell lines, as part of the ENCODE4 Deeply Profiled Cell Lines collection. Our analyses show that RNA turnover dynamics differ widely between transcripts of different genes and among different classes of RNA but remain primarily cell-type independent. Gene set enrichment analysis (GSEA) reveals that transcripts encoding proteins belonging to the same pathway often show similar turnover dynamics. Furthermore, transcript isoforms exhibit distinct turnover profiles, suggesting that RNA turnover plays an important role in regulating mRNA isoform choice. Finally, splicing of newly synthesized transcripts appears to occur in a binary manner, largely happening either fully on both sides of 47,163 exons or not at all. These data sets generated as part of the ENCODE Project illustrate the intricate and coordinated regulation of RNA dynamics in controlling gene expression, allowing for the precise coordination of cellular functions.

INTRODUCTION

Gene expression is regulated at many levels in mammalian cells. Transcription initiation is directed by epigenetic marks allowing regulatory DNA elements to be accessible to activating or repressive transcription factors (The ENCODE Project Consortium et al. 2020). Epigenetic regulation to compact chromatin restricts transcription and we recently reported that only about 20% of the genome in human cells is actively producing measurable amounts of RNA (McShane et al. 2024). Additionally, chromatin

organization dictates the proximity of genes to regulatory elements (Djebali et al. 2012; Grubert et al. 2020; Vierstra et al. 2020). The abundance and quality of synthesized RNAs are then regulated through different processing and degradation pathways in the nucleus and the cytoplasm. For example, the nuclear RNA exosome participates in the processing and degradation of rRNA precursors, as well as in the degradation of aberrant tRNA and mRNA transcripts, along with RNA generated from enhancer elements (eRNAs) and promoter upstream regions (PROMPTs) (Bousquet-Antonelli et al. 2000; Houseley et al. 2006; Darnell 2013; Pan et al. 2015; Kilchert et al. 2016). Moreover, RNA degradation may occur in the cytoplasm in association with translation through mechanisms such as nonsense-mediated decay (NMD) (Losson and Lacroute 1979; Maquat et al. 1981; Schoenberg and Maquat 2012; Braunschweig et al. 2014), nonstop-mediated decay (NSD) (Frischmeyer et al. 2002; van Hoof et al. 2002) or no-go decay (NGD) (Doma and Parker 2006).

RNA turnover can be assessed by measuring the decay of steady-state RNA, either after transcriptional inhibition or following a chase of metabolically bromouridine-labeled steady-state RNA (Tani and Akimitsu 2012; Tani et al. 2012; Imamachi et al. 2014). While transcription inhibition with pharmacological inhibitors such as actinomycin D is stressful to cells, metabolic labeling of RNA with bromouridine for extended periods of time has minimal effects on transcription, post-transcriptional processing, and cell viability (Tani et al. 2012; Bedi et al. 2021; Huang et al. 2025). In approach-to-equilibrium experiments, actively synthesizing RNAs in cells are labeled for increasing durations until the labeled population reaches steady-state. Decay rates can then be determined by comparing the abundance of labeled RNA at different labeling times with

the steady-state RNA levels (Lugowski et al. 2018). Another widely used approach involves labeling nascent RNA with a short pulse of a uridine analog and subsequently comparing the abundance of nascent transcripts to that of the steady-state transcriptome. Techniques using this methodology include SLAM seq (Herzog et al. 2017), TimeLapse-seq (Schofield et al. 2018), and TUC-seq (Riml et al. 2017). Nano-ID can be used to assess isoform-specific turnover (Maier et al. 2020), while Dyrec-seq utilizes double metabolic labeling with bromouridine and 4-thiouridine (4sU) to measure both synthesis and degradation rates (Kawata et al. 2020). Finally, INSPEcT is a computational method that estimates RNA turnover rates from steady-state RNA-seq data, eliminating the need for metabolic labeling (Furlan et al. 2020).

When comparing nascent RNA synthesis to steady-state RNA levels in cells, it is assumed that RNA degradation follows first-order exponential decay kinetics (Lam et al. 2001). However, the kinetics of decay may be differentially affected by quality control processes that remove aberrant RNA, as well as by cytoplasmic degradation of mature mRNAs. To assess RNA dynamics without relying on steady-state RNA comparisons, we used BruChase-seq, which measures the abundance of different RNAs at various time points after pulse-labeling with bromouridine and an ensuing uridine chase (Paulsen et al. 2013; Paulsen et al. 2014). By collecting RNA at different time points after labeling, transcriptomes of defined ages can be studied and compared in an unbiased way (Rabani et al. 2011). In this study, Bru-labeled RNA was collected either 30 minutes after Bru-labeling (nascent RNA) or following a 2- or 6-hour chase in uridine.

Using Bru-seq, we have shown that co-transcriptional splicing removes only about half of the introns from pre-mRNA in human cells (Bedi et al. 2021). Similar

findings have been reported in studies using long-read sequencing of nascent RNA, where transcripts were found to be either fully spliced or completely unspliced, with a smaller group of transcripts showing a mixture of spliced and unspliced introns (Herzel et al. 2018; Drexler et al. 2020). However, after a 6-hour chase time, the transcripts are almost completely spliced. This scenario suggests that the intron-containing transcripts are either removed from the transcriptome by quality control mechanisms and degradation (Tian et al. 2020) or undergo post-transcriptional splicing (Drexler et al. 2020).

A recent study using the same Bru-seq and BruChase-seq datasets analyzed here reported that allele-specific genetic variants can affect RNA stability by altering conserved binding sequences of miRNAs and RNA-binding proteins (Huang et al. 2025). Many of the genes that have these allele-specific variants are involved in immune system diseases, linking altered regulation of RNA stability to human diseases (Huang et al. 2025). The objective of the current study was to generate extensive datasets describing nascent transcription, splicing kinetics, and RNA stability across 16 human cell lines. These datasets generated for ENCODE4 provide a comprehensive insight into the regulation of RNA processing and its role in coordinating gene expression to maintain cellular homeostasis.

RESULTS

Using BruChase-seq to assess RNA degradation dynamics across 16 cell lines

It is well established that certain RNA species are highly unstable, such as introns, enhancer RNAs (eRNAs), and promoter upstream transcripts (PROMPTs) produced

upstream of most active gene promoters in the divergent direction (Andersson and Sandelin 2020). The functions of eRNAs and PROMPTs remain poorly characterized; these transcripts are short (<4 kb) and are rapidly degraded by the RNA exosome (Magnuson et al. 2015; Kilchert et al. 2016), making their detection difficult in steady-state RNA isolations. Additionally, RNA is generated from transcriptional readthrough at the ends of genes (Rodriguez-Molina et al. 2023; McShane et al. 2024). To assess the relative turnover of these different classes of RNA, we performed nascent RNA Bru-seq and BruChase-seq. Newly synthesized RNA was captured at time zero (Bru-seq or 0h) or after a 2- or 6-hour uridine chase interval (BruChase-seq, 2h and 6h) across 16 human cell lines (Fig. 1A). The 2h chase period was selected to capture early degradation of RNA, which is expected to occur primarily in the nucleus, while the 6h chase period was chosen to capture the degradation of predominantly mature, spliced mRNA which is expected to take place in the cytoplasm. These experiments were performed on two independent growths of each cell line, and the resulting nascent RNA libraries were sequenced to a depth of approximately 50 million paired-end reads (Supplemental Table S1). To assess the relative amounts of Bru-labeled RNA synthesized by different cell lines during the 30 minute labeling period (Supplemental Fig. S1A), as well as the relative amounts of Bru-labeled RNA remaining after the 2h or 6h chase periods (Supplemental Fig. S1B), we used exogenous Bru-labeled RNA spike-in controls consisting of the *Drosophila melanogaster* (dm6) transcriptome and selected *in vitro*-transcribed *Arabidopsis thaliana* RNA. The rationale is that Bru-labeled cellular RNA and spike-in RNA compete for capture by anti-BrdU antibodies. Thus, the more Bru-labeled RNA synthesized by the cells, the less spike-in RNA will be recovered from

the sample. Using the inverse fraction of spike-in RNA reads as an estimate of Bru-labeling efficiency, we found that the cell lines differed significantly (Supplemental Fig. S1A). These differences may be due to variability in bromouridine uptake, homeostatic uridine concentration in cells, efficiency of cellular conversion of BrU to BrUTP, or inherent rates of RNA synthesis.

As RNA ages, the fraction of intronic signal is expected to decrease due to splicing and degradation, while the fraction of exonic signal would increase in relationship to all other RNAs. Compared to the stability of exons, intronic stability was only 0.23-fold and 0.11-fold after a 2-hour and 6-hour chase period, respectively (Fig. 1B, Supplemental Table S2). PROMPTs and eRNAs showed degradation patterns similar to intronic sequences, whereas transcriptional readthrough RNA was comparatively more stable. These results show that different types of RNA species undergo degradation with different kinetics.

We next compared the fraction of reads mapped to exons, introns, exon antisense, intron antisense, and intergenic sequences across the 16 cell lines. In addition, ambiguous reads, referring to reads that could not be unambiguously mapped, were included in the analysis. At time 0h, the fraction of intronic signal accounted for 0.72-0.87 and exonic reads accounted for 0.09-0.13 of total reads across the cell lines. The intergenic (unannotated) read fraction made up 0.05-0.09, and the ambiguous read fraction 0.03-0.04 of all reads (Fig. 1C, Supplemental Table S3). The fraction of antisense RNA comprised less than 0.02 of all the reads. Following the 2h and 6h chase periods, the fraction of exonic reads increased and that of intronic reads decreased, while the fractions of antisense, intergenic, and ambiguous reads remained

relatively constant. We observed a large discrepancy in intronic signal across the 16 cell lines at the later time point. For example, after a 6-hour chase, the fraction of intronic signal was much higher in IMR90 cells (0.61) than in Panc1 cells (0.28), a result that was highly consistent across both biological replicates (Supplemental Table S3). This may suggest that either splicing or intron turnover rates differ between cell lines.

To explore whether differences in intronic abundance among the 16 cell lines correspond to splicing efficiency, we estimated the splicing index (SI) of common introns expressed in all 16 cell lines, as previously described (Bedi et al. 2021) and detailed in the Materials and Methods section. Co-transcriptional SI, defined as a measure of splicing occurring during the 30-minute bromouridine-labeling period (0h), showed that the median SI values ranged from 0.30 in PC9 cells to 0.61 in HMEC cells (Fig. 1D, Supplemental Table S4). Surveying the RNA two hours after labeling (2h) showed a marked increase, with median SI values rising to 0.75 for K562 cells and to approximately 0.87 or higher for most of the other cell lines. Finally, following a 6-hour chase, median SI values increased to over 0.9 for most cell lines. These results indicate that, on average, only about half of the splice junctions flanking introns are co-transcriptionally spliced, although splicing across the transcriptome becomes more complete over time. This may result from either post-transcriptional splicing activities or from RNA degradation machinery purging the transcriptome of intron-containing transcripts. Indeed, it has been shown using BruChase-seq technology that intron-containing transcripts are considerably less stable than spliced transcripts (Tian et al. 2020), suggesting targeted degradation of unspliced transcripts. The high abundance of introns in IMR90 cells after a 6-hour chase period (61%) does not correspond to a

conversely low splicing index (median SI > 0.87), suggesting that these introns are spliced out but not efficiently degraded during the chase period in this cell line. We surmise that, across cell lines, the remaining intronic signal originates either from introns that have been spliced out but not yet degraded by nucleases, or from introns with non-coding RNA functions protected from degradation. If transcription initiation within these introns were generating small ncRNAs, we would expect higher read counts in the introns compared to the gene; however, we did not observe such a pattern.

Cell type and pathway-specific RNA dynamics

To compare gene expression across the different cell lines, we calculated normalized counts for exons (Reads Per Kilobase per Million mapped reads, RPKM) at 0, 2, and 6 hours for RNA derived from genes of all biotypes (Supplemental Table S5). To evaluate Bru-labeled RNA degradation over time in the 16 cell lines, we determined the relative RNA stability of expressed genes, measured as the $\log_2$ fold change, during the early (0-2h) and late (2-6h) chase periods. Only exonic reads were used, as introns are rapidly degraded post-splicing and display turnover kinetics distinct from those of exons in mature mRNA (see Fig. 1B). We found that correlations in RNA stabilities among the 16 cell lines were generally stronger at the 2h chase than at the 6h chase period, and that the grouping of cell lines based on RNA stability values for 17,049 genes, as revealed by hierarchical clustering, largely differed between the two chase periods (Fig. 2A,B).

To assess whether turnover rates for protein-coding gene transcripts change over time, we compared the relative stability of individual genes during the early and late chase periods. We selected the Panc1 cell line to showcase our results, as it showed low intron abundance and a high splicing index at 6 hours compared to other cell lines (Fig. 1C,D). Using an expression threshold of >0.5 RPKM in the 0h sample, we found that the relative stability increased significantly for 161 transcripts and decreased significantly for 114 transcripts in Panc1 cells during the first 2 hours after synthesis ($-\log_{10}p_{adj} > 2$, $\log_2FC > \pm 1.5$) (Fig. 2C, Supplemental Table S6). During the 2h to 6h period, relative stability increased significantly for 66 transcripts and decreased significantly for 885 transcripts (Fig. 2D, Supplemental Table S6). In many of the 16 cell lines, there are more transcripts with higher relative turnover between 2 and 6 hours than between 0 and 2 hours (Supplemental Fig. S2). Based on the read coverage profile for *FAF2* (Fig. 2E), reads are distributed across the entire gene at 0 hours but become concentrated in the exonic regions at 2 and 6 hours. This gene exemplifies transcripts that show high relative stability across both time periods. In contrast, *DUSP1* transcripts exhibit low relative stability at both time points (Fig. 2F), and *MAT2A* transcripts display high relative stability during the first 2 hours but are extensively degraded during the subsequent 4h chase period (Fig. 2G). These findings demonstrate that degradation rates differ between transcripts and do not necessarily follow first order kinetics (Palumbo et al. 2015).

To better understand which cellular processes are associated with relative RNA stability, we performed Gene Set Enrichment Analysis (Subramanian et al. 2005), ranking genes using regularized log transformation (rlog) in DESeq2 (Love et al. 2014)

for comparisons performed between 0 and 2 hours, and between 2 and 6 hours. In Panc1 cells, Hallmark gene sets such as “androgen response” and “mitotic spindle” were enriched among high-stability transcripts following the early chase period (0-2h) (Fig. 2H, Supplemental Table S6), while “MYC targets V1” and “protein secretion” were enriched after the late chase period (2-6h) (Fig. 2I). Gene sets enriched among low-stability transcripts included “TNFA signaling via NFkB” and “oxidative phosphorylation” during the early period (Fig. 2J), and “TNFA signaling via NFkB” and “p53 pathway” for the late period (Fig. 2K). Genes in KEGG pathways such as “ribosome”, “oxidative phosphorylation” (mitochondria), “proteasome” and “spliceosome” produced transcripts that were relatively unstable during the early chase period but showed high stability during the late period. (Fig. 2L). This trend was consistent when analyzing transcript stability of these pathways across all 16 cell lines (Supplemental Fig. S3A).

Several Hallmark gene sets were consistently enriched for either high or low transcript stability across the 16 cell lines during the early chase period (0-2h) (Supplemental Fig. S3B). Examples of gene sets enriched for high stability RNA include “androgen response” (NES > 1 in all cell lines) and “mitotic spindle” (NES > 1 in 15/16 cell lines), whereas gene sets enriched for low stability RNA include “TNF signaling via NFkB” (NES < -1 in 14/16 cell lines) and “oxidative phosphorylation” (NES < -1 in 15/16 cell lines). Transcripts exhibiting high relative stability during the 2-6h chase period included “MYC targets” and “oxidative phosphorylation”, whereas “TNFA signaling via NFkB” and “hypoxia” gene sets are associated with high turnover. We observed a similar enrichment trend for KEGG pathways among low stability transcripts, with coordinated patterns across the cell lines (Supplemental Fig. S3C). These results

suggest that transcripts coding for proteins belonging to a specific pathway tend to share similar turnover rates, and that this pattern is consistent across cell lines.

Post-transcriptional regulation of mitochondrial encoded transcripts

The circular mitochondrial genome is transcribed into three polycistronic transcripts that include 2 rRNA genes, 22 tRNA genes and 13 protein-coding genes required for the mitochondrial oxidative phosphorylation system (OXPHOS) (Anderson et al. 1981). These polycistronic transcripts are processed into individual transcripts via the “tRNA punctuation model” (Ojala et al. 1981), although additional cleavage sites have also been observed in these transcripts (Mercer et al. 2011). It has been noted that individual mRNAs levels are regulated by differential degradation (Piechota et al. 2006), and that Suv3p helicase is involved in processing and turnover of mitochondrial transcripts (Szczesny et al. 2010). In our analysis pipeline, mitochondrial reads are mapped to the mitochondrial genome but are filtered during the pre-processing step (see Materials and Methods, alignment files are available on the ENCODE portal). Analysis of Bru-seq and BruChase-seq data from the 16 cell lines reveals that expression of the mitochondrial genes is differentially regulated at the post-transcriptional level. The distribution of read counts across the mitochondrial genome shows that *ND1*, *ND2* and *ND3* genes produce relatively unstable transcripts, while *CO1*, *CO3*, *ATP6*, *ATP8* and *ND4* genes generate relatively stable transcripts (Supplemental Fig. S4). However, K562 and PC3 cells exhibit higher stability of the *ND2* transcript compared to all other cell lines. Consistent with a previous study, we observed antisense RNA expression for the *ND5* gene (Mercer et al. 2011). This transcript is generally very unstable, except in A673, Caco2,

HUVEC and K562 cells, where this ncRNA was relatively stable (Supplemental Fig. S4). Thus, our data suggest both gene- and cell type-specific post-transcriptional regulation of mitochondrial gene expression.

Unique RNA turnover dynamics for individual genes

We next compared the degradation dynamics of transcripts from individual genes across the 16 cell lines. In Figure 3A, we plotted read coverages for four distinct genes in Panc1 cells that demonstrate widely different degradation dynamics to highlight differences in relative transcript stability. To compare the degradation trajectories of individual gene transcripts during the two chase periods across the 16 cell lines, we scaled the $\log_2$ FC relative stability values between 0 and 1 for each cell line (Supplemental Table S7). When we plotted the scaled RNA stabilities for the two chase periods for the *MTPN* gene, we found that the data points from all 16 cell lines clustered in the upper right corner, suggesting high stability across both time intervals (Fig. 3B). In contrast, transcripts from the *MYC* gene show a high degree of degradation during both time periods across all 16 cell lines. Both *MTPN* and *MYC* transcripts converge around the diagonal, indicating that their degradation rates were comparable during both time intervals. On the other hand, most data points for the *TRMT10C* transcript fell below the diagonal, reflecting high initial stability followed by lower stability, whereas the scaled stability values for *GAPDH* fell above the diagonal, indicating fast initial turnover followed by increased stability (Fig. 3B). These results show that some transcripts exhibit different degradation rates during the first 2 hours compared to the following 4 hours after synthesis, suggesting that not all transcripts follow first-order degradation

kinetics (Palumbo et al. 2015). The distributions of scaled RNA stability values for 4,415 expressed genes in Panc1 cells are shown for both the first 2 hours (Fig. 3C) and the subsequent 4 hours (Fig. 3D). The distribution of scaled stabilities for 4,415 genes at the 2h and 6h chase periods is shown for individual cell lines in Supplemental Fig. S5, and for all 16 cell lines combined, along with comparisons of variance (standard deviation), in Supplemental Fig. S6.

The patterns of transcript stability were generally similar across the cell lines, suggesting that the transcript, rather than the cell type, is the primary determinant of these degradation patterns. Additional examples of transcripts showing similar RNA dynamics across the cell lines are shown in Supplemental Fig. S7A. However, we identified exceptions in which transcripts exhibited cell type-specific RNA stability trajectories (Fig. 3E,F and Supplemental Fig. S7B). We next explored whether splicing efficiency may influence RNA stability. We assessed the correlation between splicing index at 0 hours and transcript stability during the 0-2h time interval for 2,693 genes across the 16 cell lines (Supplemental Table. S4). We found no correlation between splicing efficiency and transcript stability across the cell lines (Supplemental Fig. S8). We then averaged the splicing and stability values and calculated the standard deviation across the cell lines, identifying two genes that showed high variation in both parameters (Supplemental Fig S8, lower right). Both *RPL22L1* (Fig. 3E) and *MRPL17* (Fig. 3F) demonstrated a moderate correlation between splicing efficiency and transcript stability. The *MRPL17* transcript exhibited particularly low stability in K562 cells (Fig. 3F, left), where it was transcribed at a very high rate at 0 hours (Fig. 3G); however, Bru- RNA levels decreased to an average level by 6 hours. It is possible that the high

transcription rate of *MRPL17* in K562 cells overwhelms the splicing capacity, resulting in poorly spliced transcripts that are subsequently degraded. While poor splicing was associated with low stability for some transcripts, this was clearly not a general trend. It is possible that using the average splicing index across a transcript is not suitable for assessing a transcript's degradation trajectory. Perhaps the splicing index of the most poorly spliced intron, together with the contributions of RNA-binding proteins and miRNAs, may more accurately reflect a transcript's propensity for degradation and should be explored in future studies (Jens and Rajewsky 2015).

We next plotted the scaled stability values for transcripts coding for proteins within the same functional pathways (Supplemental Fig. S9). For each transcript, we averaged the stability values from the 2h vs 0h and 6h vs 2h comparisons across all 16 cell lines. For most pathways, the scaled stability values were centered around the diagonal, suggesting similar degradation rates across both time periods (Supplemental Fig. S9). In contrast, for the 64 transcripts expressed across all 16 cell lines and representing the GSEA “KEGG ribosome” pathway, the degradation trajectories differed between the two time points, with the initial chase period showing a higher rate of degradation than the later period. Similar patterns were also observed for transcripts representing the “oxidative phosphorylation (OXPHOS)”, “proteasome” and “spliceosome” KEGG pathways. These findings are consistent with the data presented in Figure 2L and Supplemental Fig. S3C. Moreover, when examining specific genes from these pathways, we observe high concordance across the 16 cell lines, such that individual transcripts show higher degradation in the first 2h chase period than in the following 4 hours (Supplemental Fig. S10). These data suggest that transcript stability is

most often transcript-specific, displaying similar dynamics across the cell lines, although some cell type-specific regulation was observed.

Isoform-specific post-transcriptional processing

Most human genes generate multiple transcript isoforms through the use of alternative promoters, splice junctions or poly(A) sites (Marasco and Kornblihtt 2023). While the selection of a specific promoter for transcriptional initiation is regulated by transcription factor binding and epigenetic marks (Neph et al. 2012; Villar et al. 2014; Stricker et al. 2017), the choice of splicing junctions is determined by proteins binding to splicing enhancer and silencer sequences within the RNA molecule (Marasco and Kornblihtt 2023). What is less understood is whether isoform-selective RNA turnover contributes to the emergence of a dominant isoform from a pool of synthesized isoforms. To address whether isoform-selective post-transcriptional processing occurs in human cells, we analyzed the Bru-seq (0h) and BruChase-seq (2h and 6h) data using the RSEM software tool (Li and Dewey 2011).

We expressed the isoform abundances as a percentage of all isoforms for each gene and time point. Both *ETV3* and *ATF4* genes (Figs. 4A,B,C) show opposing abundance patterns for their respective isoforms across most of the 16 cell lines, with one isoform's abundance decreasing and the other's increasing over the time course. The abundances of *NDUFS5* isoforms vary between cell lines (Fig. 4D). While both isoforms of the gene appear to coexist at similar levels throughout the time course in 10 cell lines, isoform #1 is dominant in the remaining 6 cell lines. This dominance is established within the first 30 minutes of production (0h) and is maintained throughout

the chase periods. *HNRNPR*, which has 9 isoforms, also shows differential stability of its isoforms, with some increasing in abundance and others decreasing over the time course (Fig. 4E). Using an isoform percentage of 50% as the threshold for annotating an isoform as dominant (or major), approximately 30% of the genes we studied showed a pattern in which one isoform established dominance at 0 hours and remained dominant throughout the time course (“always dominant”, AD) (Fig. 4F). An example of this is isoform #1 of *NDUFS5*, which was dominant in 6 cell lines (Fig. 4D). Around 20% of the genes had an isoform that gained dominance (GD), such as isoform #2 of *ETV3* and *ATF4*, starting below 50% prevalence but rising above this level during the time course. In addition, roughly 15% of the genes had an isoform that lost dominance (LD) over time. *ETV3* and *ATF4* are examples of genes that are classified under both GD and LD categories. The abundance patterns for the remaining genes could not be clearly delineated. These results suggest that many genes initially generate multiple isoforms, which are then regulated by isoform-specific turnover to yield the desired stable isoform.

Intron retention by lack of splicing or stable spliced intron

Introns are spliced co-transcriptionally from nascent RNA as it exits the RNA polymerase (Oesterreich et al. 2016; Rogalska et al. 2023). Thus, it is expected that introns are removed in a sequential manner as the RNA molecules exit the RNA polymerase. However, several studies have reported that the order of splicing does not always follow the direction of transcription (Drexler et al. 2020; Choquet et al. 2023; Choquet et al. 2025a), and that co-transcriptional splicing is often incomplete (Bedi et al.

2021). To assess whether transcripts show consistent splicing index (SI) values for individual introns, or if each intron is removed independently of each other, we analyzed multi-intron genes that produce a single annotated isoform with RPKM ≥ 0.5 at 0 hours (n=626 genes, 7,231 introns) (Supplemental Table S8).

When analyzing the SI values of individual multi-intron transcripts, we found that splicing efficiency differed markedly along the transcript. Transcripts synthesized from the *DUSP12* gene in the A673 cell line show that most introns remain after the 30-minute Bru-labeling of nascent RNA and are not spliced out evenly (Fig. 5A). Following a 2h chase period, introns 1-4 are almost completely spliced out (SI > 0.93) and subsequently degraded. In contrast, intron 5 is poorly spliced even at the 2h time point but is efficiently spliced and degraded by 6 hours (Fig. 5B). This example suggests that co-transcriptional splicing is not coordinated across all introns of this transcript, and that one intron is initially retained due to incomplete splicing. On the other hand, intron 4 in the *STAM* gene is spliced out by the 2h chase time (Fig. 5C,D), yet persists even in the later chase period. For annotation, we used GENCODE v29, which does not indicate the presence of a non-coding RNA in intron 4 of the *STAM* gene. A similar pattern is observed for this intron across all 16 cell lines (Supplemental Fig. S11A,B). It is possible that, after splicing, this intron serves a non-coding function in the cells and is therefore not targeted for immediate degradation. To assess the frequency of stable introns in our dataset of single isoform transcripts across the 16 cell lines, we analyzed introns with a SI value at the 6-hour time point greater than the mean for that cell line, and a fold change in read density (6h vs 0h) at least 3 standard deviations above the mean for that

cell line (Supplemental Table S9). We identified 404 introns across the 16 cell lines that were considered stable according to our chosen cut-off values (Supplemental Figure S12). We found that 128 stable introns were unique to a specific cell type, while only one stable intron was shared among 12 cell lines. We also found evidence for introns in single isoform genes that showed minimal splicing at the 6h time point, and these were termed “retained introns”. These were defined as introns with SI values at the 6h time point at least three standard deviations below the mean for that cell line, and a fold change in read density (6h vs 0h) at least 3 standard deviations above the mean for that cell line. We found 388 retained introns across the 16 cell lines, where 51 were unique to a single cell line, while four introns were retained across 14 cell lines (Supplemental Figure S13).

Binary co-transcriptional splicing pattern

Long read sequencing has shown that many transcripts exhibit a binary “all or none” splicing pattern (Herzel et al. 2018; Drexler et al. 2020; Reimer et al. 2021; Merens et al. 2024). Since Bru-seq assays were performed using 150 nucleotide paired-end sequencing of fragments 500-600 bp in length, it was possible to use this data to assess two splicing events around an exon on the same sequenced fragment. While focusing on single-isoform genes, we compiled the splicing pattern categories (Fig. 6A) around 47,163 exons from 4,578 genes. The two largest groups of splicing categories observed in all 16 cell lines were “spliced-spliced” and “unspliced-unspliced” (Fig. 6B, Supplemental Table S8). We also observed that, of the two smaller groups (“spliced-unspliced” and “unspliced-spliced”), about twice as many exons had only the upstream

434 intron spliced compared to exons with only the downstream intron spliced. To explore
435 whether splicing efficiency correlates with transcription level (exonic RPKM), individual
436 transcripts with three different expression levels in HCT116 cells were selected. *VMP1*,
437 which had the highest transcription of the three genes, showed very inefficient splicing
438 across its 12 exons (Fig. 6C). *AATF*, which had an intermediate expression level,
439 showed a balance between fully spliced and unspliced states across its 12 exons.
440 Finally, *HDAC1* showed the most complete splicing around its 14 exons, while its
441 expression was the lowest of the three genes. This trend of an inverse association
442 between splicing efficiency and expression level was also observed for *VMP1* when
443 splicing patterns were plotted for each of the 16 cell lines, ordered by increasing
444 expression (exonic RPKM) (Supplemental Figures S14-S16). However, this trend was
445 not observed for the other two genes when examining the relationship across all cell
446 lines. Finally, upon evaluating the distribution of the different splicing patterns following
447 a 6h chase, we found that the “spliced-spliced” fraction increased at the expense of the
448 “unspliced-unspliced” fraction (Fig. 6D, Supplemental Tables S10-S12). Taken together,
449 these findings suggest that co-transcriptional splicing occurs in a binary manner by two
450 major modes of transcription: one that leads to complete splicing and one that leads to
451 no splicing. Our study concurs with and extends earlier findings (Herzel et al. 2018;
452 Drexler et al. 2020; Reimer et al. 2021) by demonstrating that this pattern holds true
453 across 16 human cell lines when examining splicing around 47,163 exons. Furthermore,
454 for at least some transcripts, we show that the fraction of unspliced transcripts
455 correlates with high transcription rate, and that this unspliced fraction rapidly diminishes
456 over time. Whether this increase in splicing index over time is due to post-transcriptional

splicing (Choquet et al. 2025a; Choquet et al. 2025b) or targeted purging of the pool of unspliced transcripts (Herzel et al. 2018) will need to be elucidated in future studies.

DISCUSSION

RNA turnover is a critical component in the regulation of gene expression. While substantial attention has been given to mechanisms of transcription initiation mediated by transcription factors and epigenetic modifications, less is known about how RNA turnover is regulated (He et al. 2023). In this study, we used Bru-seq and BruChase-seq technologies to assess nascent RNA synthesis, splicing and degradation. BruChase-seq differs from other methods used to study RNA degradation in that it does not depend on comparisons of nascent RNA synthesis with steady-state RNA levels. Instead, Bru-labeled RNA is tracked as it ages during uridine chase periods, after which it is captured, and relative RNA abundance is assessed genome-wide. While other studies assessing steady-state cellular RNA have shown that degradation occurs through first-order kinetics, our BruChase-seq approach revealed clear exceptions to this premise. Performed across 16 human cell lines as part of the ENCODE4 project, our study showed that several transcripts exhibited different degradation kinetics during the two chase time periods (2 or 6 hours). Additionally, we found that RNA degradation varies widely among transcripts, is mostly transcript-dependent, and generally does not associate with splicing efficiency.

As expected, intronic sequences showed a rapid turnover profile compared to exonic sequences (Fig. 1B). The same was observed for readthrough transcripts, PROMPTs and eRNAs. The rapid turnover of the vast amount of intronic RNA is critical

for the recycling of ribonucleotides, enabling their participation in new RNA synthesis. We found that cell lines differed considerably in the turnover of intronic sequences, with the normal fibroblast IMR90, for example, retaining over 60% of intronic signal after 6 hours of chase (Fig. 1C). This high intronic signal in IMR90 was not due to a lack of splicing but is rather likely the result of slow degradation of spliced introns. Co-transcriptional splicing indices ranged from 0.3 to 0.6 across the 16 cell lines but increased to nearly 0.9 after a 2h chase period (Fig. 1D).

It has been previously suggested that physiological logic governs the decay rates of functionally linked transcripts (Lam et al. 2001; Wang et al. 2002; Yang et al. 2003; Raghavan and Bohjanen 2004; Sharova et al. 2008; Friedel et al. 2009; Rabani et al. 2011; Schwanhausser et al. 2011). This pattern is observed for transcripts encoding multimeric protein complexes in both yeast (Wang et al. 2002) and in human cells (Friedel et al. 2009). Moreover, transcripts associated with regulatory functions, such as transcription factors, typically have short half-lives, enabling cells to respond rapidly to environmental stimuli and challenges (Yang et al. 2003; Sharova et al. 2008; Rabani et al. 2011). This trend was observed for some, but not all, transcripts coding for transcription factors, with *KLF6*, *KLF10* and *MYC* being very unstable, and *TP53*, *HIF1A* and *SMAD5* being relatively stable (manuscript in preparation). We observed a striking pattern of RNA degradation for transcripts encoding protein complexes such as the ribosome, proteasome, spliceosome and oxidative phosphorylation machinery, where the initial rate of degradation was extensive but dramatically slowed during the later chase period. (Fig. 2L, Supplemental Fig. S3A, S8, S9). Although speculative, it is possible that the cellular strategy for obtaining stoichiometric amounts of subunits for

multimeric protein complexes involves synthesizing these transcripts in excess, followed by rapid degradation of this transcript pool to generate balanced amounts of transcripts for translation.

Most genes in eukaryotic cells produce multiple transcript isoforms through alternative splicing or selection of alternative promoters. Alternative splicing can lead to frameshifts, resulting in rapid degradation, or give rise to unique proteins that contribute to the diversification of the proteome (Marasco and Kornblihtt 2023; Rogalska et al. 2023). The mechanisms that regulate alternative splicing are not fully understood, but specific sequences within exons and introns are thought to attract or block the binding of RNA-binding proteins involved in the splicing process (Wang and Burge 2008). In addition, epitranscriptomic modifications of pre-mRNA and sequence-mediated structural features can affect splice-site selection, leading to the generation of preferred transcript isoforms (Rogalska et al. 2023). Cells may also generate a plethora of transcript isoforms from a gene, subsequently removing unwanted isoforms through selective degradation. Using the BruChase-seq approach, we observed several genes producing multiple isoforms at 0 hours, followed by the increased prevalence of one or more isoforms at the expense of others at 2 and 6 hours (Fig. 4). This finding suggests that a prominent mechanism of isoform selection involves selective post-transcriptional degradation of unwanted isoforms. Future studies are needed to further elucidate the mechanisms underlying this selective degradation.

Splicing is thought to occur immediately after nascent RNA exits the RNA polymerase, when splice sites become exposed to the splicing machinery (Oesterreich et al. 2016). However, sequencing of nascent RNA has provided evidence that many

introns remain in transcripts long after the RNA polymerase has finished transcription (Herzel et al. 2018; Drexler et al. 2020; Bedi et al. 2021; Reimer et al. 2021). It has also been noted that the order of splicing does not always correspond to the sequential order of introns in the transcript (Drexler et al. 2020; Choquet et al. 2023; Choquet et al. 2025a). Our previous study using Bru-seq technology corroborated these findings, showing that only about half of all introns are co-transcriptionally spliced and that splicing does not always occur in a linear order (Bedi et al. 2021). Here, we extended these studies by analyzing splicing patterns across 16 cell lines and measuring the splicing index of transcripts over time. We found examples of introns that showed poor splicing even 2 hours after synthesis, while neighboring introns were efficiently spliced (Fig. 5A, Supplemental Fig. S13). We also identified introns that were efficiently spliced at 2 hours but whose intronic signal remained (Fig 5C, Supplemental Fig. S12), suggesting selective retention of these introns, perhaps because they provide a specific function for the cell.

We performed 150 bp paired-end sequencing on fragments size-selected at approximately 500-600 bp in our Bru-seq and BruChase-seq studies. Since many exons are around 150 nucleotides long, we were able to examine two splicing events on the same molecule on either side of 47,163 exons. Compiling splicing events for the four possible combinations of splicing around these exons, we found that most molecules were either “spliced-spliced” or “unspliced-unspliced” across the 16 cell lines (Fig. 6B). This finding is consistent with previous studies using long-read sequencing of nascent RNA that have demonstrated a binary splicing pattern (Herzel et al. 2018; Drexler et al. 2020; Kwon et al. 2021). Examining the relationship between splicing efficiency and

expression levels, we did not find an association between efficient splicing and high expression. This finding contrasts with other studies that suggest an “economy of scale”, in which RNA polymerase II and highly transcribed genes are located near nuclear speckles where efficient splicing takes place (Ding and Elowitz 2019; Bhat et al. 2024). When assessing splicing over time, the fraction of “spliced-spliced” molecules increased at the expense of the “unspliced-unspliced” fraction, which diminished over time. One interpretation of this increase in splicing index is that these transcripts are spliced post-transcriptionally. Alternatively, all splicing may occur co-transcriptionally, while transcripts that remain unspliced undergo a time-dependent purging to eliminate them from the transcriptome. Why would cells expend time and energy transcribing genes into pre-mRNA that is not spliced and destined for degradation? A possible explanation is that this excessive transcription is utilized for DNA damage scanning in a mutation-suppressing mechanism, as previously proposed (Ljungman 2022). How can it be reconciled that intron splicing in a transcript could have very different kinetics, as shown for *DUSP12* in Fig. 5, while most splicing occurs in an all-or-none fashion as shown in Fig. 6? It may be that, for transcripts spliced co-transcriptionally, some introns are initially detained and require a cellular signal to be post-transcriptionally spliced and exported to the cytoplasm for translation (Boutz et al. 2015).

These studies using Bru-seq and BruChase-seq across 16 cell lines as part of ENCODE4 provide comprehensive datasets on RNA synthesis and degradation in human cells. Additionally, many complementary transcription and chromatin assays were performed on the same cell growths, making it possible to directly compare data from many different assays. We show here examples of the complexity of RNA

dynamics in human cells, as well as clear patterns of cellular logic in the post-transcriptional regulation of gene expression. Known regulators of RNA stability include miRNAs and RNA-binding proteins, and in an accompanying study based on these Bru-seq and Bru-Chase datasets, allele-specific SNVs overlapping binding sites of miRNAs and RNA-binding proteins were associated with altered RNA stability (Huang et al. 2025). Importantly, genes with allele-specific SNVs alter RNA stability function in many disease pathways, highlighting the importance of accurate RNA stability regulation for normal homeostasis. Regulation of RNA stability is an understudied area, and future studies will certainly uncover the importance of precise regulation of RNA dynamics for maintaining cellular homeostasis, as well as elucidate the ways in which defects in this regulation can contribute to the development and progression of human disease.

MATERIALS AND METHODS

Growth conditions for Cell lines

The cell lines analyzed in this study are part of the ENCODE4 Deeply Profiled Cell Lines collection and were cultured as two independent growths (Supplemental Table S13, McShane et al. 2024). Cells were grown according to manufacturer's recommendations in a humidified incubator at 37°C and 5% CO₂, and were harvested at approximately 80% confluency between passages 6 and 12. For consistency, cell line names are reported without dashes or spaces in this study.

Bromouridine labeling, Uridine Chase, and sequencing library preparation

594 Cells were incubated in 2 mM bromouridine (MilliporeSigma) for 30 min to label nascent
595 RNA, then immediately lysed (0h) or chased in 20 mM uridine (MilliporeSigma) for either
596 2 hours (2h) or 6 hours (6h), as previously described (Paulsen et al. 2013; Paulsen et
597 al. 2014). Cells were then lysed in TRIzol reagent (Invitrogen), total RNA was isolated,
598 and a spike-in cocktail consisting of Bru-labeled and unlabeled RNA was added
599 (Magnuson et al. 2022). Bru-labeled RNA was then captured from total RNA using anti-
600 BrdU antibodies (BD Biosciences). Stranded library preparation was performed
601 according to protocols linked to each experiment on the ENCODE portal
602 (<https://www.encodeproject.org/>) and as detailed elsewhere (McShane et al. 2024).
603 Experimental metadata related to each ENCODE library is also available on the
604 ENCODE portal.

605

606 **Alignment parameters**

607 Alignment was carried out as previously described (McShane et al. 2024). Briefly, Bru-
608 seq and BruChase-seq libraries were sequenced on the Illumina NovaSeq 6000
609 platform (150 bp paired-end reads) to a depth of approximately 50 million read pairs per
610 sample. Raw reads were trimmed using BBDuk (BBTools, v38.46) (Bushnell 2014) to
611 remove adapter sequences and then pre-aligned to the ribosomal RNA (rRNA)
612 repeating unit (GenBank U13369.1), the mitochondrial genome (chrM), EBV genome
613 (from the hg38 analysis set), and selected spike-in sequences (*Drosophila*
614 *melanogaster* (dm6) rRNA and dm6 mitochondrial genome) using Bowtie 2 (v2.3.3)
615 (Langmead and Salzberg 2012). Reads aligning to these sequences were recorded and
616 removed. The remaining reads were then aligned to the reference genome (GRCh38,

GENCODE v29 annotation) and specific spike-in references (dm6 genome, *E. coli* genome (K-12 MG1655), and selected *Arabidopsis thaliana* gene sequences) using STAR (v2.7.0f) (Dobin et al. 2013).

Estimation of Bru-labeling efficiency

The RNA spike-in cocktail was prepared and added to the Bru-seq and 2h BruChase-seq samples according to their total RNA content, as previously described (Magnuson et al. 2022). For the 6h BruChase-seq assay, 1/4th of these amounts was used. The relative Bru-labeling efficiency for each cell line was calculated as the inverse of the average “fracBru” statistic (fraction of reads mapping to spike-in sequences, Supplemental Table S1) from replicate samples. For comparison across time points, inverted “fracBru” values at 2 and 6 hours were normalized to the 0 hour sample for each of the 16 cell lines. These values were averaged, and the standard deviation was determined (n=16). For the 6h chase samples, the “fracBru” values were multiplied by 4 to account for the reduced spike-in quantity.

RNA abundance calculation for features of interest

Uniquely mapping reads (MAPQ $\geq$ 255) to intergenic regions, introns and exons on either the sense or antisense strand were counted using custom scripts. The proportion of these reads relative to all uniquely mapping reads in the standard chromosomes was calculated. Reads that mapped to more than one gene in the standard chromosomes were assigned to the ambiguous class. Promoter upstream regions, intergenic enhancer regions, and readthrough (RT) segments were obtained similarly to those described

640 elsewhere (McShane et al. 2024). Briefly, PROMPTs (promoter upstream transcripts)
641 were characterized within a 2 kb divergent region located upstream of a subset of
642 protein-coding genes and were required not to overlap an annotated feature
643 (GENCODE v29, any biotype) on the divergent strand. Intergenic enhancers were
644 obtained from BruUV-seq peak calls that were intersected with the GENCODE v29
645 gene annotation to remove those overlapping annotated genes. Intergenic RT segments
646 were extracted using classifications defined by McShane et. al. (McShane et al. 2024)
647 where segments were retained if they were tagged as classes Ia or Ib, meaning they did
648 not overlap a gene on the same strand.

649 Counts for protein-coding gene features (intron_sense and exon_sense),
650 PROMPTs, eRNA and RT regions were obtained by calculating the fractional coverage
651 of sequenced reads at each base in a strand-specific manner, and then summing these
652 values along the entire feature. These values were then added for each biological
653 replicate at all three time points for 14 of the 16 cell lines, excluding HMEC and HUVEC
654 (Supplemental Table S2). For each cell line and time point, normalized counts
655 (*sample_norm_cts*) for each feature were calculated by dividing the total feature counts
656 by the number of uniquely mapped reads to the genome. For each feature at a given
657 time point, the *sample_norm_cts* values from 14 cell lines were summed to obtain the
658 total normalized counts (*sum_norm_counts*). The relative stability of intron_sense RNA,
659 RT RNA, PROMPTs and eRNA (represented as “feature” in the equation below) at the
660 2h and 6h time points was calculated using exon_sense RNA as the reference,
661 according to the following formula:

Relative stability of a feature

$$= \frac{\text{feature sum_norm_counts at 2h or 6h} / \text{feature sum_norm_counts at 0h}}{\text{exon_sense sum_norm_counts at 2h or 6h} / \text{exon_sense sum_norm_counts at 0h}}$$

Measurement of splicing metrics

Splicing Index calculations were performed as previously described (Bedi et al. 2021).

Briefly, an intron-centric splicing index was calculated for common introns across the 16 cell lines in both Bru-seq and BruChase-seq assays, after merging the reads from the two replicates per cell line per assay. Splicing Index (SI) was defined by the formula:

$$a/[a + (b + c)/2]$$

where a = split reads (i.e., reads spanning exon-exon junctions); b = 5' exon-intron junction reads (with a minimum of 10 bp overlap on either side of the junction); c = 3' intron-exon junction reads (with a minimum of 10 bp overlap on either side of the junction). The junction read sum (jrsum) was defined as $a + b + c$, and a threshold of ≥ 10 was applied. Gene-level splicing index was calculated as the mean of all intronic splicing indices for that gene.

The various splicing categories, defined by examining the upstream and downstream introns flanking each exon, were identified by reads spanning the associated exon-exon or exon-intron boundaries. The number of counts for each splicing category at the three time points can be found in Supplemental Tables S10-12. For downstream analyses, only those categories in which events were captured for both the upstream and downstream introns on either side of an exon were considered, as shown in Fig 6A.

684

685 **Dynamics of protein-coding RNA**

686 To compare the transcript degradation trajectories across the two chase periods in 16
687 cell lines, reads were counted in exons only ("exonic") for each gene in Bru-seq (0h)
688 and BruChase-seq (2h and 6h) samples. For each cell line, exonic counts from two
689 biological replicates per condition were normalized and compared using DESeq2
690 (DESeq2 v1.18.1, R v3.4.3) (Love et al. 2014). The resulting $\log_2$ fold change values
691 were used to represent transcript stability between the 0h and 2h, and 2h and 6h time
692 points. Genes not rejected by independent filtering in any time point comparison
693 (DESeq2) for all 16 cell lines were retained (17,049 genes for exonic stabilities).
694 Pairwise correlations (Pearson's r) were calculated, and hierarchical clustering was
695 performed using Euclidean distance and the complete linkage method.

696 An expression filter of exonic RPKM > 0.5 at the earlier time point (0h for the 0h
697 vs 2h comparison, and 2h for the 2h vs 6h comparison) was applied per cell line to
698 ensure that only genes with robust expression were included as starting points. Volcano
699 plots were generated using $\log_2FC$ and $-\log_{10}padj$ significance values from two
700 biological experiments, visualized with VolcanoR (Goedhart and Luijsterburg 2020).
701 Gene set enrichment analysis (GSEA v4.2.3) (Subramanian et al. 2005) was performed
702 using regularized $\logFC$ -ranked lists to identify Hallmark and KEGG pathway gene set
703 enrichments. Gene sets were obtained from the Molecular Signatures Database
704 (MSigDB, <https://www.gsea-msigdb.org/gsea/msigdb>).

705 To compare RNA stability trajectories across cell lines, we selected common
706 genes with exonic RPKM > 0.5 at the earlier comparison time point (0h or 2h) present in

all 16 cell lines, resulting in a set of 4,415 genes. Log₂FC values for these genes were then scaled from 0 to 1 within each cell line for each time point comparison (0h Vs 2h, and 2h Vs 6h), with 1 representing the highest stability and 0 the lowest stability. This scaling enabled direct comparison of RNA dynamics for the 4,415 genes across cell lines. Log₂FC (LFC) values were scaled independently for each cell line using the formula:

$$Scaled\ LFC_i = (LFC_i - \min(LFC)) / (\max(LFC) - \min(LFC))$$

where *Scaled LFC_i* is the normalized value for the *i*th gene; *LFC_i* is the original log₂FC value for the *i*th gene; *min(LFC)* is the minimum log₂FC value in the cell line; and *max(LFC)* is the maximum log₂FC value in the cell line. The process of scaling modifies only the scale, leaving the underlying distribution unchanged. The scaled RNA stability values for the two time periods were then plotted in a scatterplot.

Isoform-specific RNA dynamics

For gene isoform-specific analysis, raw reads from both replicates of each cell line for the three assays (Bru-seq 0h, BruChase-seq 2h and BruChase-seq 6h) were merged and processed through the RSEM pipeline (Li and Dewey 2011). Merged reads were first trimmed using BBduk (BBTools v38.46). Filtered (trimmed) reads were aligned to a premap reference set consisting of human rRNA, human mitochondrial genome, EBV genome and select spike-in references (*D. melanogaster* (dm6) rRNA, dm6 mitochondrial genome). Reads that did not align to premap references were then aligned to the reference genome (GRCh38, GENCODE v29 annotation), specific spike-

in references (dm6 genome, *E. coli* genome (K-12 MG1655) and selected *A. thaliana* gene sequences) using RSEM with the following options:

```
rsem-calculate-expression --num-threads 16 --star --star-gzipped-read-file --output-genome-bam --sort-bam-by-coordinate --sort-bam-memory-per-thread 8G --temporary-folder /temp --keep-intermediate-files --time --paired-end
```

The RSEM output provided normalized isoform counts (transcripts per million, TPM) add isoform percentage quantification files. For downstream analysis, we shortlisted genes that were expressed at ≥ 1 TPM at the Bru-seq (0h) timepoint and were common across all 16 cell lines, resulting in 1,076 genes with at least two isoforms (2,549 transcript isoforms). Isoform percentage changes across the three timepoints (0h, 2h, and 6h) were assessed for these 2,549 transcript isoforms. An isoform with a percentage $\geq 50\%$ at a given time point was defined as the dominant or major isoform. For each cell line, the percentage of genes with isoforms exhibiting specific patterns, such as “always dominant” (AD), “gained dominance” (GD), or “lost dominance” (LD) was determined. A shiny tool (https://umich-biostatistics.shinyapps.io/ENCODE_TPMgte1_MergedReps_IsoPct/) was developed to visualize isoform percentages for these common genes over the time course.

Data Access

The Bru-seq and BruChase-seq datasets generated in this study are available on the ENCODE portal (<https://www.encodeproject.org/>) under the “Deeply Profiled Cell Lines” collection. Accession numbers for the key datasets analyzed in this study can be found in Supplemental Table S13. Custom analysis scripts developed for this manuscript are

available through GitHub at <https://github.com/LjungmanLab/ENCODE-RNA-dynamics-paper> and in Supplemental_Data_File_1.

Competing Interest Statement

The authors have no competing interests to declare.

ACKNOWLEDGEMENTS

We thank the personnel at University of Michigan Advanced Genomics Core for professional technical assistance and Manhong Dai and Fan Meng for administration and maintenance of the University of Michigan Molecular and Behavioral Neuroscience Institute (MBNI) computing cluster. This work was supported by grants from NHGRI (UM1 HG009382) and the Rogel Cancer Center Support Grant (P30 CA046592) by the use of the following Cancer Center Shared Resource(s): Cancer Data Science.

Author contributions: ML developed the Bru-seq and BruChase-seq technologies, performed analyses, made figures and wrote the manuscript. MP performed all the Bru-seq and BruChase-seq experiments. IV and KB performed the major analyses for the manuscript with the help of AM, BM and MA. TEW developed the Bru-seq analysis pipeline used for the bulk of the analyses. All authors discussed and edited the manuscript.

FIGURE LEGENDS

Figure 1. Using BruChase-seq for genome-wide assessment of RNA turnover and splicing. (A) Experimental outline. (B) Relative stability of different RNA categories

compared to protein-coding exons. The values plotted represent averages from 16 human cell lines. Relative stability is shown for 2 hours (2h, blue) and 6 hours (6h, red). (C) Sequencing read distribution of nascent RNA for different genic features at 0h (top), 2h (middle) and 6h (bottom) across 16 human cell lines. Feature types include exon-sense (dark blue), intron-sense (orange), exon-antisense (grey), intron-antisense (yellow), intergenic (light blue), and ambiguous (green). Data from biological replicate #1 was used. (D) Splicing index (SI) across 16 human cell lines at 0, 2 and 6 hours after Bru-labeling. SI values from 24,272 introns are displayed as violin plots with overlaid boxplots showing the median (black solid line) and the interquartile range (white box).

Figure 2. Gene and pathway-specific relative RNA stability. (A) Pairwise correlations (Pearson's r) were used to cluster cell lines according to the relative stability of transcripts from all expressed genes during the 0-2h chase period, and (B) during the 2-6h chase period. Hierarchical clustering was performed using Euclidean distance and complete linkage methods. (C) Volcano plot showing the relative stability of transcripts after the 2h chase period (2vs0h) in Panc1 cells. (D) Volcano plot showing the relative stability of transcripts during the subsequent 4h chase period (6vs2h) in Panc1 cells. (E,F,G) Read coverage profiles for (E) *FAF2*, (F) *DUSP1*, and (G) *MAT2A* in Panc1 cells at 0, 2 and 6 hours. Negative values on the y-axis denote RPKM expression values for the negative strand genes. (H,I,J,K) Gene Set Enrichment Analysis (GSEA) of transcripts ranked by relative RNA stability in Panc1 cells shows the following: upregulated Hallmark pathways at (H) 2 hours and (I) 6 hours, and downregulated Hallmark pathways at (J) 2 hours and (K) 6 hours. (L) GSEA analyses in Panc1 cells of

four KEGG pathways involving multiprotein complexes show decreased RNA stability at 2 hours but increased RNA stability at 6 hours.

Figure 3. Different RNA stability dynamics for selected genes. (A) Gene traces displaying read counts across selected genes in Panc1 cells at three time points. Negative RPKM values on the y-axis denote gene expression on the negative strand. (B) Scaled stability values for all transcripts of genes depicted in (A). (C,D) Distribution of scaled stability values for all transcripts in Panc1 cells during the (C) 0-2h time period and (D) 2-6h time period. (E) Distribution of scaled stability values for *RPL22L1* (left), and correlation between splicing index and stability of *RPL22L1* during the 0-2h interval (right). (F) Distribution of scaled stability values for the *MRPL17* transcript (left), and correlation between splicing index and stability of *MRPL17* during the 0-2h interval (right). (G) RPKM values of the *MRPL17* transcript across the 16 cell lines at 0h and 6h, with values for K562 cells highlighted in light green.

Figure 4. Isoform-specific post-transcriptional processing. (A) Coverage of captured Bru-labeled RNA across *ETV3* isoforms in A673 cells. (B,C,D,E) Isoform percentages, as computed by the RSEM pipeline, for transcript isoforms generated by (B) *ETV3*, (C) *ATF4*, (D) *NDUFS5*, and (E) *HNRNPR* genes at the three time points. Each of the 16 cell lines is denoted by a different colored line. (F) Percentage of genes (1,076 genes with at least two isoforms) under study whose isoforms showed varying patterns of dominance over time: always dominant (AD), gained dominance (GD), and lost dominance (LD).

821

822 **Figure 5. Splicing patterns differ within and between transcripts.** (A) Gene traces
823 of the *DUSP1* gene in A673 cells at 0, 2 and 6 hours, showing read coverage (RPKM, y-
824 axis), with individual introns numbered in black and the corresponding splicing index
825 (SI) values indicated in red. While introns 1-4 show SI values >0.93 at 2h, intron 5
826 displays an SI of 0.51 but approaches complete splicing by 6h. (B) SI values for the
827 introns along the *DUSP12* transcript (left plot) and corresponding relative intron
828 abundances (right plot) are shown for the different time points. (C) Gene traces of
829 introns 2-5 of the *STAM* gene in A673 cells, showing read coverage (RPKM, y-axis) and
830 retention of intron 4 throughout the time course, despite efficient splicing at 2h. (D) SI
831 values along the *STAM* transcript (left plot) and relative intron abundance (right plot) in
832 A673 cells.

833

834 **Figure 6. Coordinated splicing of introns flanking 47,163 exons.** (A) Diagram
835 illustrating the formation of four distinct splicing pattern categories. (B) Relative
836 proportions for the four co-transcriptional splicing patterns across 47,163 introns in 16
837 cell lines. (C) Fraction of splicing categories for transcripts synthesized from the *VMP1*,
838 *AATF* and *HDAC1* genes in HCT116 cells. Exonic RPKM values at 0 hours for each
839 gene are indicated. (D) Fraction of splicing categories averaged across the 16 cell lines
840 for each of the three time points.

REFERENCES

- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* **290**: 457-465.
- Andersson R, Sandelin A. 2020. Determinants of enhancer and promoter activities of regulatory elements. *Nat Rev Genet* **21**: 71-87.
- Bedi K, Magnuson BR, Narayanan I, Paulsen M, Wilson TE, Ljungman M. 2021. Co-transcriptional splicing efficiencies differ within genes and between cell types. *RNA* doi:10.1261/rna.078662.120.
- Bhat P, Chow A, Emert B, Ettlin O, Quinodoz SA, Strehle M, Takei Y, Burr A, Goronzy IN, Chen AW et al. 2024. Genome organization around nuclear speckles drives mRNA splicing efficiency. *Nature* **629**: 1165-1173.
- Bousquet-Antonelli C, Presutti C, Tollervey D. 2000. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102**: 765-775.
- Boutz PL, Bhutkar A, Sharp PA. 2015. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev* **29**: 63-80.
- Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ. 2014. Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res* **24**: 1774-1786.
- Bushnell, Brian. "BBMap: A Fast, Accurate, Splice-Aware Aligner." 2014.

862 Chen R, Ishak CA, De Carvalho DD. 2021. Endogenous Retroelements and the Viral
863 Mimicry Response in Cancer Therapy and Cellular Homeostasis. *Cancer Discov*
864 **11**: 2707-2725.

865 Choquet K, Baxter-Koenigs AR, Dulk SL, Smalec BM, Rouskin S, Churchman LS. 2023.
866 Pre-mRNA splicing order is predetermined and maintains splicing fidelity across
867 multi-intronic transcripts. *Nat Struct Mol Biol* **30**: 1064-1076.

868 Choquet K, Chaumont LP, Bache S, Baxter-Koenigs AR, Churchman LS. 2025a.
869 Genetic regulation of nascent RNA maturation revealed by direct RNA nanopore
870 sequencing. *Genome Res* **35**: 712-724.

871 Choquet K, Patop IL, Churchman LS. 2025b. The regulation and function of post-
872 transcriptional RNA splicing. *Nat Rev Genet* **26**: 378-394.

873 Darnell JE, Jr. 2013. Reflections on the history of pre-mRNA processing and highlights
874 of current knowledge: a unified picture. *RNA* **19**: 443-460.

875 Ding F, Elowitz MB. 2019. Constitutive splicing and economies of scale in gene
876 expression. *Nat Struct Mol Biol* **26**: 424-432.

877 Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde
878 J, Lin W, Schlesinger F et al. 2012. Landscape of transcription in human cells.
879 *Nature* **489**: 101-108.

880 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
881 Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*
882 **29**: 15-21.

883 Doma MK, Parker R. 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in
884 translation elongation. *Nature* **440**: 561-564.

885 Drexler HL, Choquet K, Churchman LS. 2020. Splicing Kinetics and Coordination
886 Revealed by Direct Nascent RNA Sequencing through Nanopores. *Mol Cell* **77**:
887 985-998 e988.

888 The ENCODE Project Consortium, Moore JE, Purcaro MJ, Pratt HE, Epstein CB,
889 Shores N, Adrian J, Kawli T, Davis CA, Dobin A et al. 2020. Expanded
890 encyclopaedias of DNA elements in the human and mouse genomes. *Nature*
891 **583**: 699-710.

892 Friedel CC, Dolken L, Ruzsics Z, Koszinowski UH, Zimmer R. 2009. Conserved
893 principles of mammalian transcriptional regulation revealed by RNA half-life.
894 *Nucleic Acids Res* **37**: e115.

895 Frischmeyer PA, van Hoof A, O'Donnell K, Guerrerio AL, Parker R, Dietz HC. 2002. An
896 mRNA surveillance mechanism that eliminates transcripts lacking termination
897 codons. *Science* **295**: 2258-2261.

898 Furlan M, Galeota E, Gaudio ND, Dassi E, Caselle M, de Pretis S, Pelizzola M. 2020.
899 Genome-wide dynamics of RNA synthesis, processing, and degradation without
900 RNA metabolic labeling. *Genome Res* **30**: 1492-1507.

901 Goedhart J, Luijsterburg MS. 2020. VolcanoR is a web app for creating, exploring,
902 labeling and sharing volcano plots. *Sci Rep* **10**: 20560.

903 Grubert F, Srivas R, Spacek DV, Kasowski M, Ruiz-Velasco M, Sinnott-Armstrong N,
904 Greenside P, Narasimha A, Liu Q, Geller B et al. 2020. Landscape of cohesin-
905 mediated chromatin loops in the human genome. *Nature* **583**: 737-743.

906 He S, Valkov E, Cheloufi S, Murn J. 2023. The nexus between RNA-binding proteins
907 and their effectors. *Nat Rev Genet* **24**: 276-294.

908 Herzel L, Straube K, Neugebauer KM. 2018. Long-read sequencing of nascent RNA
 909 reveals coupling among RNA processing events. *Genome Res* **28**: 1008-1019.
 910 Herzog VA, Reichholf B, Neumann T, Rescheneder P, Bhat P, Burkard TR, Wlotzka W,
 911 von Haeseler A, Zuber J, Ameres SL. 2017. Thiol-linked alkylation of RNA to
 912 assess expression dynamics. *Nat Methods* **14**: 1198-1204.
 913 Houseley J, LaCava J, Tollervey D. 2006. RNA-quality control by the exosome. *Nat Rev*
 914 *Mol Cell Biol* **7**: 529-539.
 915 Huang E, Fu T, Zhang L, Yan G, Yamamoto R, Terrazas S, Nguyen TL, Gonzalez-
 916 Figueroa C, Khanbabaei A, Bahn JH et al. 2025. Genetic variants affecting RNA
 917 stability influence complex traits and disease risk. *Nat Genet*
 918 doi:10.1038/s41588-025-02326-8.
 919 Imamachi N, Tani H, Mizutani R, Imamura K, Irie T, Suzuki Y, Akimitsu N. 2014. BRIC-
 920 seq: a genome-wide approach for determining RNA stability in mammalian cells.
 921 *Methods* **67**: 55-63.
 922 Jens M, Rajewsky N. 2015. Competition between target sites of regulators shapes post-
 923 transcriptional gene regulation. *Nat Rev Genet* **16**: 113-126.
 924 Kawata K, Wakida H, Yamada T, Taniue K, Han H, Seki M, Suzuki Y, Akimitsu N. 2020.
 925 Metabolic labeling of RNA using multiple ribonucleoside analogs enables the
 926 simultaneous evaluation of RNA synthesis and degradation rates. *Genome Res*
 927 **30**: 1481-1491.
 928 Kilchert C, Wittmann S, Vasiljeva L. 2016. The regulation and functions of the nuclear
 929 RNA exosome complex. *Nat Rev Mol Cell Biol* **17**: 227-239.

930 Kwon DH, Chou J, Yip SM, Reimers MA, Zhang L, Wright F, Dhawan MS, Borno HT,
 931 Desai A, Aggarwal RR et al. 2021. Differential treatment outcomes in BRCA1/2-,
 932 CDK12-, and ATM-mutated metastatic castration-resistant prostate cancer.
 933 *Cancer* doi:10.1002/cncr.33487.

934 Lam LT, Pickeral OK, Peng AC, Rosenwald A, Hurt EM, Giltane JM, Averett LM, Zhao
 935 H, Davis RE, Sathymoorthy M et al. 2001. Genomic-scale measurement of
 936 mRNA turnover and the mechanisms of action of the anti-cancer drug
 937 flavopiridol. *Genome Biol* **2**: RESEARCH0041.

938 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat*
 939 *Methods* **9**: 357-359.

940 Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data
 941 with or without a reference genome. *BMC Bioinformatics* **12**: 323.

942 Ljungman M. 2022. Transcription and genome integrity. *DNA Repair (Amst)* **118**:
 943 103373.

944 Losson R, Lacroute F. 1979. Interference of nonsense mutations with eukaryotic
 945 messenger RNA stability. *Proc Natl Acad Sci U S A* **76**: 5134-5137.

946 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and
 947 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

948 Lugowski A, Nicholson B, Rissland OS. 2018. DRUID: a pipeline for transcriptome-wide
 949 measurements of mRNA stability. *RNA* **24**: 623-632.

950 Magnuson B, Bedi K, Narayanan IV, Bartkowiak B, Blinkiewicz H, Paulsen MT,
 951 Greenleaf A, Ljungman M. 2022. CDK12 regulates co-transcriptional splicing and
 952 RNA turnover in human cells. *iScience* **25**: 105030.

953 Magnuson B, Veloso A, Kirkconnell KS, Lima LC, Paulsen MT, Ljungman EA, Bedi K,
 954 Prasad J, Wilson TE, Ljungman M. 2015. Identifying transcription start sites and
 955 active enhancer elements using BruUV-seq. *Sci Rep* **5**: 17978.
 956 Maier KC, Gressel S, Cramer P, Schwalb B. 2020. Native molecule sequencing by
 957 nano-ID reveals synthesis and stability of RNA isoforms. *Genome Res* **30**: 1332-
 958 1344.
 959 Maquat LE, Kinniburgh AJ, Rachmilewitz EA, Ross J. 1981. Unstable beta-globin
 960 mRNA in mRNA-deficient beta o thalassemia. *Cell* **27**: 543-553.
 961 Marasco LE, Kornblihtt AR. 2023. The physiology of alternative splicing. *Nat Rev Mol*
 962 *Cell Biol* **24**: 242-254.
 963 McShane A, Narayanan IV, Paulsen MT, Ashaka M, Blinkiewicz H, Yang NT, Magnuson
 964 B, Bedi K, Wilson TE, Ljungman M. 2024. Characterizing nascent transcription
 965 patterns of PROMPTs, eRNAs, and readthrough transcripts in the ENCODE4
 966 deeply profiled cell lines. *bioRxiv* doi:10.1101/2024.04.09.588612.
 967 Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood AM, Haugen E,
 968 Bracken CP, Rackham O, Stamatoyannopoulos JA et al. 2011. The human
 969 mitochondrial transcriptome. *Cell* **146**: 645-658.
 970 Merens HE, Choquet K, Baxter-Koenigs AR, Churchman LS. 2024. Timing is
 971 everything: advances in quantifying splicing kinetics. *Trends Cell Biol* **34**: 968-
 972 981.
 973 Neph S, Stergachis AB, Reynolds A, Sandstrom R, Borenstein E, Stamatoyannopoulos
 974 JA. 2012. Circuitry and dynamics of human transcription factor regulatory
 975 networks. *Cell* **150**: 1274-1286.

976 Oesterreich FC, Herzel L, Straube K, Hujer K, Howard J, Neugebauer KM. 2016.
 977 Splicing of Nascent RNA Coincides with Intron Exit from RNA Polymerase II. *Cell*
 978 **165**: 372-381.

979 Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in
 980 human mitochondria. *Nature* **290**: 470-474.

981 Palumbo MC, Farina L, Paci P. 2015. Kinetics effects and modeling of mRNA turnover.
 982 *Wiley Interdiscip Rev RNA* **6**: 327-336.

983 Pan K, Lee JT, Huang Z, Wong CM. 2015. Coupling and coordination in gene
 984 expression processes with pre-mRNA splicing. *Int J Mol Sci* **16**: 5682-5696.

985 Paulsen MT, Veloso A, Prasad J, Bedi K, Ljungman EA, Magnuson B, Wilson TE,
 986 Ljungman M. 2014. Use of Bru-seq and BruChase-seq for genome-wide
 987 assessment of the synthesis and stability of RNA. *Methods* **67**: 45-54.

988 Paulsen MT, Veloso A, Prasad J, Bedi K, Ljungman EA, Tsan YC, Chang CW, Tarrier
 989 B, Washburn JG, Lyons R et al. 2013. Coordinated regulation of synthesis and
 990 stability of RNA during the acute TNF-induced proinflammatory response. *Proc*
 991 *Natl Acad Sci U S A* **110**: 2240-2245.

992 Piechota J, Tomecki R, Gewartowski K, Szczesny R, Dmochowska A, Kudla M,
 993 Dybczynska L, Stepień PP, Bartnik E. 2006. Differential stability of mitochondrial
 994 mRNA in HeLa cells. *Acta Biochim Pol* **53**: 157-168.

995 Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, Garber M, Gnirke A,
 996 Nusbaum C, Hacohen N, Friedman N et al. 2011. Metabolic labeling of RNA
 997 uncovers principles of RNA production and degradation dynamics in mammalian
 998 cells. *Nat Biotechnol* **29**: 436-442.

999 Raghavan A, Bohjanen PR. 2004. Microarray-based analyses of mRNA decay in the
1000 regulation of mammalian gene expression. *Brief Funct Genomic Proteomic* **3**:
1001 112-124.

1002 Rajurkar M, Parikh AR, Solovyov A, You E, Kulkarni AS, Chu C, Xu KH, Jaicks C,
1003 Taylor MS, Wu C et al. 2022. Reverse Transcriptase Inhibition Disrupts Repeat
1004 Element Life Cycle in Colorectal Cancer. *Cancer Discov* **12**: 1462-1481.

1005 Reimer KA, Mimoso CA, Adelman K, Neugebauer KM. 2021. Co-transcriptional splicing
1006 regulates 3' end cleavage during mammalian erythropoiesis. *Mol Cell*
1007 doi:10.1016/j.molcel.2020.12.018.

1008 Riml C, Amort T, Rieder D, Gasser C, Lusser A, Micura R. 2017. Osmium-Mediated
1009 Transformation of 4-Thiouridine to Cytidine as Key To Study RNA Dynamics by
1010 Sequencing. *Angew Chem Int Ed Engl* **56**: 13479-13483.

1011 Rodriguez-Molina JB, West S, Passmore LA. 2023. Knowing when to stop:
1012 Transcription termination on protein-coding genes by eukaryotic RNAPII. *Mol Cell*
1013 **83**: 404-415.

1014 Rogalska ME, Vivori C, Valcarcel J. 2023. Regulation of pre-mRNA splicing: roles in
1015 physiology and disease, and therapeutic prospects. *Nat Rev Genet* **24**: 251-269.

1016 Schoenberg DR, Maquat LE. 2012. Regulation of cytoplasmic mRNA decay. *Nat Rev*
1017 *Genet* **13**: 246-259.

1018 Schofield JA, Duffy EE, Kiefer L, Sullivan MC, Simon MD. 2018. TimeLapse-seq: adding
1019 a temporal dimension to RNA sequencing through nucleoside recoding. *Nat*
1020 *Methods* **15**: 221-225.

1021 Schwanhaussner B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach
 1022 M. 2011. Global quantification of mammalian gene expression control. *Nature*
 1023 **473**: 337-342.

1024 Sharova LV, Sharov AA, Nedorezov T, Piao Y, Shaik N, Ko MS. 2008. Database for
 1025 mRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of
 1026 Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Res*
 1027 doi:10.1093/dnares/dsn030.

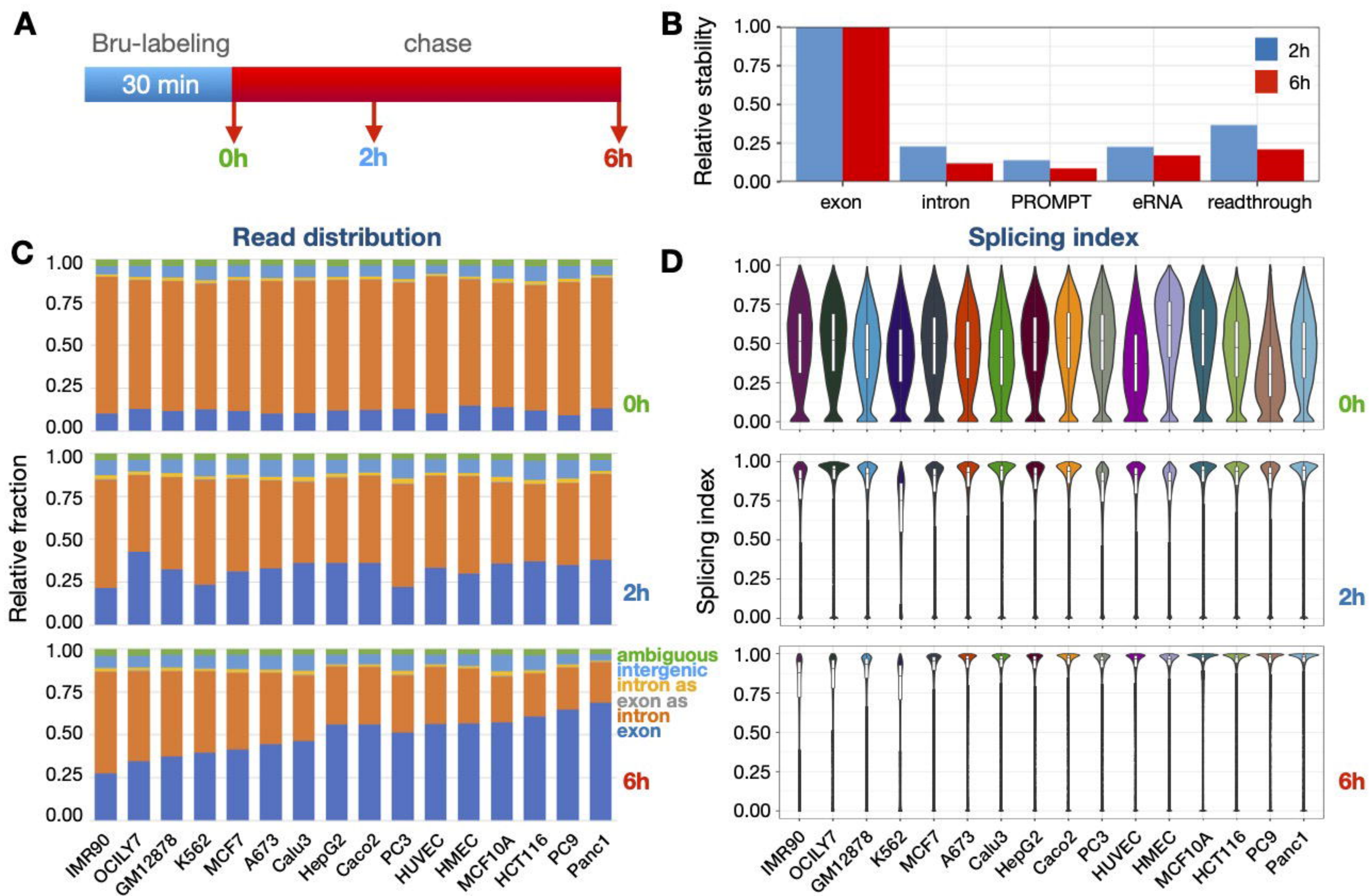
1028 Stricker SH, Kofler A, Beck S. 2017. From profiles to function in epigenomics. *Nat Rev*
 1029 *Genet* **18**: 51-66.

1030 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA,
 1031 Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set
 1032 enrichment analysis: a knowledge-based approach for interpreting genome-wide
 1033 expression profiles. *Proc Natl Acad Sci U S A*. 2005 Oct 25;102(43):15545-50.
 1034 doi: 10.1073/pnas.0506580102.

1035 Szczesny RJ, Borowski LS, Brzezniak LK, Dmochowska A, Gewartowski K, Bartnik E,
 1036 Stepień PP. 2010. Human mitochondrial RNA turnover caught in flagranti:
 1037 involvement of hSuv3p helicase in RNA surveillance. *Nucleic Acids Res* **38**: 279-
 1038 298.

1039 Tani H, Akimitsu N. 2012. Genome-wide technology for determining RNA stability in
 1040 mammalian cells: historical perspective and recent advantages based on
 1041 modified nucleotide labeling. *RNA Biol* **9**: 1233-1238.

1042 Tani H, Mizutani R, Salam KA, Tano K, Ijiri K, Wakamatsu A, Isogai T, Suzuki Y,
 1043 Akimitsu N. 2012. Genome-wide determination of RNA stability reveals hundreds
 1044 of short-lived noncoding transcripts in mammals. *Genome Res* **22**: 947-956.
 1045 Tian Y, Zeng Z, Li X, Wang Y, Chen R, Mattijssen S, Gaidamakov S, Wu Y, Maraia RJ,
 1046 Peng W et al. 2020. Transcriptome-wide stability analysis uncovers LARP4-
 1047 mediated NFkappaB1 mRNA stabilization during T cell activation. *Nucleic Acids*
 1048 *Res* **48**: 8724-8739.
 1049 van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition
 1050 and degradation of mRNAs lacking a termination codon. *Science* **295**: 2262-
 1051 2264.
 1052 Vierstra J, Lazar J, Sandstrom R, Halow J, Lee K, Bates D, Diegel M, Dunn D, Neri F,
 1053 Haugen E et al. 2020. Global reference mapping of human transcription factor
 1054 footprints. *Nature* **583**: 729-736.
 1055 Villar D, Flicek P, Odom DT. 2014. Evolution of transcription factor binding in
 1056 metazoans - mechanisms and functional implications. *Nat Rev Genet* **15**: 221-
 1057 233.
 1058 Wang Y, Liu CL, Storey JD, Tibshirani RJ, Herschlag D, Brown PO. 2002. Precision and
 1059 functional specificity in mRNA decay. *Proc Natl Acad Sci U S A* **99**: 5860-5865.
 1060 Wang Z, Burge CB. 2008. Splicing regulation: from a parts list of regulatory elements to
 1061 an integrated splicing code. *RNA* **14**: 802-813.
 1062 Yang E, van Nimwegen E, Zavolan M, Rajewsky N, Schroeder M, Magnasco M, Darnell
 1063 JE, Jr. 2003. Decay rates of human mRNAs: correlation with functional
 1064 characteristics and sequence attributes. *Genome Res* **13**: 1863-1872.



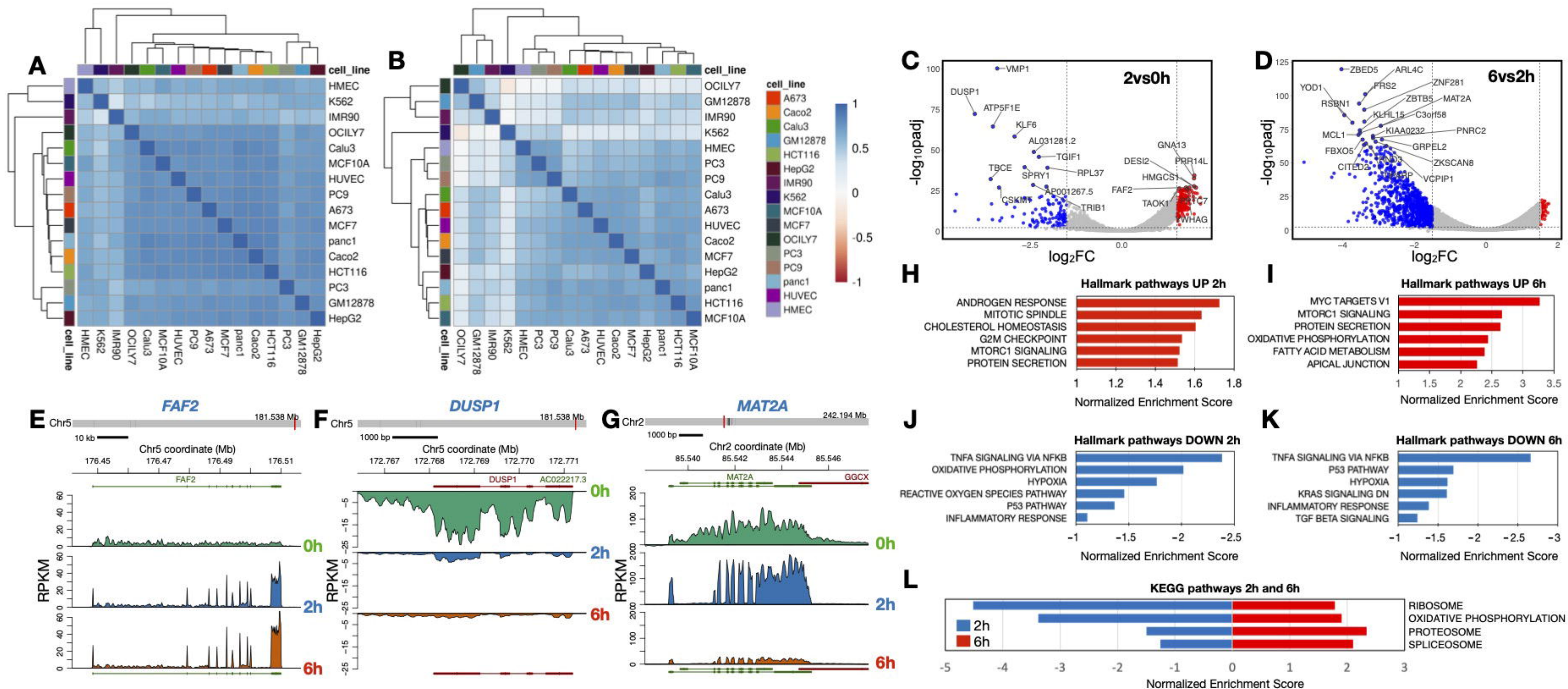


Figure 2

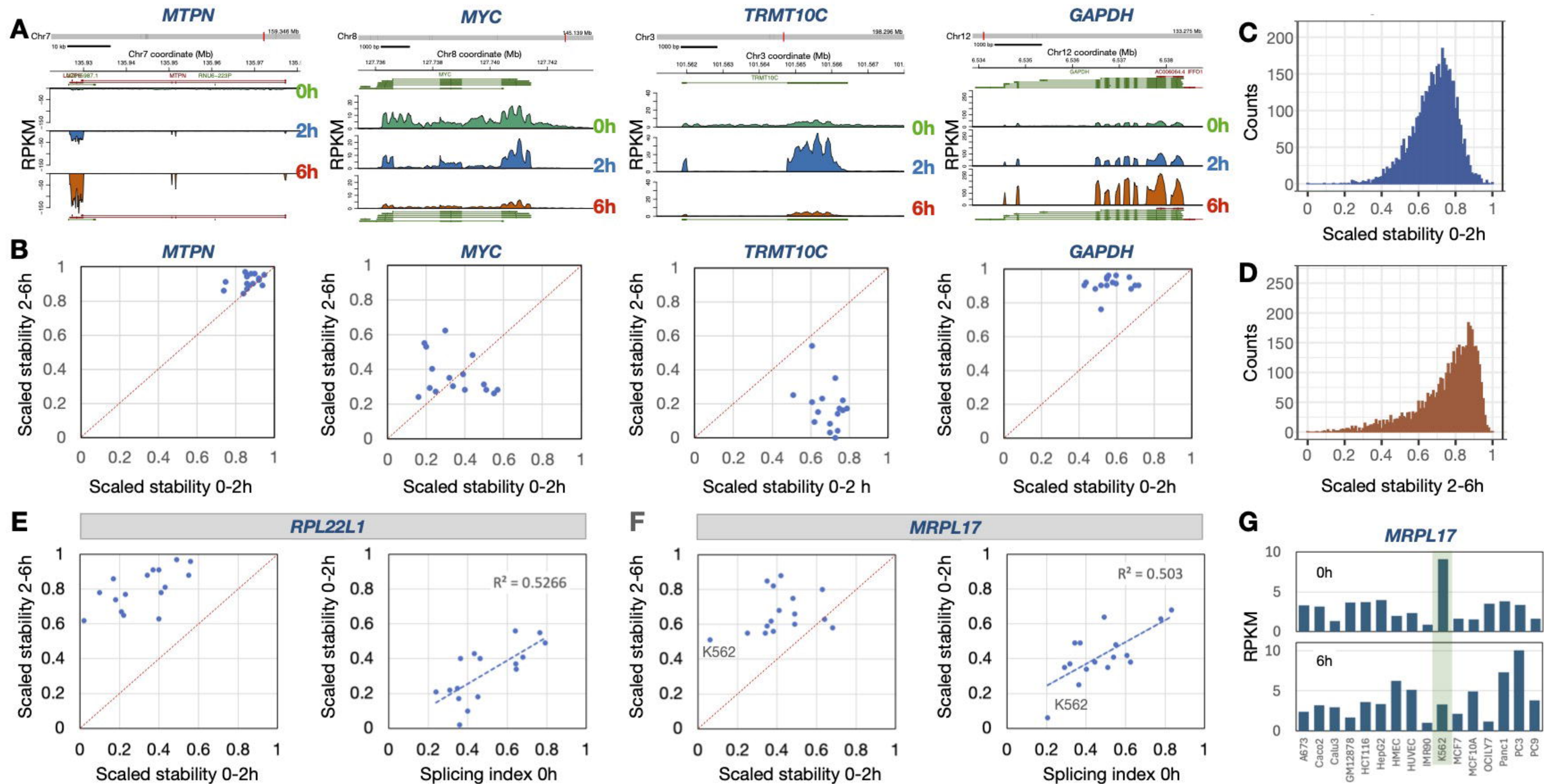


Figure 3

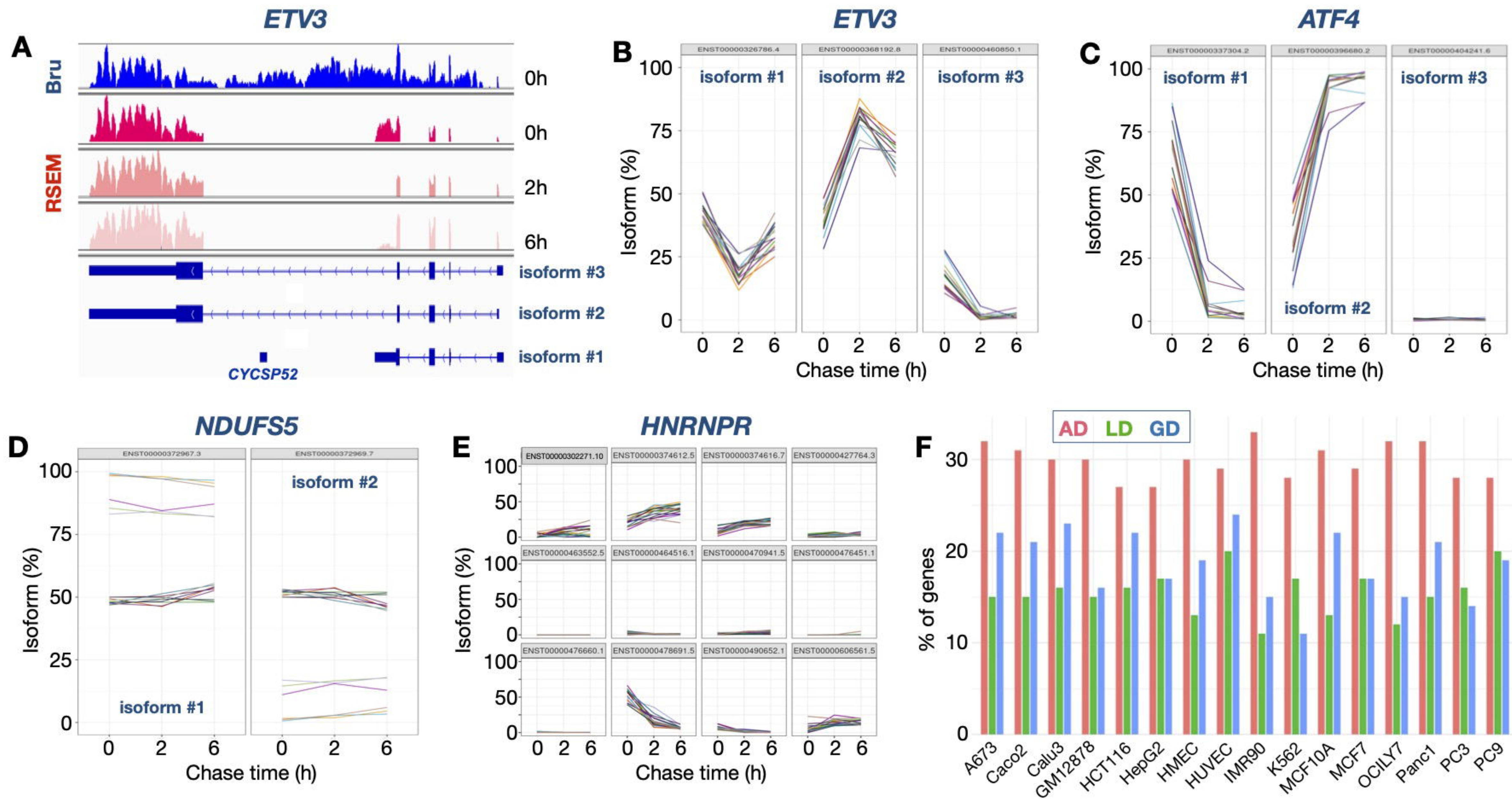


Figure 4

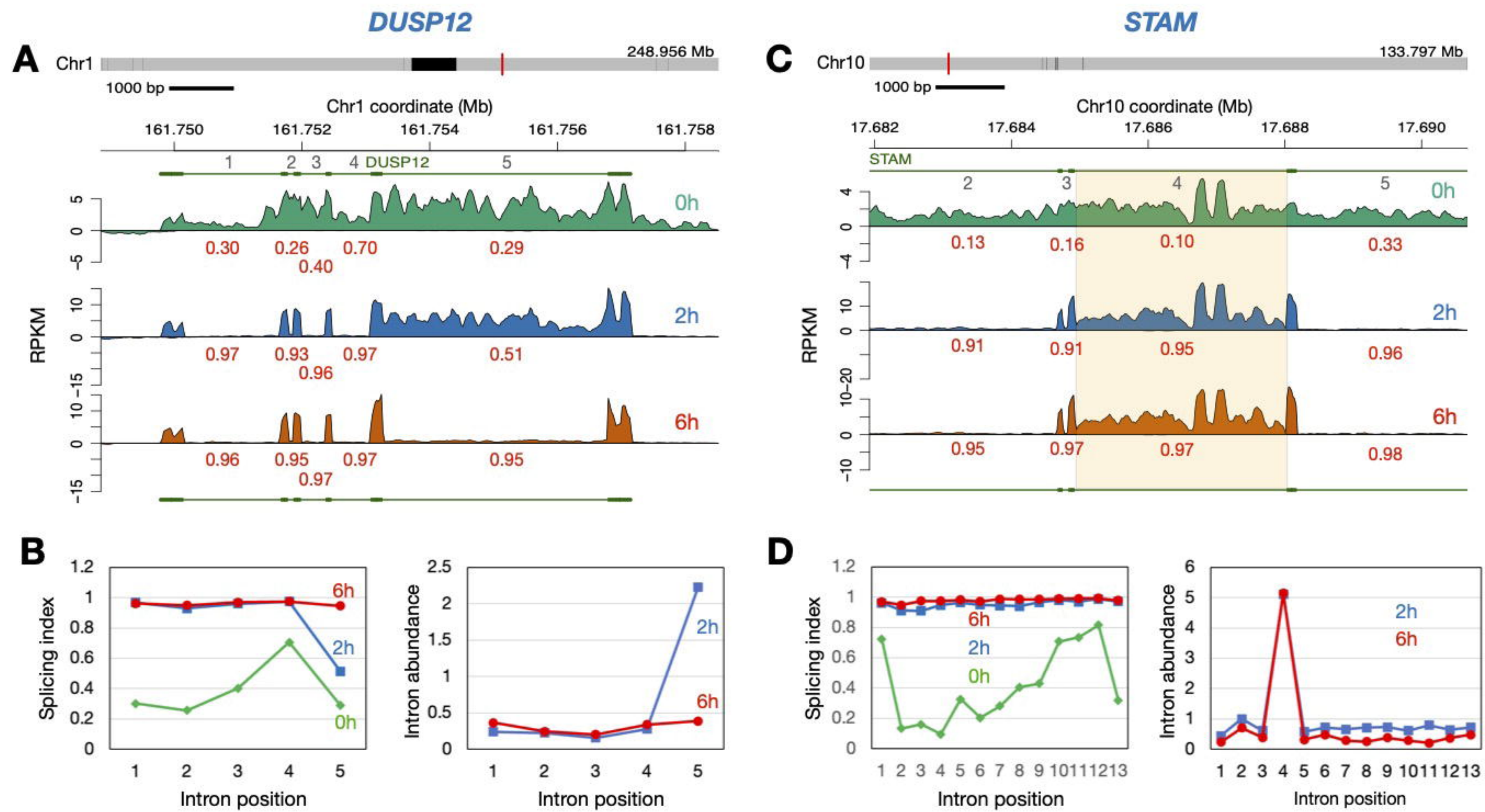


Figure 5

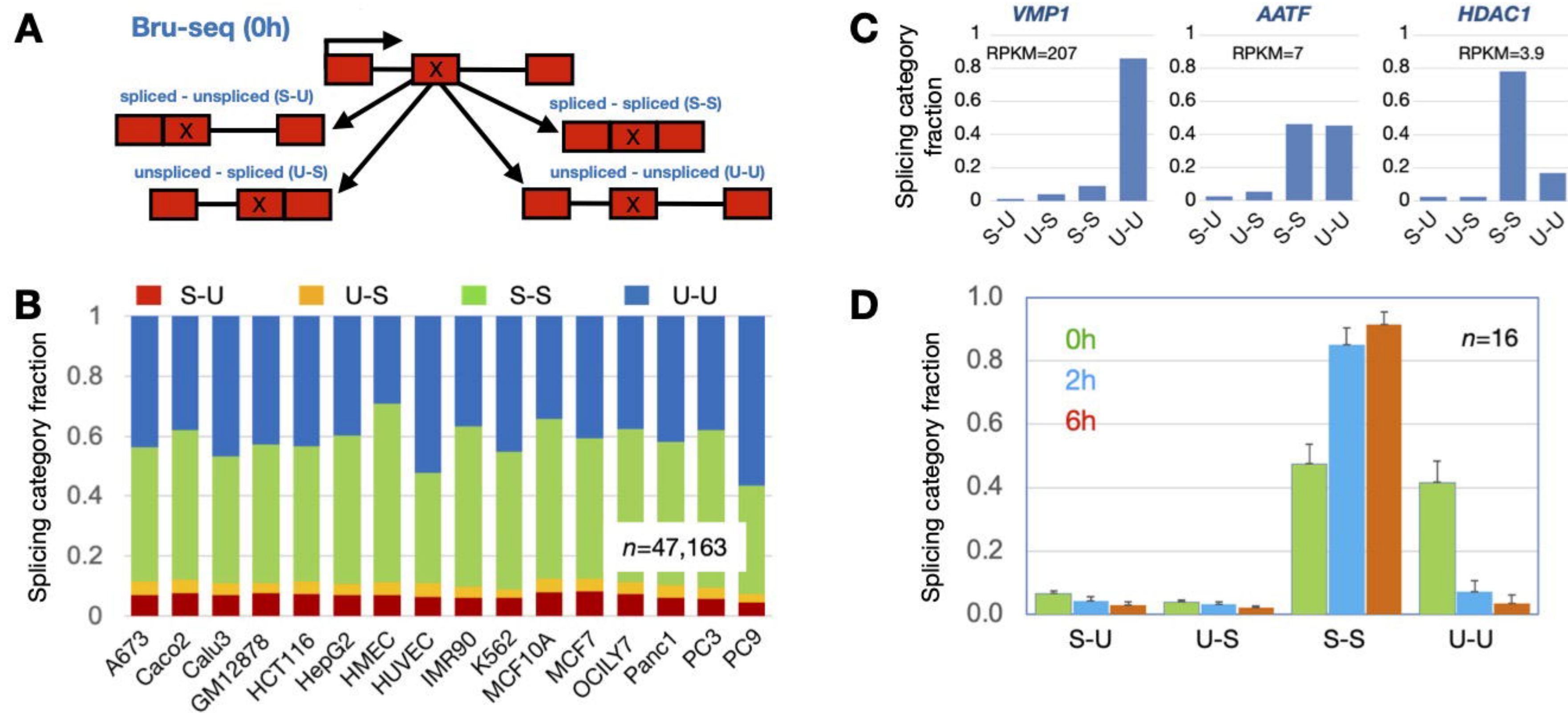


Figure 6