**ADVANCED REVIEW**

WIREs RNA | WILEY

# Uncovering the impacts of alternative splicing on the proteome with current omics techniques

**Marina Reixachs-Solé**[1,2] | **Eduardo Eyras**[1,2,3,4]

[1]The John Curtin School of Medical Research, Australian National University, Canberra, Australian Capital Territory, Australia

[2]EMBL Australia Partner Laboratory Network and the Australian National University, Canberra, Australian Capital Territory, Australia

[3]Catalan Institution for Research and Advanced Studies, Barcelona, Spain

[4]Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain

**Correspondence**
Eduardo Eyras, The John Curtin School of Medical Research, Australian National University, Canberra, ACT, Australia.
Email: eduardo.eyras@anu.edu.au

**Edited by:** Jeff Wilusz, Editor-in-Chief

**Abstract**

The high-throughput sequencing of cellular RNAs has underscored a broad effect of isoform diversification through alternative splicing on the transcriptome. Moreover, the differential production of transcript isoforms from gene loci has been recognized as a critical mechanism in cell differentiation, organismal development, and disease. Yet, the extent of the impact of alternative splicing on protein production and cellular function remains a matter of debate. Multiple experimental and computational approaches have been developed in recent years to address this question. These studies have unveiled how molecular changes at different steps in the RNA processing pathway can lead to differences in protein production and have functional effects. New and emerging experimental technologies open exciting new opportunities to develop new methods to fully establish the connection between messenger RNA expression and protein production and to further investigate how RNA variation impacts the proteome and cell function.

This article is categorized under:

RNA Processing > Splicing Regulation/Alternative Splicing
Translation > Regulation
RNA Evolution and Genomics > Computational Analyses of RNA

**KEYWORDS**

alternative splicing, proteome, RNA processing, transcriptomics, translatomics

## 1 | INTRODUCTION

With the human genome being almost complete, the estimated number of protein-coding genes has stabilized at around 20,000 (Lander et al., 2001; Pertea et al., 2018). However, there is nearly six times the number of annotated proteins (115,885 in Ensembl GRCh38.p13, September 2021). The key to this diversity lies in the fact that every step in the production and processing of messenger RNA (mRNA), from transcription to translation, provides an opportunity for molecular diversification. Among all mRNA processing steps, perhaps the most well-studied mechanism and possibly the one that provides the greatest chance for molecular variation in eukaryotes is splicing. First described in 1977, splicing is a sequential set of biochemical reactions to remove introns from the pre-mRNA molecule (Berget et al., 1977;

Chow et al., 1977). The discovery of splicing came with the observation that the same adenoviral gene produced different transcripts (Berget et al., 1977; Chow et al., 1977). The modular structure of the pre-mRNA, with exons and introns, facilitates the assembly of different configurations of exons to generate alternative transcript isoforms (Gilbert, 1978), that is, alternative splicing. Soon after the discovery of splicing, the first case of alternative splicing was described in humans in the immunoglobulin μ gene, where one of the isoforms encoded the membrane-bound antibody. In contrast, the other isoform encoded the secreted form of the protein (Singer et al., 1980). In recent decades, high-throughput sequencing techniques have provided evidence indicating that more than 95% of the genes may undergo alternative splicing (Barbosa-Morais et al., 2012; Pan et al., 2008; Wang et al., 2008). These results have highlighted the vast impact that alternative splicing has on shaping the transcriptome. However, it remains unclear how frequently transcriptome variation results in measurable protein and functional changes.

A shift in the relative abundances of the transcript isoforms produced from a gene locus may lead to the production of protein variants, which may in turn impact gene function through various mechanisms, including the alteration of protein–protein interactions (Ellis et al., 2012; Weatheritt et al., 2012; Wojtowicz et al., 2007). However, in vivo studies using cutting-edge proteomics have failed so far to validate these functional impacts, sparking an intense debate on the field (Blencowe, 2017; Tress et al., 2017). In addition to the intrinsic differences and limitations of RNA sequencing and proteomics technologies, multiple regulatory mechanisms controlling transcript stability and translation could hinder the identification of direct links between transcript and protein variation (Braun & Young, 2014). Despite the abundant evidence supporting the functional impact of alternative splicing (Baralle & Giudice, 2017), whether it leads to fundamentally distinct proteins remains unknown for most genes. Understanding the regulation and interdependency of splicing with other regulatory processes is crucial to explain how the transcriptome encodes the functional complexity of cells. Moreover, to fill in the knowledge gaps, from gene expression to the production of different functional proteins, it is essential to gain insight into the mechanisms of RNA processing in the context of both the transcriptome and the translatome. In this review, we describe the mechanisms by which changes in RNA processing may result in the translation of different protein products, the high-throughput techniques that have been developed to uncover these mechanisms, and the results from recent studies regarding the impact of alternative splicing in protein production and gene function.

## 2 | MECHANISMS OF TRANSCRIPT DIVERSITY

The splicing reaction starts with the recognition of sequence elements in the precursor messenger RNA (pre-mRNA) by a small nuclear ribonucleoprotein (snRNP) complex, the spliceosome, which acquires its catalytic activity while assembling on the pre-mRNA (Zhan et al., 2018; Figure 1a). In this process, intron sequences are excised from the pre-mRNA, and exon sequences are spliced together to form the mRNA. Several sequence elements in the pre-mRNA are essential for splicing. Most introns are characterized by the nucleotides GU at the 5′ splice site (SS) and AG at the 3′ SS, with very few exceptions (Figure 1a; Thanaraj & Clark, 2001). The branch point (BP) consists of an A nucleotide with a catalytic role near the 3'SS, and a polypyrimidine tract (PPT) is located between the BP and the 3′ SS. However, BPs may also occur far upstream of the 3'SS (Corvelo et al., 2010). The BP and PPT show specific sequence motif biases, but these may vary between species (Schwartz et al., 2008). The core spliceosome comprises five main snRNPs (U1, U2, U4, U5, and U6) that are sequentially assembled in the mRNA molecule, where U1 recognizes the 5′ SS and U2 the BP. Several conformational and compositional rearrangements involving U2, U5, and U6 take place to allow two trans-esterifications: the first one between the 5′ SS and the BP and the second between the 5′ end of the first exon and the 3′ end of the second exon (Figure 1a).

## 2.1 | Alternative splicing regulation

Exon recognition is not straightforward, especially in metazoans, considering that the sequence surrounding the splice sites and the branch point can be highly variable, and introns are on average thousands of times larger than exons (Iwata & Gotoh, 2011; Lander et al., 2001; Long & Deutsch, 1999). For this reason, for efficient splicing to occur, additional cis-acting elements on the pre-mRNA sequence and trans-acting proteins are often required (Figure 1b; Lim & Burge, 2001). These additional regulatory elements, which work together with the core splicing signals (3′SS, 5′SS, and BP), are not just essential for catalysis but often act as modulators of the splicing process. They are classified according to their location on intron or exons and their role in either enhancing or repressing splicing, resulting in four main
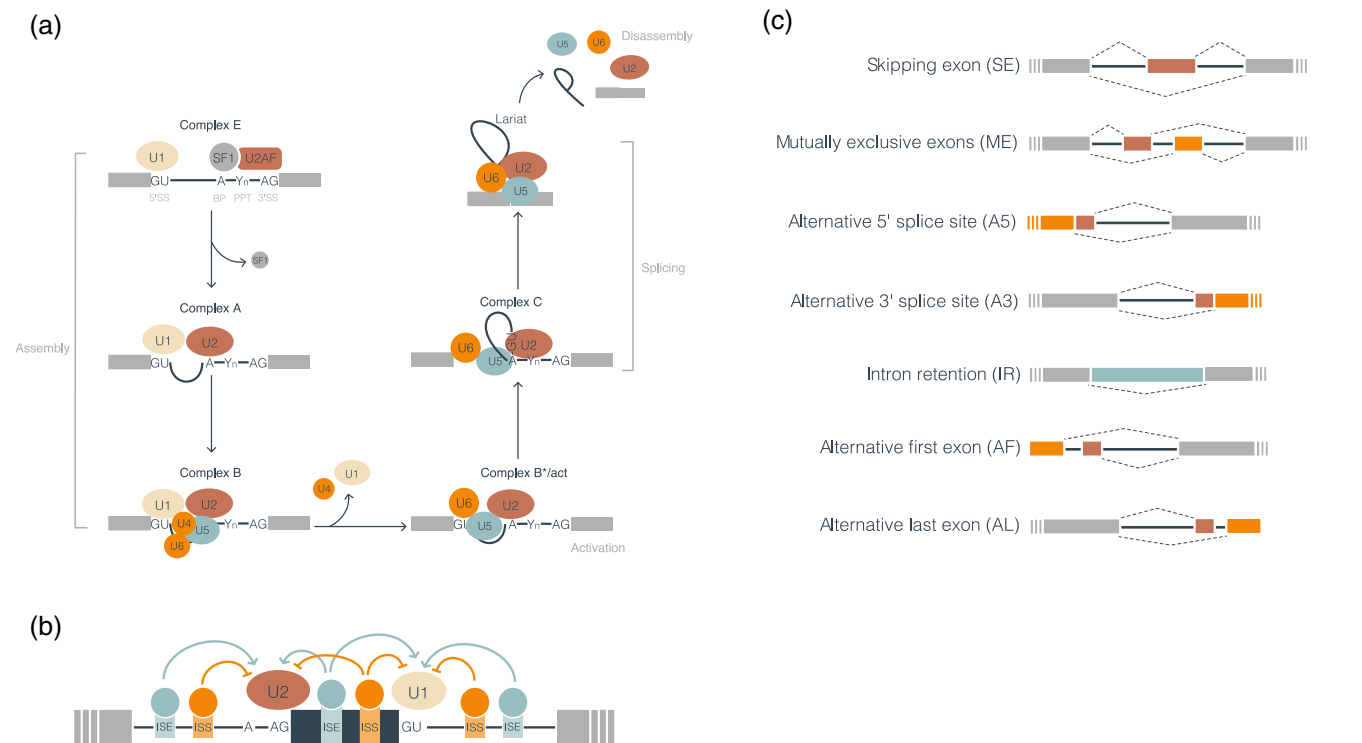
**FIGURE 1** The mechanisms of splicing and alternative splicing. (a) The complexes and processing steps of splicing for the major spliceosome are depicted. In the assembly phase, U1, U2, U4, U5, and U6 are sequentially assembled in the mRNA molecule resulting in the formation of complex E, complex A, and finally, complex B. The activation phase consists of conformational and compositional rearrangements of complex B involving U2, U5, and U6. Two transesterifications occur as the catalytic complex C adapts its conformation: the first between the 5′ SS and the BP and the second between the 5′ end of the first exon and the 3′ end of the second exon. Finally, the snRNPs still attached to the mRNA disassemble to start a new cycle. (b) Illustration of the regulation of splicing by *cis*- and *trans*-acting factors. Cis-acting factors are represented as boxes in introns and exons. *Trans*-acting factors are represented as circles bound to their corresponding *cis*-acting elements: exonic splicing enhancer (ESE), exonic splicing silencer (ESS), intronic splicing enhancer (ISE), or intronic splicing silencer (ISS). Silencers are depicted in orange and enhancers in light blue. (c) Schematic representation of alternative splicing events. Gray boxes represent constitutive exons, red and orange boxes represent alternative exons, and a light blue box represents an exonized intron. Discontinuous ends in the exons indicate that the transcript continues in that direction. Thick blue lines represent introns in the transcript before processing, and dashed lines above and below represent the alternative processing of the exons

classes: exonic splicing enhancer (ESE) or silencer (ESS), and intronic splicing enhancer (ISE) or silencer (ISS; Wang & Burge, 2008). Whereas most of these regulatory elements are located near the splice sites, distant ones have also been described to modulate exon inclusion (Lovci et al., 2013). Generally, these sequence motifs are short and represent binding sites for RNA binding proteins (RBPs; Fu & Ares, 2014). These RBPs are the *trans*-acting factors that regulate exon definition and recognize the cis-acting elements with variable affinity and specificity.

Two prominent RBP families that regulate splicing are SR proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs), which often display opposite regulatory roles. SR proteins have one or two RNA-recognition motifs and arginine/serine (RS)-rich domains that mediate protein–protein interactions (Jeong, 2017). They generally bind ESEs to enhance exon inclusion but can exceptionally act as inhibitors by binding ISSs close to the intron boundaries (Bradley et al., 2015; Shen & Mattox, 2012). On the other hand, hnRNPs also show a modular structure with multiple RNA binding and auxiliary domains (Geuens et al., 2016). They bind ESSs to repress splicing, but as SR proteins, they can have an enhancing role depending on their position relative to other sequence elements, for example, by binding ISEs (Erkelenz et al., 2013). Multiple other RBPs have been found to play a role in splicing regulation in different physiological and disease contexts (Yee et al., 2019).

Numerous efforts have attempted to provide a complete census of proteins involved in splicing regulation, uncover their cognate sequence motifs, and establish their role in splicing, which have been jointly dubbed as the splicing code (Barash et al., 2010). These have been primarily based on cross-linking and immunoprecipitation (CLIP) followed by high-throughput sequencing, and have allowed identifying many RBPs and their recognition motifs, as well as other splicing regulators including coactivators or proteins involved in other RNA-processing steps (König et al., 2011; Ule

et al., 2003; Van Nostrand et al., 2020). However, the splicing code remains challenging to decipher due to the many proteins and sequence elements contributing to the splicing outcome. First, the strength of the splice sites, measured as their similarity to the consensus, dictates the susceptibility to regulation by *trans*-acting factors. Strong splice-site motifs facilitate splice-site recognition, but weaker splice sites are more favorable for effective splicing modulation (Baralle & Baralle, 2018). Second, it is still challenging to determine if the binding of an RBP will positively or negatively regulate splicing since many of them have dual roles depending on the context (Fu & Ares, 2014). Finally, and most importantly, splicing patterns are driven by complex combinations of proteins and sequence motifs with cooperative and competitive relationships (Papasaikas et al., 2015; Ule & Blencowe, 2019).

## 2.2 | Variation of the splicing products

The variation in splice site selection is generally described as local variations, which can be categorized into five primary events (Figure 1c; Box 1). The selection of alternative splice sites can make exons shorter or longer, generating alternative 5′ SS (A5) and alternative 3′SS (A3) splicing events. Splicing variation can also affect complete exons that may be spliced in or out of the mature transcript, described as a skipping exon (SE) event. Mutually exclusive exons (ME) are a complex case where one of two adjacent exons may be skipped, and they both are rarely observed together in a mature transcript. Finally, intron retention (IR) occurs when an intron is included in the mature transcript. Different combinations of primary events increase the number of isoforms produced from a single gene changing the UTRs or the coding sequence. Exons are classified as alternatives when they are differentially included in isoforms from the same gene, possibly described according to the event types mentioned above, or constitutive when present in all the transcript variants. On the other hand, these alternative splicing events coexist with other forms of transcript variation. The choice of alternative promoters can induce the selection of alternative first (AF) exons, and alternative polyadenylation (APA) can lead to the inclusion of alternative last (AL) exons (Figure 1c). Due to the involvement of regulatory RBPs that are common to splicing regulation, AF and AL events are often included when studying alternative splicing.

One or more alternative splicing events can give rise to a wide range of transcript-isoform outputs, from two to a few hundred per gene (Ray et al., 2020; Tung et al., 2020), leading to complex patterns of coordinate splicing events

---

**BOX 1    Computational approaches for alternative splicing detection**

Genome-wide detection of alternative splicing (AS) can be based on full-length isoforms, exons, exon–exon junctions, or events. Isoform-based methods use transcript abundances to calculate differential transcript usage (DTU), as in Cuffdiff2 from the Cufflinks package (Trapnell et al., 2010) or SUPPA (Alamancos et al., 2015; Trincado et al., 2018), or isoform switches, as in iso-kTSP (Sebestyén et al., 2015). Both exon- and junction-based analysis methods are generally based on short-read counts. Exon-based methods such as DEXSeq (Anders et al., 2012) identify differential usage of individual exons and exon fragments, which can mask splicing changes since they do not consider neighboring exons. Junction-based methods, such as LeafCutter (Li et al., 2018), use short reads that span two consecutive exons. Event-based methods like rMATS (Shen et al., 2014), MAJIQ (Vaquero-Garcia et al., 2016), or SUPPA (Trincado et al., 2018) analyze alternative splicing choices represented in the form of events, as described in Figure 1c. Isoform and event-based methods are easier to interpret in relation to the AS-related functional impacts than methods based on exons or junctions. The inclusion level of an alternatively spliced event is measured as the Percentage (or proportion) Spliced In (PSI or $\Psi$), and its calculation varies depending on the method. For transcript isoforms, a relative abundance measure similar to PSI can be defined to study DTU and isoform switches (Climente-González et al., 2017). Event- and isoform-based approaches are complementary. However, splicing events are easy to interpret and validate experimentally, but they do not capture as much complex variation as using full-length isoforms (Sebestyén et al., 2015). For this reason, in some cases, both analyses are implemented in the same tool. MISO provides a differential expression of events and isoforms from aligned RNA-seq reads, whereas SUPPA (Alamancos et al., 2015; Trincado et al., 2018) uses transcript abundances for both isoform usage and PSI calculation, providing a direct correspondence between events and isoforms.

(Grosso et al., 2008; Licatalosi & Darnell, 2010; Tapial et al., 2017). How they exactly contribute to functional diversification is still an active area of research. Complex alternative splicing patterns have been described to contribute to the definition of embryonic stem cells, to cell lineage differentiation processes, or to the epithelial–mesenchymal transition (Agosto & Lynch, 2018; Kalsotra & Cooper, 2011; Venables et al., 2013). Differentiated cells and tissues also show specific splicing patterns, as it has been described for the brain, heart, skeletal muscle, testis, and liver, as well as for erythrocytes and immune cells, among others (Baralle & Giudice, 2017; Bhate et al., 2015; Licatalosi & Darnell, 2010; Lynch, 2004; Nakka et al., 2018; Pimentel et al., 2016; Tapial et al., 2017). Finally, alternative splicing is also related to physiologic response to changing conditions, such as DNA damage response, thermal regulation, or even adaptation to stress conditions (Biran et al., 2020; Haltenhof et al., 2020; Shkreta & Chabot, 2015).

## 2.3 | Interplay of splicing with other RNA processing steps

During transcription initiation, expression of different transcription factors (TFs) and changes in chromatin state can lead to the selection of alternative promoters, thereby triggering transcription from an alternative transcription start site (Davuluri et al., 2008). As a result, different exons may be included at the 5′ end of the transcript that can modify the N termini of the protein (Leppek et al., 2018). This may happen as a direct sequence change of the open reading frame (ORF) of the resulting transcript (Tasic et al., 2002) or through a change in the translation start site mediated by alterations in the regulatory sequences of the 5′ untranslated region (5′UTR) of the mRNA (Figure 2). On the other hand, there is often no observable effect in the protein product. Instead, changes in the 5′UTR may affect the translation efficiency, that is, two or more of the transcript isoforms produced from the gene locus encode the same protein but at a different rate. (Tamarkin-Ben-Harush et al., 2017; Figure 2).

The coupling of splicing with transcription has been widely studied through the RNA Polymerase II (RNAPII) function. From the kinetic point of view, it has been shown that elongation rates of RNAPII can affect SS recognition with different outcomes depending on the SS strength (Fong et al., 2014; Saldi et al., 2017). On the other
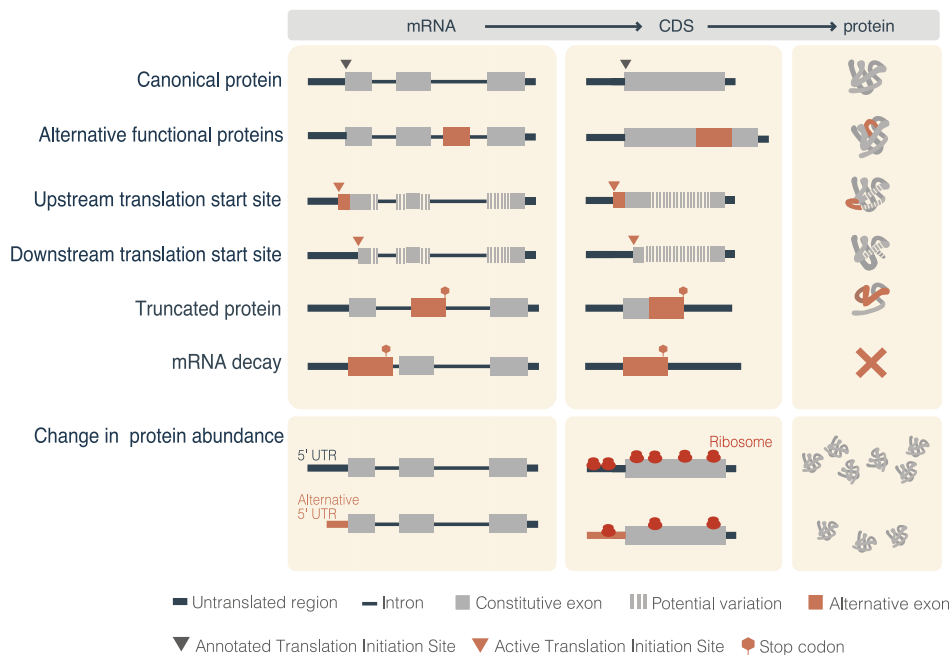


**FIGURE 2** How alternative splicing potentially reshapes the protein products. Alternative splicing (AS) can reshape the proteome in diverse ways. These can be separated into two impacts: changing the protein sequence (upper panels) or altering the amount of protein produced (lower panels). Splicing changes affecting the coding sequences (CDS) can generate alternative proteins with different amino acid sequences, structures, and functions. Similarly, changes in the 5′ end of the mRNA molecule can lead to an alternative initiation site, which results in proteins with different amino-terminal ends. Both mechanisms could also lead to a shift in the reading frame. Premature stop codons generated through a splicing change can translate into truncated proteins or lead to mRNA decay (Box 2). Finally, translation efficiency can be affected by the inclusion of alternative 5′UTRs through alternative splicing

hand, the recruitment model is supported by known interactions of several SFs with RNAPII, mainly through the C-terminal domain (CTD) of the large subunit or with other members of the elongation complex to regulate exon inclusion (Huang et al., 2012; de la Mata & Kornblihtt, 2006; Hsin & Manley, 2012; Das et al., 2007). For example, direct binding of U2AF on the CTD has been shown to recruit Prp19, crucial for spliceosome activation (David et al., 2011). A new approach in a recent study proposes multiple interactions between U1 and RNAPII that support a model of co-transcriptional assembly of the spliceosome on the RNAPII (Zhang, Aibara, et al., 2021). Beyond RNAPII, it has also been observed that the use of alternative promoters, enhancers, and TFs affect splicing rates (Han et al., 2017; Alasoo et al., 2019).

Chromatin state has also been described as affecting the splicing outcome (Agirre et al., 2015, 2021; González-Vallinas et al., 2015). Indeed, DNA methylation and histone posttranslational modifications (PTMs) can also modulate splicing and potentially participate in alternative splicing. Alternatively, spliced exons show lower methylation levels than constitutive ones (Gelfman et al., 2013). Moreover, depending on the factors mediating the crosstalk between methylation and splicing, the effect can be either promote or repress exon inclusion (Shukla et al., 2011; Maunakea et al., 2013; Yearim et al., 2015; Brito et al., 2020). Similarly, distinct histone modification signatures are also known to correlate with specific splicing patterns. This has been observed for *FGFR2*, where combinations of histone methylation marks regulate two mutually exclusive isoforms (Luco et al., 2010). This and other similar cases have led to the proposal of a chromatin code that can predict alternative splicing (Agirre et al., 2015, 2021).

Alternative polyadenylation (APA) is emerging as a relevant mechanism to regulate gene expression in multiple biological processes (Ren et al., 2020). APA can result in significantly different 3′ UTRs, affecting the transcript stability, export, and translation rates (Mayr, 2016; Kishor et al., 2020). Longer 3′ UTRs generally contain more regulatory structures and are related to decreased mRNA stability and translation efficiency (Moqtaderi et al., 2018). Moreover, if APA occurs upstream of the last exon, it can affect the 3′ UTR and the coding region by changing the exon composition at the 3′ end of the transcript. APA in coding regions can regulate gene expression and contribute to protein diversification (Tian & Manley, 2016). Apart from the core 3′ end processing factors, APA can be controlled by the interplay with other RNA-processing steps, including transcription (chromatin state, TFs, and RNAPII activity) and splicing (Elkon et al., 2013; Gruber & Zavolan, 2019). RNA methylation dynamics have also been shown to regulate APA usage, promoting inclusion of terminal exons and therefore controlling the 3′ UTR length (Bartosovic et al., 2017).

Splicing regulators also influence other RNA biology processes that can alter the transcriptome beyond alternative splicing itself. For instance, members of the SR and hnRNP protein families can regulate transcription, processing, stability, and transport (Huang & Steitz, 2005; Ji et al., 2013; Geuens et al., 2016).

Furthermore, translation can be regulated to modulate the amount of protein produced and generate alternative proteins by the alternative selection of Translation Initiation Sites (TISs). Alternative TISs, often downstream of the canonical one, can be selected through several mechanisms, including leaky scanning of the ribosome, presence of non-AUG codons, or through uORFs affecting the capacity of the ribosome to reinitiate translation (Kozak, 2005; Kochetov, 2008). In other cases, the start site of an uORF becomes the actual start site of a 5′ extended version of the translated ORF (Kochetov, 2008). In addition to the variation in the N-termini of the protein, alternative TISs can also shift the translating frame, potentially resulting in different protein products and often creating premature stop codons (Orr et al., 2020).

Finally, biochemical RNA modifications are emerging as a mechanism to modulate the stability and translation efficiency of mRNAs (Boo & Kim, 2020). RNA modifications are abundant in ribosomal (rRNA), and transfer RNAs (tRNA), and their alterations impact translation through the functional impairment of rRNAs and tRNAs (Liu, Clark, et al., 2016; Ma, Wang, et al., 2019). There is now evidence that RNA modifications are also abundant in mRNAs (Dominissini et al., 2012; Squires et al., 2012), and the modifications 5-methylcytidine and N4-acetylcytidine have been linked to enhanced translation efficiency (Arango et al., 2018; Schumann et al., 2020; Huang et al., 2019). Given that there are more than 100 possible modifications (Linder & Jaffrey, 2019), this opens up multiple new possible mechanisms to modulate the protein output from mRNAs.

# 3 | TOWARD IDENTIFYING THE EFFECTS OF SPLICING ON PROTEIN DIVERSITY

Transcriptomics studies have provided insight into how AS gives rise to transcript diversity in the cells. However, given the plethora of mechanisms to control and regulate protein synthesis after pre-mRNA splicing (Box 2), the transcriptome is not a trustworthy representation of the proteome. Then, the key to determining the functional impact of AS

> **BOX 2    Regulation of RNA processing after pre-mRNA splicing**
>
> Cellular mRNAs are subject to multiple layers of regulatory control. A significant proportion of this control occurs at the level of translation. Translation control is a primary regulator of gene expression that can act globally or specifically target certain mRNAs (Schwanhäusser et al., 2011). Global regulation occurs at the initiation step, where reversible protein phosphorylation controls the activity of initiation factors such as eIF2 and eIF4F (Jackson et al., 2010). Molecule-specific regulation requires elements in the mRNA sequence able to influence translation, such as inhibitory secondary structures (hairpins), upstream ORFs (uORFs), or internal ribosome entry sequences (IRES; Gebauer et al., 2012). Other gene regulatory mechanisms involve small noncoding RNAs. Regulation by micro-RNAs (miRNAs) is the most widespread one (Filipowicz et al., 2008), with thousands predicted to operate on about 60% of the protein-coding gene in humans. Their predominant role is silencing by affecting the translation and stability of mRNAs (Friedman et al., 2009; Seiler et al., 2018). In addition to these regulatory mechanisms, there are quality control mechanisms to prevent the production of aberrant proteins that could accumulate and be detrimental to the cell. One of the most studied of these surveillance pathways is Non-Sense Mediated Decay (NMD), which degrades mRNA molecules carrying a premature termination codon (PTC; Kervestin & Jacobson, 2012). Besides its function as a broad surveillance mechanism, there is evidence that NMD act as a specific gene expression regulator for functional transcripts (Nickless et al., 2017). Other ribosome-associated surveillance pathways are Non-stop decay (NSD), which degrades mRNAs lacking a stop codon (Frischmeyer et al., 2002), and No-go decay, which targets mRNAs bound by stalled translation elongation complexes (Pisareva et al., 2011).

is to understand how splicing events reshape the protein products but also to discriminate which alternative transcript isoforms will be translated into protein.

Most genes have one transcript that encodes the most ubiquitously expressed and evolutionary conserved protein isoform, and best represents that gene's function. This has motivated the definition of principal transcripts, which simplifies the transcriptome based on protein sequence, structure, function, and conservation (Rodriguez et al., 2013, 2018). However, this definition clashes with evidence indicating that the most expressed transcript isoform from a gene may vary between tissues and disease conditions (Gonzàlez-Porta et al., 2013; Sebestyén et al., 2015). Changes in the most expressed protein-coding transcript isoform can thus result in a diversification of proteins and their function, leading to changes in catalytic activities, interactions with other proteins, subcellular localization, and loss of function (Yang et al., 2016; Figure 2).

## 3.1 | Proteomics evidence of alternative splicing

One of the primary sources of evidence to support the production of multiple proteins from the same gene locus through AS is mass spectrometry (MS; Blakeley et al., 2010). In a general MS workflow, proteins are digested with enzymes, and the generated peptides are ionized and loaded into a mass spectrometer. The generated spectra are used to computationally reconstruct peptide sequences using a database of annotated proteins as guide (Aebersold & Mann, 2003; Deutsch et al., 2008). Even though the existence of alternatively spliced isoforms from RNA-seq analysis is incontestable, MS experiments have offered very little support for alternative protein products associated with AS, raising controversy in the field (Tress et al., 2017; Blencowe, 2017).

Although the detection power may vary significantly depending on the annotation database used, several studies have provided evidence for hundreds to about one thousand detectable alternative protein isoforms, often showing a tissue-specific pattern (Wilhelm et al., 2014; Kim et al., 2014; Tay et al., 2015). A study of 30 human samples identified proteins corresponding to 84% of the total annotated protein-coding genes and mapped peptides uniquely to approximately 400 alternative protein isoforms (Kim et al., 2014). In contrast, analysis of eight large-scale human proteomics datasets with highly conservative filters yielded evidence for about 200 alternative protein isoforms (Ezkurdia et al., 2015; Abascal et al., 2015). It has also been claimed that there is an overestimation of alternative protein isoforms resulting from considering cases where peptides only map to one side of the variable exon in a splicing event rather than both sides, and from not taking into account the potential false positives in the MS data processing (Tress et al., 2017).

Although a significant amount of high-throughput MS data have been produced to date for multiple cell models and tissues, studies using this data to search for protein isoform variants have almost invariably been performed in steady-state conditions (Alfaro et al., 2017; Nusinow et al., 2020). This contrasts with the fact that most RNA splicing changes have been reported in association with changing physiological conditions or between normal and disease states. Supporting this notion, several studies have provided evidence for several genes with detectable protein isoform variation across the cell cycle in leukemia cells (Ly et al., 2014).

Integration of RNA-seq experiments with MS for the identification of protein isoforms, also known as proteogenomics, has been implemented to reduce peptide mapping uncertainty and to detect novel peptides (Jeong et al., 2018; Liu et al., 2017; Wu et al., 2019; Lau et al., 2019; Agosto et al., 2019). In the context of depletion of a spliceosome component and using quantitative proteomics, it was shown that out of 450 transcripts with significant changes in RNA relative abundance, up to 160 showed consistent changes in protein production using quantitative proteomics (Liu et al., 2017). Similarly, a splice-junction-centric approach applied to public proteogenomics data led to the identification of 1500 alternative protein isoforms (Lau et al., 2019). Most of the junctions (~60%) arose from SE events, consistent with what has been observed in RNA-seq experiments.

Proteogenomics studies have raised the issue of the poor correlation between transcript and protein abundances, which is usually below 0.5 (Pearson correlation coefficient) and varies across tissues (Kosti et al., 2016; Eraslan et al., 2019; Liu et al., 2017). Despite the general lack of consistency between transcript and protein abundances, it has been observed that above certain levels of transcript expression, RNA-seq is a good predictor of protein expression (Vogel & Marcotte, 2012). A recent study has modeled the effect of protein synthesis regulation after pre-mRNA splicing and has managed to explain 30% of the variance in the protein-to-mRNA ratios (Eraslan et al., 2019).

Many of the studies mentioned above point out that spectra generated in MS experiments might not provide enough resolution to distinguish particular isoforms (Tay et al., 2015; Liu et al., 2017; Kosti et al., 2016; Eraslan et al., 2019). On the one hand, these limitations could be attributed to the digestion enzymes used in proteomics. Trypsin is the standard protease in shotgun proteomics. It efficiently digests at K or R residues, resulting in small size (<3 kDa) peptides with a positive charge at the C-terminus, which is optimal for the MS protocols (Huang et al., 2005; Laskay et al., 2013). The main drawbacks of the peptides produced are the short lengths (around six aa) and the limited proportion of the covered proteome (Giansanti et al., 2016). Other proteases such as chymotrypsin, LysC, LysN, AspN, GluC, ArgC, elastase, and proteinase K have been shown to cover complementary fractions of the proteome, and proteases such as GluC improve the detection of peptides containing PTMs, which are otherwise overlooked (Janssen et al., 2019); suggesting that combinations of enzymes may provide the best approach to obtain a comprehensive peptide identification (Dau et al., 2020; Giansanti et al., 2016). Regarding splicing products, different digestion enzymes used in proteomics are biased toward identifying certain splicing junctions with small overlap among them (Wang et al., 2018). On the other hand, since transcript isoforms share a considerable proportion of their sequence, isoform-specific peptides are hard to identify (Tran et al., 2017). Nonetheless, proteogenomic approaches have improved the detection of peptides that can be uniquely matched to protein isoforms (Agosto et al., 2019). An added complexity in MS analysis is sequence identification from peptide mass spectra. Since the application of de novo peptide sequencing to MS remains challenging, the current analysis relies on database searches and can lead to false positives and neglect certain peptides (Bogdanow et al., 2016; Karunratanakul et al., 2019). In particular, a significant proportion of missed and unassigned peptides carry PTMs that shift the mass, which strongly impacts the database searches (Chick et al., 2015). Taken together, all these factors support the notion that there is an underestimation by MS experiments of the number of protein isoforms occurring in the cell.

## 3.2 | Predicting the functional impact of alternative splicing from the sequence

In parallel to the proteomics approaches, other studies have tried to address the functional impact of alternative splicing using computational models to predict the function of the resulting coding and protein sequences. The most straightforward approach has been to analyze the effect of splicing events on the coding sequence (CDS). For instance, in VastDB (Tapial et al., 2017), predictions are classified into ORF-preserving events, which have the potential to generate a fully functional protein, NMD targets, or ORF-disrupting events, including both frame-shifted and truncated CDS. However, these predictions are based on the impact of specific local events in isolation, without considering what happens elsewhere in the transcript, which does not reflect the complete splicing change. Indeed, a study of the functional impacts of isoform switches in cancer indicated that about 70% of these changes correspond to a single local splicing event. In

contrast, the rest include changes that cannot be described in terms of the primary events (Climente-González et al., 2017). Addressing this issue, the method ISOTOPE calculates the modified ORFs resulting from all the splicing alterations in the gene to calculate possible novel epitopes in tumor samples (Trincado et al., 2021).

Other studies have measured the impact in terms of the alteration in functional domains. Such approaches map protein domains and functional elements to transcript isoforms to identify which functions are lost or gained due to AS (Hiller et al., 2005; Sulakhe et al., 2019; de la Fuente et al., 2020; Climente-González et al., 2017). For instance, one of these studies found that a large proportion of the isoform switches observed in cancer impacted functional domain families that are also targeted by cancer driver mutations (Climente-González et al., 2017). In a different study, integrating the functional annotations with differential transcript usage (DTU) showed that ~80% of the genes with DTU between two different mouse neural precursor cells present differences in the content of the functional features (de la Fuente et al., 2020).

The impact on protein structure can also be used to evaluate the potential functional variability resulting from AS. Variable exonic regions have been observed to generally code for disordered regions in proteins, which has been interpreted as enabling functional and regulatory diversity as disorder regions confer conformational flexibility that can be accommodated in the context of structured proteins (Hegyi et al., 2011; Romero et al., 2006). Modeling the impact of splicing on protein structure faces significant challenges since very few protein isoforms have a resolved 3D protein structure, and it remains difficult to obtain structures for proteins without a stable fold (Melamud & Moult, 2009). To address these challenges, Hao et al. (2015) implemented a semi-supervised learning approach based on structural information, and that does not require a negative set to conclude that 32% of alternatively protein isoforms result from SE events lead to functional proteins. Cross-species conservation has also been used as evidence for the functionality of protein isoform variants (Mazin et al., 2021; Rodriguez et al., 2013; Tapial et al., 2017). In fact, phylogenetically conserved isoform variants generally show higher expression from mRNA and a higher detection rate in proteomics experiments than nonconserved ones (Ezkurdia et al., 2012; Cummings et al., 2020).

## 3.3 | Effects on functional networks of protein interactions

One of the most intriguing functional effects of AS, and with potentially the farthest-reaching consequences, is the modulation of protein–protein interaction (PPI) networks. It has been proposed that proteins with associated tissue-specific splicing events are central hubs in interaction networks, having distinct interaction partners in different tissues (Buljan et al., 2012). Similarly, isoform switches in various cancer types have also been proposed to remodel PPIs involving cancer drivers and cancer-related pathways (Climente-González et al., 2017). Using in vitro experiments, it was observed that the inclusion of neuronally regulated exons significantly promoted or disrupted protein interactions using co-immunoprecipitation assays (Ellis et al., 2012). In addition, yeast-two-hybrid experiments revealed how to splice variants expand the network of interactions in the autism spectrum disorders (Corominas et al., 2014). This mechanism was further studied at genome-scale to reveal that less than 50% of interactions are shared among pairs of protein isoforms (Yang et al., 2016). Despite the vast range of the predicted isoform-specific interactions in tissues and disease conditions, most PPI databases only capture the interactome of canonical protein isoforms and describe these in terms of one protein per gene. Recently, the database DIGGER has provided a way to retrieve interactions between domains and proteins at the isoform and exon levels (Louadi et al., 2021). Understanding the rewiring of functional networks driven by AS can shed light on how even minor changes in the mRNA and protein sequences can trigger significant functional effects on the cell.

All the described studies have added multiple layers of evidence for an impact of the transcriptome variation on the proteome and its functional repertoire that appears to be more extensive than what can be directly observed from proteomics experiments alone. However, one could argue that even controlling for NMD or using additional structural and functional information, most studies have been based on computational predictions or in vitro experiments and require further experimental validation using cell or in vivo models, circling back to proteomics and its limitations. Below we describe alternative methodologies that address these limitations.

## 4 | TRANSLATOMICS: FILLING THE GAP BETWEEN MRNA AND THE PROTEIN

The basis for translatome analysis was already set more than 50 years ago (Warner et al., 1963; Steitz, 1969). However, it has been only recently that, thanks to combining those principles with high-throughput sequencing, translatomics

has emerged as a field with the potential to capture the elements involved in mRNA translation in a transcriptome-wide manner (Zhao et al., 2019). We will focus on how translatomics techniques provide evidence of active translation of mRNAs and how this has been used to measure how AS can impact the proteome.

## 4.1 | High-throughput measurement of translating mRNAs

The most widely used techniques for capturing translating mRNAs, polysome, and ribosome profiling, are based on the high-throughput sequencing of transcripts bound by ribosomes (Figure 3). Polysome profiling experiments are based on sucrose gradient separation and ultracentrifugation of full-length mRNAs, which have different sedimentation speeds according to the number of ribosomes sitting on them (Lackner et al., 2007; Mašek et al., 2011). This makes possible the stratification of different fractions containing mRNAs engaged by one (monosomes) or multiple (polysomes) ribosomes, indicating potentially different translation activities (Chassé et al., 2016; Floor & Doudna, 2016). Ribosomes are detached to perform high-throughput sequencing in each fraction, like standard RNA sequencing for transcriptomics analyses, providing a straightforward measure of translation rates (Maslon et al., 2014). Ribosome profiling or Ribo-seq follows a different approach (Ingolia et al., 2009). The segments of mRNA not bound by fully assembled ribosomes are
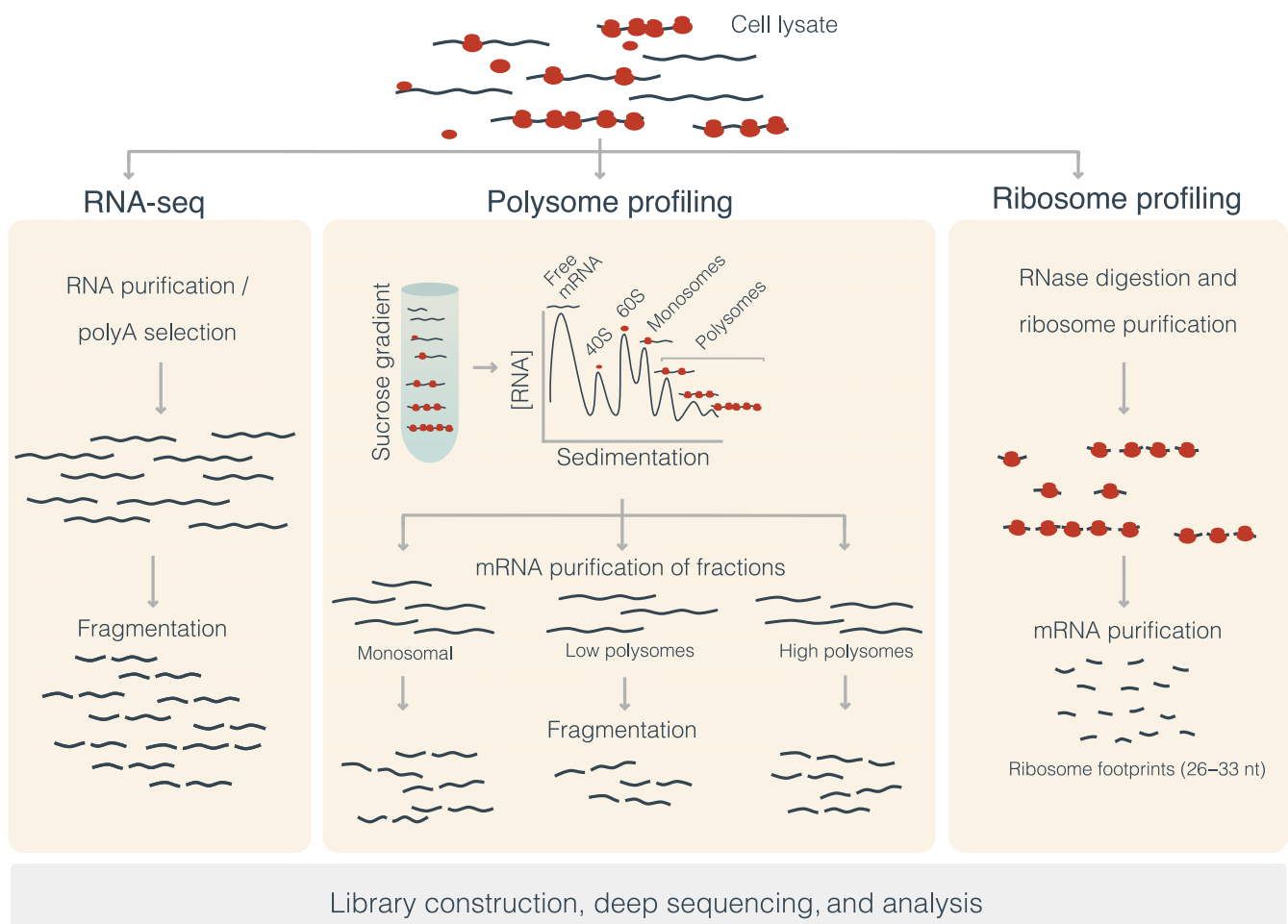


**FIGURE 3** High-throughput approaches to capture the translatome. RNA sequencing can capture all mRNAs in the cell lysate regardless of its translation status (left panel). Polysome profiling (middle panel) separates fractions of mRNAs according to the number of ribosomes attached to the mRNAs, which is assumed to correlate with the translation activity. Downstream sequencing of RNA in each polysomal fraction is performed as a standard RNA-seq experiment, and the relative mRNAs abundances are interpreted as mRNA translation activities. In ribosome profiling or Ribo-seq (right panel), mRNAs bound by the ribosomes are digested, producing footprints of approximately 30 nt representing the mRNA sites occupied by the ribosome. These fragments are then sequenced after detaching the ribosomes

digested, leaving intact only ribosome-protected mRNA fragments of about 26–33 nucleotides (ribosome footprints), which are isolated for sequencing.

One advantage of Ribo-seq over polysome profiling is that it provides the exact position where ribosomes are sitting in the mRNA molecule, that is, the ribosome footprint. Positional information enables the identification of start codon positions, codon composition effects on translation, detection of translating uORFs, or stalling (Zhang, Wang, et al., 2021; Ishimura et al., 2014). Moreover, the ribosome footprints can be assessed in terms of their uniformity along with the transcript and their deviations from the expected three-nucleotide periodicity in the translation frames. Both measures are considered to be essential to determine active translation (Ji, 2018). However, Ribo-seq experiments only capture fully assembled elongating ribosomes, that is, they do not provide information about the translation initiation and termination steps. A modification of the technique named translation complex profile sequencing (TCP-seq) enabled to additionally retrieval footprints of the small ribosome subunit, expanding the picture to all stages of translation: initiation, elongation, and termination (Archer et al., 2016; Shirokikh et al., 2017). Translatomics approaches have been shown to significantly improve correlation with proteomics compared with transcriptomics-based approaches (Wang et al., 2013; Blevins et al., 2019). Hence they represent promising alternatives to overcome the limitations of proteomics in the study of splicing-associated alterations of the proteome.

## 4.2 | Measuring alternative splicing from the translatome

Most of the initial studies on translatome profiling were conducted at gene level (Gonzalez et al., 2014; Ruiz-Orera et al., 2014; Ingolia et al., 2009). However, as these experiments involve sequencing of transcriptome or transcriptome fragments, it is possible to apply transcript-isoform-aware tools to translatomics data to investigate how splicing impacts translation dynamics. In this regard, polysome profiling indicated that ~30% of splicing events show differences in polyribosome association that could be explained by remodeling of cis-acting elements (Sterne-Weiler et al., 2013). Similarly, translation regulation by SRSF1 was observed from polysomal fractions in mRNAs that display SRSF1-dependent splicing changes, providing evidence for SRSF1 involvement in the coupling of nuclear mRNA processing events with translation (Maslon et al., 2014). Furthermore, different isoforms from the same gene have been shown to have variable translation potential during neural development using mRNAs from high polysome fractions (Blair et al., 2017).

There have also been several translatome studies seeking evidence of translation of alternative isoforms and AS events using Ribo-seq experiments. In 2016, Weatheritt et al. mapped ribosome footprints across the exon–exon junctions of alternative splicing events, showing that 75% of skipping exon events from sufficiently expressed transcripts were engaged by the ribosomes (Weatheritt et al., 2016). However, it can be argued that ribosome engagement alone does not necessarily imply active translation of the full-length isoform. To address this question, several groups have developed methods for isoform-level quantification from Ribo-seq data (Wang et al., 2016; Reixachs-Solé et al., 2020; Calviello et al., 2020). ORFquant detected alternative ORFs being translated in some genes, although most genes showed one dominant isoform (Calviello et al., 2020). ORQAS pipeline showed higher sensitivity in detecting alternative isoforms, finding 33–53% of the expressed genes in mouse and human samples have more than one translated isoform. Additionally, 40% of the splicing changes at the RNA level were shown to be concordant with differences in translation (Reixachs-Solé et al., 2020).

These and other studies have consistently reported that ribosome engaged splicing events are enriched in characteristics and functional elements that are relevant to specific physiological conditions, such as phosphorylation sites in regenerating liver, methylation islands in models of fragile X syndrome, or microexons (exons of length 3–51 nt) in glioma samples (Shah et al., 2020; Bhate et al., 2015; Reixachs-Solé et al., 2020). These results also provided additional evidence for the contribution of alternative splicing to the definition of a specific cell, organ, and organismal phenotypes; and highlighted the relevance of studying the functional impact of alternative splicing in multiple physiological conditions and diseases.

Long read sequencing technologies enable the discovery and reconstruction of full-length transcripts (Kovaka et al., 2019; Tang et al., 2020; Hu et al., 2021). This significantly improves the ability to identify potentially functional transcripts (Troskie et al., 2021) and provides the opportunity to define more accurate and complete references for translatomics studies (Wang et al., 2019; Sun et al., 2021). A recent study showed increased detection of actively translated alternative isoforms in rat hippocampus using full-length transcripts compared to the available reference annotation (Wang et al., 2019). In another study, single-molecule long-read sequencing combined with ribosome profiles identified evidence of translation in RNAs in sperm that would be otherwise challenging to assemble using short

sequencing reads (Sun et al., 2021). As long reads may solve the mapping uncertainty present in short reads, they could also improve ribosome profiling quantification by adapting existing methods that use transcript abundances as a guide to resolving ambiguity in Ribo-seq mapping (Wang et al., 2016; Reixachs-Solé et al., 2020).

# 5 | CONCLUSION

Given the lack of accurate models to establish protein abundances from mRNA sequencing and the intrinsic limitations of the different approaches (summarized in Table 1), integration of multiple omics techniques addressing different stages of protein production remains to date the best strategy to study the impact of AS on the proteome and function. Although there is still some disagreement on the precise extent of the functional effects of AS, there is plenty of evidence that AS influences protein products. One of the most challenging obstacles is that existing proteomics methods may not provide enough sampling depth to keep up with the large number of splice variants identified from RNA-seq. Therefore, it remains challenging to validate all the predicted variants at the protein level. An integral approach to overcome this limitation is to combine RNA-seq information with proteomics analysis, which has been shown to significantly improve peptide matches (Ma, Liu, et al., 2019).

Measuring the relation between transcriptome and translatome changes has emerged as a powerful way to identify alternative splicing events that affect the protein product and, therefore, potentially impact function (Weatheritt et al., 2016; Reixachs-Solé et al., 2020). Knowledge of the mRNA abundances from the transcriptome can help overcome the drawbacks of translatomics analysis. We can find a clear example of this in ribosome profiling. The short length of the fragments represents one of the main limitations for isoform quantification that can be addressed by considering isoform abundances from RNA-seq (Wang et al., 2016; Reixachs-Solé et al., 2020). At the same time, it has been shown that the incorporation of ribosome footprints as priors can improve the accuracy of protein isoform abundances estimated from MS (Carlyle et al., 2018). New protocols such as TCP-seq have the potential to shed light on how splicing variation in noncoding regions (5′ UTR) can influence translation (Archer et al., 2016; Shirokikh et al., 2017). Thus, analysis of TCP-seq data in an isoform-aware fashion could uncover new mechanisms of translational control coupled with transcript initiation and alternative splicing. One potential limitation of polysome profiling and ribosome footprinting experiments is that they involve lengthy and complex protocols. Improvements on these protocols to make them easier to implement and combine with high-throughput platforms could benefit the study of the impacts of alternative splicing on the proteome (Yoshikawa et al., 2021).

**TABLE 1** Capabilities and limitations of current omics techniques for the detection of splicing variants

|  | Transcriptome RNA-seq | Translatome | | Proteome MS |
|  |  | Polysome profiling | Ribo-seq |  |
| --- | --- | --- | --- | --- |
| Sample processing | RNase digestion or chemical fragmentation | | RNase digestion | Protease digestion |
| Fragment length | 75–150 nt | | ~30 nt | ~10 aa |
| Sampling biases | Differences in PCR efficiency, overrepresentation of highly expressed mRNAs, lack of data for mRNAs with low expression. | | | Variable coverage depending on the protease used |
| Molecular biases | Underrepresentation of non-poly A+ RNAs (in RNA-seq with poly-A+ selection), nuclear RNAs, lncRNA[a] | | | Underrepresentation of transmembrane proteins, PTMs, nonannotated variants, and translated noncoding regions |
|  | Overrepresentation of rRNA (if rRNA depletion is not conducted), globin RNAs in blood samples, and so forth | | | Tissue-specific overrepresentation of proteins |
| Splicing events detected | Events in coding regions, UTRs, and NMD targets | Events in coding regions, UTRs, and NMD targets[b] | Events in coding regions and NMD targets[c] | Events involving coding regions |

[a]lncRNA generally has low expression and are underrepresented but can still be translated into short peptides (Ruiz-Orera et al., 2014).
[b]NMD targets may still be engaged by the ribosome but not translated; hence they will be pulled down in monosome or low polysomal fractions (Heyer & Moore, 2016; Karousis et al., 2020).
[c]With Ribo-seq, elongating ribosomes can be discriminated from stalled ones using measures of periodicity and uniformity (Ji, 2018).

The integration of polysome profiling and ribosome footprinting with long-read sequencing would be especially beneficial to the study of the translatome. Long-read sequencing approaches would remove the uncertainty of transcript reconstruction and abundance estimation inherent in short reads. Full-length transcripts and their abundances could be thus defined for the same conditions as in the Ribo-seq experiments to describe translation variation and complexity accurately. In addition, long-reads can extend the number of relevant RNAs from traditionally noncoding coding or nonfunctional species, such as lncRNAs and pseudogenes (Troskie et al., 2021; Ruiz-Orera et al., 2014). Similarly, integration of long-read sequencing with polysome profiling would also improve the characterization of translating mRNAs. Overall, everything points to long-read sequencing as a necessary technology to obtain a more reliable interpretation of Ribosome profiling experiments and potentially of the assignment of MS peptides to transcripts.

Finally, it is worth highlighting that ribosome engagement and translation rates do not necessarily correlate with stable protein levels. Protein degradation, either co-translational or posttranslational, can eliminate aberrant proteins that escaped previous quality control steps and tune cellular activity in specific conditions by modulating levels of functional proteins (Liu, Beyer, & Aebersold, 2016; Joazeiro, 2019). Thus, despite all the advances to link the transcriptome variations to protein production and function, we still need to measure the proteome. Proteomics has been limiting so far to provide evidence of these protein variants for which there are multiple other sources of evidence. On the other hand, emerging single-molecule protein sequencing technologies (Alfaro et al., 2021) could give a genuinely unbiased measurement of proteomes and, more precisely, unveil the extent to which alternative splicing impacts the proteome and, thereby, the function of genes.

## CONFLICT OF INTEREST
The authors declare no conflicts of interest.

## AUTHOR CONTRIBUTIONS
**Marina Reixachs-Solé:** Conceptualization (lead); data curation (lead); resources (lead); visualization (lead); writing – original draft (lead); writing – review and editing (lead). **Eduardo Eyras:** Conceptualization (supporting); writing – original draft (supporting); writing – review and editing (supporting).

## DATA AVAILABILITY STATEMENT
Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID
*Marina Reixachs-Solé* https://orcid.org/0000-0001-9446-8082
*Eduardo Eyras* https://orcid.org/0000-0003-0793-6218

## RELATED WIREs ARTICLES
The evolution of posttranscriptional regulation
Microexons: Discovery, regulation, and function
Intron retention and its impact on gene expression and protein diversity: A review and a practical guide

## REFERENCES
Abascal, F., Ezkurdia, I., Rodriguez-Rivas, J., Rodriguez, J. M., del Pozo, A., Vázquez, J., Valencia, A., & Tress, M. L. (2015). Alternatively spliced homologous exons have ancient origins and are highly expressed at the protein level. *PLoS Computational Biology*, *11*, e1004325.
Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*, *422*, 198–207.
Agirre, E., Bellora, N., Alló, M., Pagès, A., Bertucci, P., Kornblihtt, A. R., & Eyras, E. (2015). A chromatin code for alternative splicing involving a putative association between CTCF and HP1α proteins. *BMC Biology*, *13*, 31.
Agirre, E., Oldfield, A. J., Bellora, N., Segelle, A., & Luco, R. F. (2021). Splicing-associated chromatin signatures: A combinatorial and position-dependent role for histone marks in splicing definition. *Nature Communications*, *12*, 682.
Agosto, L. M., Gazzara, M. R., Radens, C. M., Sidoli, S., Baeza, J., Garcia, B. A., & Lynch, K. W. (2019). Deep profiling and custom databases improve detection of proteoforms generated by alternative splicing. *Genome Research*, *29*, 2046–2055.
Agosto, L. M., & Lynch, K. W. (2018). Alternative pre-mRNA splicing switch controls hESC pluripotency and differentiation. *Genes & Development*, *32*, 1103–1104.

Alamancos, G. P., Pages, A., Trincado, J. L., Bellora, N., & Eyras, E. (2015). Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA*, *21*, 1521–1531.

Alasoo, K., Rodrigues, J., Danesh, J., Freitag, D. F., Paul, D. S., & Gaffney, D. J. (2019). Genetic effects on promoter usage are highly context-specific and contribute to complex traits. *eLife*, *8*, e41763.

Alfaro, J. A., Bohländer, P., Dai, M., Filius, M., Howard, C. J., van Kooten, X. F., Ohayon, S., Pomorski, A., Schmid, S., Aksimentiev, A., Anslyn, E. V., Bedran, G., Cao, C., Chinappi, M., Coyaud, E., Dekker, C., Dittmar, G., Drachman, N., Eelkema, R., ... Joo, C. (2021). The emerging landscape of single-molecule protein sequencing technologies. *Nature Methods*, *18*, 604–617.

Alfaro, J. A., Ignatchenko, A., Ignatchenko, V., Sinha, A., Boutros, P. C., & Kislinger, T. (2017). Detecting protein variants by mass spectrometry: A comprehensive study in cancer cell-lines. *Genome Medicine*, *9*, 62.

Anders, S., Reyes, A., & Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Research*, *22*, 2008–2017.

Arango, D., Sturgill, D., Alhusaini, N., Dillman, A. A., Sweet, T. J., Hanson, G., Hosogane, M., Sinclair, W. R., Nanan, K. K., Mandler, M. D., Fox, S. D., Zengeya, T. T., Andresson, T., Meier, J. L., Coller, J., & Oberdoerffer, S. (2018). Acetylation of Cytidine in mRNA promotes translation efficiency. *Cell*, *175*, 1872–1886.e24.

Archer, S. K., Shirokikh, N. E., Beilharz, T. H., & Preiss, T. (2016). Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature*, *535*, 570–574.

Baralle, F. E., & Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nature Reviews. Molecular Cell Biology*, *18*, 437–451.

Baralle, M., & Baralle, F. E. (2018). The splicing code. *Biosystems*, *164*, 39–48. https://pubmed.ncbi.nlm.nih.gov/29122587/

Barash, Y., Calarco, J. A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B. J., & Frey, B. J. (2010). Deciphering the splicing code. *Nature*, *465*, 53–59.

Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., Kim, T., Misquitta-Ali, C. M., Wilson, M. D., Kim, P. M., Odom, D. T., Frey, B. J., & Blencowe, B. J. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science (New York, N.Y.)*, *338*(6114), 1587–1593.

Bartosovic, M., Molares, H. C., Gregorova, P., Hrossova, D., Kudla, G., & Vanacova, S. (2017). N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3′-end processing. *Nucleic Acids Research*, *45*, 11356–11370.

Berget, S. M., Moore, C., & Sharp, P. A. (1977). Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, *74*, 3171–3175.

Bhate, A., Parker, D. J., Bebee, T. W., Ahn, J., Arif, W., Rashan, E. H., Chorghade, S., Chau, A., Lee, J. H., Anakk, S., et al. (2015). ESRP2 controls an adult splicing programme in hepatocytes to support postnatal liver maturation. *Nature Communications*, *6*, 1–12.

Biran, J., Gliksberg, M., Shirat, I., Swaminathan, A., Levitas-Djerbi, T., Appelbaum, L., & Levkowitz, G. (2020). Splice-specific deficiency of the PTSD-associated gene PAC1 leads to a paradoxical age-dependent stress behavior. *Scientific Reports*, *10*(1), 9559. https://pubmed.ncbi.nlm.nih.gov/32533011/

Blair, J. D., Hockemeyer, D., Doudna, J. A., Bateup, H. S., & Floor, S. N. (2017). Widespread translational remodeling during human neuronal differentiation. *Cell Reports*, *21*, 2005–2016.

Blakeley, P., Siepen, J. A., Lawless, C., & Hubbard, S. J. (2010). Investigating protein isoforms via proteomics: A feasibility study. *Proteomics*, *10*, 1127–1140.

Blencowe, B. J. (2017). The relationship between alternative splicing and proteomic complexity. *Trends in Biochemical Sciences*, *42*, 407–408.

Blevins, W. R., Tavella, T., Moro, S. G., Blasco-Moreno, B., Closa-Mosquera, A., Díez, J., Carey, L. B., & Albà, M. M. (2019). Extensive post-transcriptional buffering of gene expression in the response to severe oxidative stress in baker's yeast. *Scientific Reports*, *9*, 11005.

Bogdanow, B., Zauber, H., & Selbach, M. (2016). Systematic errors in peptide and protein identification and quantification by modified peptides. *Molecular & Cellular Proteomics*, *15*, 2791–2801.

Boo, S. H., & Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. *Experimental & Molecular Medicine*, *52*, 400–408.

Bradley, T., Cook, M. E., & Blanchette, M. (2015). SR proteins control a complex network of RNA-processing events. *RNA*, *21*, 75–92.

Braun, K. A., & Young, E. T. (2014). Coupling mRNA synthesis and decay. *Molecular and Cellular Biology*, *34*, 4078–4087.

Brito, D. V. C., Gulmez Karaca, K., Kupke, J., Frank, L., & Oliveira, A. M. M. (2020). MeCP2 gates spatial learning-induced alternative splicing events in the mouse hippocampus. *Molecular Brain*, *13*, 1–16.

Buljan, M., Chalancon, G., Eustermann, S., Wagner, G. P., Fuxreiter, M., Bateman, A., & Babu, M. M. (2012). Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Molecular Cell*, *46*, 871–883.

Calviello, L., Hirsekorn, A., & Ohler, U. (2020). Quantification of translation uncovers the functions of the alternative transcriptome. *Nature Structural & Molecular Biology*, *27*, 717–725.

Carlyle, B. C., Kitchen, R. R., Zhang, J., Wilson, R. S., Lam, T. T., Rozowsky, J. S., Williams, K. R., Sestan, N., Gerstein, M. B., & Nairn, A. C. (2018). Isoform-level interpretation of high-throughput proteomics data enabled by deep integration with RNA-seq. *Journal of Proteome Research*, *17*, 3431–3444.

Chassé, H., Boulben, S., Costache, V., Cormier, P., & Morales, J. (2016). Analysis of translation using polysome profiling. *Nucleic Acids Research*, *45*, e15.

Chick, J. M., Kolippakkam, D., Nusinow, D. P., Zhai, B., Rad, R., Huttlin, E. L., & Gygi, S. P. (2015). A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. *Nature Biotechnology*, *33*, 743–749. https://pubmed.ncbi.nlm.nih.gov/26076430/

Chow, L. T., Gelinas, R. E., Broker, T. R., & Roberts, R. J. (1977). An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. *Cell*, *12*, 1–8.

Climente-González, H., Porta-Pardo, E., Godzik, A., & Eyras, E. (2017). The functional impact of alternative splicing in cancer. *Cell Reports*, *20*, 2215–2226.

Corominas, R., Yang, X., Lin, G. N., Kang, S., Shen, Y., Ghamsari, L., Broly, M., Rodriguez, M., Tam, S., Trigg, S. A., Fan, C., Yi, S., Tasan, M., Lemmens, I., Kuang, X., Zhao, N., Malhotra, D., Michaelson, J. J., Vacic, V., ... Iakoucheva, L. M. (2014). Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. *Nature Communications*, *5*, 3650.

Corvelo, A., Hallegger, M., Smith, C. W. J., & Eyras, E. (2010). Genome-wide association between branch point properties and alternative splicing. *PLoS Computational Biology*, *6*, e1001016.

Cummings, B. B., Karczewski, K. J., Kosmicki, J. A., Seaby, E. G., Watts, N. A., Singer-Berk, M., Mudge, J. M., Karjalainen, J., Satterstrom, F. K., O'Donnell-Luria, A. H., Poterba, T., Seed, C., Solomonson, M., Alföldi, J., Genome Aggregation Database Production Team, Genome Aggregation Database Consortium, Daly, M. J., & MacArthur, D. (2020). Transcript expression-aware annotation improves rare variant interpretation. *Nature*, *581*, 452–458. https://pubmed.ncbi.nlm.nih.gov/3246165

Das, R., Yu, J., Zhang, Z., Gygi, M. P., Krainer, A. R., Gygi, S. P., & Reed, R. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Molecular Cell*, *26*, 867–881.

Dau, T., Bartolomucci, G., & Rappsilber, J. (2020). Proteomics using protease alternatives to trypsin benefits from sequential digestion with trypsin. *Analytical Chemistry*, *92*, 9523–9527.

David, C. J., Boyne, A. R., Millhouse, S. R., & Manley, J. L. (2011). The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes & Development*, *25*, 972–983.

Davuluri, R. V., Suzuki, Y., Sugano, S., Plass, C., & Huang, T. H. M. (2008). The functional consequences of alternative promoter use in mammalian genomes. *Trends in Genetics*, *24*, 167–177.

de la Fuente, L., Arzalluz-Luque, Á., Tardáguila, M., del Risco, H., Martí, C., Tarazona, S., Salguero, P., Scott, R., Lerma, A., Alastrue-Agudo, A., Bonilla, P., Newman, J. R. B., Kosugi, S., McIntyre, L. M., Moreno-Manzano, V., & Conesa, A. (2020). tappAS: A comprehensive computational framework for the analysis of the functional impact of differential splicing. *Genome Biology*, *21*, 119.

de la Mata, M., & Kornblihtt, A. R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nature Structural & Molecular Biology*, *13*, 973–980.

Deutsch, E. W., Lam, H., & Aebersold, R. (2008). Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. *Physiological Genomics*, *33*, 18–25.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., & Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, *485*, 201–206.

Elkon, R., Ugalde, A. P., & Agami, R. (2013). Alternative cleavage and polyadenylation: Extent, regulation and function. *Nature Reviews. Genetics*, *14*, 496–506.

Ellis, J. D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J. A. A., Wang, X., Pan, Q., O'Hanlon, D., Kim, P. M. M., Wrana, J. L., & Blencowe, B. J. (2012). Tissue-specific alternative splicing remodels protein-protein interaction networks. *Molecular Cell*, *46*, 884–892.

Eraslan, B., Wang, D., Gusic, M., Prokisch, H., Hallström, B. M., Uhlén, M., Asplund, A., Pontén, F., Wieland, T., Hopf, T., Hahne, H., Kuster, B., & Gagneur, J. (2019). Quantification and discovery of sequence determinants of protein-per-mRNA amount in 29 human tissues. *Molecular Systems Biology*, *15*, e8513.

Erkelenz, S., Mueller, W. F., Evans, M. S., Busch, A., Schöneweis, K., Hertel, K. J., & Schaal, H. (2013). Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA*, *19*, 96–102.

Ezkurdia, I., del Pozo, A., Frankish, A., Rodriguez, J. M., Harrow, J., Ashman, K., Valencia, A., & Tress, M. L. (2012). Comparative proteomics reveals a significant bias toward alternative protein isoforms with conserved structure and function. *Molecular Biology and Evolution*, *29*, 2265–2283.

Ezkurdia, I., Rodriguez, J. M., Carrillo-de Santa Pau, E., Vázquez, J., Valencia, A., & Tress, M. L. (2015). Most highly expressed protein-coding genes have a single dominant isoform. *Journal of Proteome Research*, *14*, 1880–1887.

Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nature Reviews. Genetics*, *9*, 102–114.

Floor, S. N., & Doudna, J. A. (2016). Tunable protein synthesis by transcript isoforms in human cells. *eLife*, *5*, e10921.

Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X. D., & Bentley, D. L. (2014). Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes & Development*, *28*, 2663–2676.

Friedman, R. C., Farh, K. K.-H., Burge, C. B., & Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, *19*, 92–105. https://pubmed.ncbi.nlm.nih.gov/18955434/

Frischmeyer, P. A., van Hoof, A., O'Donnell, K., Guerrerio, A. L., Parker, R., & Dietz, H. C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*, *295*, 2258–2261.

Fu, X. D., & Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. *Nature Reviews. Genetics*, *15*, 689–701.

Gebauer, F., Preiss, T., & Hentze, M. W. (2012). From cis-regulatory elements to complex RNPs and back. *Cold Spring Harbor Perspectives in Biology*, *4*, a012245.

Gelfman, S., Cohen, N., Yearim, A., & Ast, G. (2013). DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon-intron structure. *Genome Research*, *23*, 789–799.

Geuens, T., Bouhy, D., & Timmerman, V. (2016). The hnRNP family: Insights into their role in health and disease. *Human Genetics*, *135*, 851–867.

Giansanti, P., Tsiatsiani, L., Low, T. Y., & Heck, A. J. R. (2016). Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. *Nature Protocols*, *11*, 993–1006.

Gilbert, W. (1978). Why genes in pieces? *Nature*, *271*, 501.

Gonzalez, C., Sims, J. S., Hornstein, N., Mela, A., Garcia, F., Lei, L., Gass, D. A., Amendolara, B., Bruce, J. N., Canoll, P., & Sims, P. A. (2014). Ribosome profiling reveals a cell-type-specific translational landscape in brain tumors. *The Journal of Neuroscience*, *34*, 10924–10936.

Gonzàlez-Porta, M., Frankish, A., Rung, J., Harrow, J., & Brazma, A. (2013). Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene. *Genome Biology*, *14*, R70.

González-Vallinas, J., Pagès, A., Singh, B., & Eyras, E. (2015). A semi-supervised approach uncovers thousands of intragenic enhancers differentially activated in human cells. *BMC Genomics*, *16*, 523.

Grosso, A. R., Gomes, A. Q., Barbosa-Morais, N. L., Caldeira, S., Thorne, N. P., Grech, G., von Lindern, M., & Carmo-Fonseca, M. (2008). Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Research*, *36*, 4823–4832.

Gruber, A. J., & Zavolan, M. (2019). Alternative cleavage and polyadenylation in health and disease. *Nature Reviews. Genetics*, *20*, 599–614.

Haltenhof, T., Kotte, A., de Bortoli, F., Schiefer, S., Meinke, S., Emmerichs, A.-K., Petermann, K. K., Timmermann, B., Imhof, P., Franz, A., Loll, B., Wahl, M. C., Preußner, M., & Heyd, F. (2020). A conserved kinase-based body-temperature sensor globally controls alternative splicing and gene expression. *Molecular Cell*, *78*, 57–69.e4.

Han, H., Braunschweig, U., Gonatopoulos-Pournatzis, T., Weatheritt, R. J., Hirsch, C. L., Ha, K. C. H., Radovani, E., Nabeel-Shah, S., Sterne-Weiler, T., Wang, J., O'Hanlon, D., Pan, Q., Ray, D., Zheng, H., Vizeacoumar, F., Datti, A., Magomedova, L., Cummins, C. L., Hughes, T. R., ... Blencowe, B. J. (2017). Multilayered control of alternative splicing regulatory networks by transcription factors. *Molecular Cell*, *65*, 539–553.e7.

Hao, Y., Colak, R., Teyra, J., Corbi-Verge, C., Ignatchenko, A., Hahne, H., Wilhelm, M., Kuster, B., Braun, P., Kaida, D., Kislinger, T., & Kim, P. M. (2015). Semi-supervised learning predicts approximately one third of the alternative splicing isoforms as functional proteins. *Cell Reports*, *12*, 183–189.

Hegyi, H., Kalmar, L., Horvath, T., & Tompa, P. (2011). Verification of alternative splicing variants based on domain integrity, truncation length and intrinsic protein disorder. *Nucleic Acids Research*, *39*, 1208–1219.

Heyer, E. E., & Moore, M. J. (2016). Redefining the translational status of 80S monosomes. *Cell*, *164*, 757–769.

Hiller, M., Huse, K., Platzer, M., & Backofen, R. (2005). Creation and disruption of protein features by alternative splicing—A novel mechanism to modulate function. *Genome Biology*, *6*, R58. https://pubmed.ncbi.nlm.nih.gov/15998447/

Hsin, J.-P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & Development*, *26*, 2119–2137.

Hu, Y., Fang, L., Chen, X., Zhong, J. F., Li, M., & Wang, K. (2021). LIQA: Long-read isoform quantification and analysis. *Genome Biology*, *22*, 182.

Huang, T., Chen, W., Liu, J., Gu, N., & Zhang, R. (2019). Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nature Structural & Molecular Biology*, *26*, 380–388.

Huang, Y., Li, W., Yao, X., Lin, Q., Yin, J., Liang, Y., Heiner, M., Tian, B., Hui, J., & Wang, G. (2012). Mediator complex regulates alternative mRNA processing via the MED23 subunit. *Molecular Cell*, *45*, 459–469.

Huang, Y., & Steitz, J. A. (2005). SRprises along a Messenger's journey. *Molecular Cell*, *17*, 613–615.

Huang, Y., Triscari, J. M., Tseng, G. C., Pasa-Tolic, L., Lipton, M. S., Smith, R. D., & Wysocki, V. H. (2005). Statistical characterization of the charge state and residue dependence of Low-energy CID peptide dissociation patterns. *Analytical Chemistry*, *77*, 5800–5813.

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science (80- )*, *324*, 218–223.

Ishimura, R., Nagy, G., Dotu, I., Zhou, H., Yang, X.-L., Schimmel, P., Senju, S., Nishimura, Y., Chuang, J. H., & Ackerman, S. L. (2014). RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science*, *345*, 455–459.

Iwata, H., & Gotoh, O. (2011). Comparative analysis of information contents relevant to recognition of introns in many species. *BMC Genomics*, *12*, 45.

Jackson, R. J., Hellen, C. U. T. T., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews. Molecular Cell Biology*, *11*, 113–127.

Janssen, K. A., Coradin, M., Lu, C., Sidoli, S., & Garcia, B. A. (2019). Quantitation of single and combinatorial histone modifications by integrated chromatography of bottom-up peptides and middle-down polypeptide tails. *Journal of the American Society for Mass Spectrometry*, *30*, 2449–2459.

Jeong, S. (2017). SR proteins: Binders, regulators, and connectors of RNA. *Molecules and Cells*, *40*, 1–9.

Jeong, S.-K., Kim, C.-Y., & Paik, Y.-K. (2018). ASV-ID, a Proteogenomic workflow to predict candidate protein isoforms on the basis of transcript evidence. *Journal of Proteome Research*, *17*, 4235–4242.

Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C. Y., Xiao, R., Burge, C. B., & Fu, X.-D. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell*, *153*, 855–868.

Ji, Z. (2018). RibORF: Identifying genome-wide translated open Reading frames using ribosome profiling. *Current Protocols in Molecular Biology*, *124*, e67.

Joazeiro, C. A. P. (2019). Mechanisms and functions of ribosome-associated protein quality control. *Nature Reviews. Molecular Cell Biology*, *20*, 368–383.

Kalsotra, A., & Cooper, T. A. (2011). Functional consequences of developmentally regulated alternative splicing. *Nature Reviews. Genetics*, *12*, 715–729.

Karousis, E. D., Gurzeler, L.-A., Annibaldis, G., Dreos, R., & Mühlemann, O. (2020). Human NMD ensues independently of stable ribosome stalling. *Nature Communications*, *11*, 4134.

Karunratanakul, K., Tang, H.-Y., Speicher, D. W., Chuangsuwanich, E., & Sriswasdi, S. (2019). Uncovering thousands of new peptides with sequence-mask-search hybrid De novo peptide sequencing framework. *Molecular & Cellular Proteomics*, *18*, 2478–2491.

Kervestin, S., & Jacobson, A. (2012). NMD: A multifaceted response to premature translational termination. *Nature Reviews. Molecular Cell Biology*, *13*, 700–712.

Kim, M.-S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabuddhe, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., … Pandey, A. (2014). A draft map of the human proteome. *Nature*, *509*, 575–581.

Kishor, A., Fritz, S. E., Haque, N., Ge, Z., Tunc, I., Yang, W., Zhu, J., & Hogg, J. R. (2020). Activation and inhibition of nonsense-mediated mRNA decay control the abundance of alternative polyadenylation products. *Nucleic Acids Research*, *48*, 7468–7482.

Kochetov, A. V. (2008). Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. *BioEssays*, *30*, 683–691.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., & Zupan, B. (2011). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural & Molecular Biology*, *17*, 909–915.

Kosti, I., Jain, N., Aran, D., Butte, A. J., & Sirota, M. (2016). Cross-tissue analysis of gene and protein expression in normal and cancer tissues. *Scientific Reports*, *6*, 24799.

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., & Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biology*, *20*, 278.

Kozak, M. (2005). Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*, *361*, 13–37.

Lackner, D. H., Beilharz, T. H., Marguerat, S., Mata, J., Watt, S., Schubert, F., Preiss, T., & Bähler, J. (2007). A network of multiple regulatory layers shapes gene expression in fission yeast. *Molecular Cell*, *26*, 145–155.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., … International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, *412*, 565–566.

Laskay, Ü. A., Lobas, A. A., Srzentić, K., Gorshkov, M. V., & Tsybin, Y. O. (2013). Proteome digestion specificity analysis for rational design of extended bottom-up and middle-down proteomics experiments. *Journal of Proteome Research*, *12*, 5558–5569.

Lau, E., Han, Y., Williams, D. R., Thomas, C. T., Shrestha, R., Wu, J. C., & Lam, M. P. Y. (2019). Splice-junction-based mapping of alternative isoforms in the human proteome. *Cell Reports*, *29*, 3751–3765.e5.

Leppek, K., Das, R., & Barna, M. (2018). Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature Reviews. Molecular Cell Biology*, *19*, 158–174. https://pubmed.ncbi.nlm.nih.gov/29165424/

Li, Y. I., Knowles, D. A., Humphrey, J., Barbeira, A. N., Dickinson, S. P., Im, H. K., & Pritchard, J. K. (2018). Annotation-free quantification of RNA splicing using LeafCutter. *Nature Genetics*, *50*, 151–158.

Licatalosi, D. D., & Darnell, R. B. (2010). RNA processing and its regulation: Global insights into biological networks. *Nature Reviews. Genetics*, *11*, 75–87.

Lim, L. P., & Burge, C. B. (2001). A computational analysis of sequence features involved in recognition of short introns. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 11193–11198.

Linder, B., & Jaffrey, S. R. (2019). Discovering and mapping the modified nucleotides that comprise the Epitranscriptome of mRNA. *Cold Spring Harbor Perspectives in Biology*, *11*(6), a032201.

Liu, F., Clark, W., Luo, G., Wang, X., Fu, Y., Wei, J., Wang, X., Hao, Z., Dai, Q., Zheng, G., Ma, H., Han, D., Evans, M., Klungland, A., Pan, T., & He, C. (2016). ALKBH1-mediated tRNA Demethylation regulates translation. *Cell*, *167*, 816–828.e16.

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell*, *165*, 535–550.

Liu, Y., Gonzàlez-Porta, M., Santos, S., Brazma, A., Marioni, J. C., Aebersold, R., Venkitaraman, A. R., & Wickramasinghe, V. O. (2017). Impact of alternative splicing on the human proteome. *Cell Reports*, *20*, 1229–1241.

Long, M., & Deutsch, M. (1999). Intron–Exon structures of eukaryotic model organisms. *Nucleic Acids Research*, *27*, 3219–3228.

Louadi, Z., Yuan, K., Gress, A., Tsoy, O., Kalinina, O. V., Baumbach, J., Kacprowski, T., & List, M. (2021). DIGGER: Exploring the functional role of alternative splicing in protein interactions. *Nucleic Acids Research*, *49*, D309–D318. https://pubmed.ncbi.nlm.nih.gov/32976589/

Lovci, M. T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M., Liang, T. Y., Stark, T. J., Gehman, L. T., Hoon, S., Massirer, K. B., Pratt, G. A., Black, D. L., Gray, J. W., Conboy, J. G., & Yeo, G. W. (2013). Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nature Structural & Molecular Biology*, *20*, 1434–1442.

Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., & Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science (80- )*, *327*, 996–1000.

Ly, T., Ahmad, Y., Shlien, A., Soroka, D., Mills, A., Emanuele, M. J., Stratton, M. R., & Lamond, A. I. (2014). A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. *eLife*, *3*, e01630.

Lynch, K. W. (2004). Consequences of regulated pre-mRNA splicing in the immune system. *Nature Reviews. Immunology*, *4*, 931–940.

Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S. K., Lv, R., Chen, K., Lu, Z., Chen, H., Shi, Y. G., Lan, F., Fan, J., Klaholz, B. P., Pan, T., Shi, Y., & He, C. (2019). N6-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. *Nature Chemical Biology*, *15*, 88–94. https://pubmed.ncbi.nlm.nih.gov/30531910/

Ma, W.-T., Liu, Z.-Y., Chen, X.-Z., Lin, Z.-L., Zheng, Z.-B., Miao, W.-G., & Xie, S.-Q. (2019). A protein identification algorithm for tandem mass spectrometry by incorporating the abundance of mRNA into a binomial probability scoring model. *Journal of Proteomics*, *197*, 53–59.

Mašek, T., Valášek, L., & Pospíšek, M. (2011). Polysome analysis and RNA purification from sucrose gradients. *Methods in Molecular Biology (Clifton, N.J.)*, *703*, 293–309.

Maslon, M. M., Heras, S. R., Bellora, N., Eyras, E., & Cáceres, J. F. (2014). The translational landscape of the splicing factor SRSF1 and its role in mitosis. *Elife*, *3*, e02028.

Maunakea, A. K., Chepelev, I., Cui, K., & Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Research*, *23*, 1256–1269.

Mayr, C. (2016). Evolution and biological roles of alternative 3′UTRs. *Trends in Cell Biology*, *26*, 227–237.

Mazin, P. V., Khaitovich, P., Cardoso-Moreira, M., & Kaessmann, H. (2021). Alternative splicing during mammalian organ development. *Nature Genetics*, *53*, 925–934.

Melamud, E., & Moult, J. (2009). Structural implication of splicing stochastics. *Nucleic Acids Research*, *37*, 4862–4872.

Moqtaderi, Z., Geisberg, J. V., & Struhl, K. (2018). Extensive structural differences of closely related 3′ mRNA isoforms: Links to Pab1 binding and mRNA stability. *Molecular Cell*, *72*, 849–861.e6.

Nakka, K., Ghigna, C., Gabellini, D., & Dilworth, F. J. (2018). Diversification of the muscle proteome through alternative splicing. *Skeletal Muscle*, *8*, 8.

Nickless, A., Bailis, J. M., & You, Z. (2017). Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell & Bioscience*, *7*, 26.

Nusinow, D. P., Szpyt, J., Ghandi, M., Rose, C. M., McDonald, E. R., 3rd, Kalocsay, M., Jané-Valbuena, J., Gelfand, E., Schweppe, D. K., Jedrychowski, M., Golji, J., Porter, D. A., Rejtar, T., Wang, Y. K., Kryukov, G. V., Stegmeier, F., Erickson, B. K., Garraway, L. A., Sellers, W. R., & Gygi, S. P. (2020). Quantitative proteomics of the cancer cell line encyclopedia. *Cell*, *180*, 387–402.e16.

Orr, M. W., Mao, Y., Storz, G., & Qian, S. B. (2020). Alternative ORFs and small ORFs: Shedding light on the dark proteome. *Nucleic Acids Research*, *48*, 1029–1042.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, *40*, 1413–1415.

Papasaikas, P., Tejedor, J. R., Vigevani, L., & Valcárcel, J. (2015). Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Molecular Cell*, *57*, 7–22.

Pertea, M., Shumate, A., Pertea, G., Varabyou, A., Breitwieser, F. P., Chang, Y.-C., Madugundu, A. K., Pandey, A., & Salzberg, S. L. (2018). CHESS: A new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Genome Biology*, *19*, 208.

Pimentel, H., Parra, M., Gee, S. L., Mohandas, N., Pachter, L., & Conboy, J. G. (2016). A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. *Nucleic Acids Research*, *44*, 838–851.

Pisareva, V. P., Skabkin, M. A., Hellen, C. U. T., Pestova, T. V., & Pisarev, A. V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *The EMBO Journal*, *30*, 1804–1817.

Ray, T. A., Cochran, K., Kozlowski, C., Wang, J., Alexander, G., Cady, M. A., Spencer, W. J., Ruzycki, P. A., Clark, B. S., Laeremans, A., He, M. X., Wang, X., Park, E., Hao, Y., Iannaccone, A., Hu, G., Fedrigo, O., Skiba, N. P., Arshavsky, V. Y., & Kay, J. N. (2020). Comprehensive identification of mRNA isoforms reveals the diversity of neural cell-surface molecules with roles in retinal development and disease. *Nature Communications*, *11*, 3328.

Reixachs-Solé, M., Ruiz-Orera, J., Albà, M. M., & Eyras, E. (2020). Ribosome profiling at isoform level reveals evolutionary conserved impacts of differential splicing on the proteome. *Nature Communications*, *11*, 1768.

Ren, F., Zhang, N., Zhang, L., Miller, E., & Pu, J. J. (2020). Alternative polyadenylation: A new frontier in post transcriptional regulation. *Biomarker Research*, *8*, 67.

Rodriguez, J. M., Maietta, P., Ezkurdia, I., Pietrelli, A., Wesselink, J. J., Lopez, G., Valencia, A., & Tress, M. L. (2013). APPRIS: Annotation of principal and alternative splice isoforms. *Nucleic Acids Research*, *41*, D110–D117.

Rodriguez, J. M., Rodriguez-Rivas, J., Di Domenico, T., Vázquez, J., Valencia, A., & Tress, M. L. (2018). APPRIS 2017: Principal isoforms for multiple gene sets. *Nucleic Acids Research*, *46*, D213–D217.

Romero, P. R., Zaidi, S., Fang, Y. Y., Uversky, V. N., Radivojac, P., Oldfield, C. J., Cortese, M. S., Sickmeier, M., LeGall, T., Obradovic, Z., & Dunker, A. K. (2006). Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 8390–8395.

Ruiz-Orera, J., Messeguer, X., Subirana, J. A., & Alba, M. M. (2014). Long non-coding RNAs as a source of new peptides. *eLife*, *3*, e03523.

Saldi, T., Cortazar, M. A., Sheridan, R. M., Bentley, D. L., Genetics, M., & Box, P. O. (2017). Coupling of RNA polymerase II transcription elongation with pre-mRNA splicing. *Journal of Molecular Biology*, *428*, 2623–2635. https://pubmed.ncbi.nlm.nih.gov/27107644/

Schumann, U., Zhang, H.-N., Sibbritt, T., Pan, A., Horvath, A., Gross, S., Clark, S. J., Yang, L., & Preiss, T. (2020). Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biology*, *18*, 40.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*, 337–342. https://pubmed.ncbi.nlm.nih.gov/21593866/

Schwartz, S. H., Silva, J., Burstein, D., Pupko, T., Eyras, E., & Ast, G. (2008). Large-scale comparative analysis of splicing signals and their corresponding splicing factors in eukaryotes. *Genome Research*, *18*, 88–103.

Sebestyén, E., Zawisza, M., & Eyras, E. (2015). Detection of recurrent alternative splicing switches in tumor samples reveals novel signatures of cancer. *Nucleic Acids Research*, *43*, 1345–1356.

Seiler, M., Peng, S., Agrawal, A. A., Palacino, J., Teng, T., Zhu, P., Smith, P. G., Buonamici, S., Yu, L., Caesar-Johnson, S. J., Demchok, J. A., Felau, I., Kasapi, M., Ferguson, M. L., Hutter, C. M., Sofia, H. J., Tarnuzzer, R., Wang, Z., Yang, L., ... Mariamidze, A. (2018). Somatic mutational landscape of splicing factor genes and their functional consequences across 33 cancer types. *Cell Reports*, *23*, 282–296.e4. https://pubmed.ncbi.nlm.nih.gov/29617667/

Shah, S., Molinaro, G., Liu, B., Wang, R., Huber, K. M., & Richter, J. D. (2020). FMRP control of ribosome translocation promotes chromatin modifications and alternative splicing of neuronal genes linked to autism. *Cell Reports*, *30*(13), 4459–4472.e6.

Shen, M., & Mattox, W. (2012). Activation and repression functions of an SR splicing regulator depend on exonic versus intronic-binding position. *Nucleic Acids Research*, *40*, 428–437.

Shen, S., Park, J. W., Lu, Z., Lin, L., Henry, M. D., Wu, Y. N., Zhou, Q., & Xing, Y. (2014). rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National Academy of Sciences*, *111*, E5593–E5601.

Shirokikh, N. E., Archer, S. K., Beilharz, T. H., Powell, D., & Preiss, T. (2017). Translation complex profile sequencing to study the in vivo dynamics of mRNA-ribosome interactions during translation initiation, elongation and termination. *Nature Protocols*, *12*, 697–731.

Shkreta, L., & Chabot, B. (2015). The RNA splicing response to DNA damage. *Biomolecules*, *5*, 2935–2977.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., & Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, *479*, 74–79.

Singer, P. A., Singer, H. H., & Williamson, A. R. (1980). Different species of messenger RNA encode receptor and secretory IgM μ chains differing at their carboxy termini. *Nature*, *285*, 294–300.

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., Suter, C. M., & Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Research*, *40*, 5023–5033.

Steitz, J. A. (1969). Polypeptide chain initiation: Nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature*, *224*, 957–964.

Sterne-Weiler, T., Martinez-Nunez, R. T., Howard, J. M., Cvitovik, I., Katzman, S., Tariq, M. A., Pourmand, N., & Sanford, J. R. (2013). Frac-seq reveals isoform-specific recruitment to polyribosomes. *Genome Research*, *23*, 1615–1623.

Sulakhe, D., D'Souza, M., Wang, S., Balasubramanian, S., Athri, P., Xie, B., Canzar, S., Agam, G., Gilliam, T. C., & Maltsev, N. (2019). Exploring the functional impact of alternative splicing on human protein isoforms using available annotation sources. *Briefings in Bioinformatics*, *20*, 1754–1768.

Sun, Y. H., Wang, A., Song, C., Shankar, G., Srivastava, R. K., Au, K. F., & Li, X. Z. (2021). Single-molecule long-read sequencing reveals a conserved intact long RNA profile in sperm. *Nature Communications*, *12*, 1361.

Tamarkin-Ben-Harush, A., Vasseur, J.-J., Debart, F., Ulitsky, I., & Dikstein, R. (2017). Cap-proximal nucleotides via differential eIF4E binding and alternative promoter usage mediate translational response to energy stress. *eLife*, *6*, e21907.

Tang, A. D., Soulette, C. M., van Baren, M. J., Hart, K., Hrabeta-Robinson, E., Wu, C. J., & Brooks, A. N. (2020). Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nature Communications*, *11*, 1438.

Tapial, J., Ha, K. C. H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallières, M., Permanyer, J., Sodaei, R., Marquez, Y., Cozzuto, L., Wang, X., Gómez-Velázquez, M., Rayon, T., Manzanares, M., Ponomarenko, J., Blencowe, B. J., & Irimia, M. (2017). An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Research*, *27*, 1759–1768.

Tasic, B., Nabholz, C. E., Baldwin, K. K., Kim, Y., Rueckert, E. H., Ribich, S. A., Cramer, P., Wu, Q., Axel, R., & Maniatis, T. (2002). Promoter choice determines splice site selection in Protocadherin α and γ pre-mRNA splicing. *Molecular Cell*, *10*, 21–33.

Tay, A. P., Pang, C. N. I., Twine, N. A., Hart-Smith, G., Harkness, L., Kassem, M., & Wilkins, M. R. (2015). Proteomic validation of transcript isoforms, including those assembled from RNA-Seq data. *Journal of Proteome Research*, *14*, 3541–3554.

Thanaraj, T. A., & Clark, F. (2001). Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. *Nucleic Acids Research*, *29*, 2581–2593.

Tian, B., & Manley, J. L. (2016). Alternative polyadenylation of mRNA precursors. *Nature Reviews. Molecular Cell Biology*, *18*, 18–30. https://pubmed.ncbi.nlm.nih.gov/27677860/

Tran, T. T., Bollineni, R. C., Strozynski, M., Koehler, C. J., & Thiede, B. (2017). Identification of alternative splice variants using unique Tryptic peptide sequences for database searches. *Journal of Proteome Research*, *16*, 2571–2578.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., & Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, *28*, 511–515.

Tress, M. L., Abascal, F., & Valencia, A. (2017). Alternative splicing may not be the key to proteome complexity. *Trends in Biochemical Sciences*, *42*, 98–110.

Trincado, J. L., Reixachs-Solé, M., Pérez-Granado, J., Fugmann, T., Sanz, F., Yokota, J., & Eyras, E. (2021). ISOTOPE: ISOform-guided prediction of epiTOPEs in cancer. *PLoS Computational Biology*, *17*, e1009411.

Trincado, J. L. J. L., Entizne, J. C. J. C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D. J. D. J., & Eyras, E. (2018). SUPPA2: Fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biology*, *19*, 40.

Troskie, R.-L., Jafrani, Y., Mercer, T. R., Ewing, A. D., Faulkner, G. J., & Cheetham, S. W. (2021). Long-read cDNA sequencing identifies functional pseudogenes in the human transcriptome. *Genome Biology*, *22*, 146.

Tung, K.-F., Pan, C.-Y., Chen, C.-H., & Lin, W. (2020). Top-ranked expressed gene transcripts of human protein-coding genes investigated with GTEx dataset. *Scientific Reports*, *10*, 16245.

Ule, J., & Blencowe, B. J. (2019). Alternative splicing regulatory networks: Functions, mechanisms, and evolution. *Molecular Cell*, *76*, 329–345.

Ule, J., Jensen, K. B., Ruggiu, M., Mele, A., Ule, A., & Darnell, R. B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science (80-)*, *302*, 1212–1215.

Van Nostrand, E. L., Pratt, G. A., Yee, B. A., Wheeler, E. C., Blue, S. M., Mueller, J., Park, S. S., Garcia, K. E., Gelboin-Burkhart, C., Nguyen, T. B., Rabano, I., Stanton, R., Sundararaman, B., Wang, R., Fu, X. D., Graveley, B. R., & Yeo, G. W. (2020). Principles of RNA processing from analysis of enhanced CLIP maps for 150 RNA binding proteins. *Genome Biology*, *21*, 90.

Vaquero-Garcia, J., Barrera, A., Gazzara, M. R., González-Vallinas, J., Lahens, N. F., Hogenesch, J. B., Lynch, K. W., & Barash, Y. (2016). A new view of transcriptome complexity and regulation through the lens of local splicing variations. *eLife*, *5*, e11752.

Venables, J. P., Lapasset, L., Gadea, G., Fort, P., Klinck, R., Irimia, M., Vignal, E., Thibault, P., Prinos, P., Chabot, B., Abou Elela, S., Roux, P., Lemaitre, J. M., & Tazi, J. (2013). MBNL1 and RBFOX2 cooperate to establish a splicing programme involved in pluripotent stem cell differentiation. *Nature Communications*, *4*, 1–10.

Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews. Genetics*, *13*, 227–232.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., & Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature*, *456*, 470–476.

Wang, H., McManus, J., & Kingsford, C. (2016). Isoform-level ribosome occupancy estimation guided by transcript abundance with Ribomap. *Bioinformatics*, *32*, 1880–1882.

Wang, T., Cui, Y., Jin, J., Guo, J., Wang, G., Yin, X., He, Q.-Y., & Zhang, G. (2013). Translating mRNAs strongly correlate to proteins in a multivariate manner and their translation ratios are phenotype specific. *Nucleic Acids Research*, *41*, 4743–4754.

Wang, X., Codreanu, S. G., Wen, B., Li, K., Chambers, M. C., Liebler, D. C., & Zhang, B. (2018). Detection of proteome diversity resulted from alternative splicing is limited by trypsin cleavage specificity. *Molecular & Cellular Proteomics*, *17*, 422–430.

Wang, X., You, X., Langer, J. D., Hou, J., Rupprecht, F., Vlatkovic, I., Quedenau, C., Tushev, G., Epstein, I., Schaefke, B., Sun, W., Fang, L., Li, G., Hu, Y., Schuman, E. M., & Chen, W. (2019). Full-length transcriptome reconstruction reveals a large diversity of RNA and protein isoforms in rat hippocampus. *Nature Communications*, *10*, 5009.

Wang, Z., & Burge, C. B. (2008). Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA*, *14*, 802–813.

Warner, J. R., Knopf, P. M., & Rich, A. (1963). A multiple ribosomal structure in protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *49*, 122–129.

Weatheritt, R. J., Davey, N. E., & Gibson, T. J. (2012). Linear motifs confer functional diversity onto splice variants. *Nucleic Acids Research*, *40*, 7123–7131.

Weatheritt, R. J., Sterne-Weiler, T., & Blencowe, B. J. (2016). The ribosome-engaged landscape of alternative splicing. *Nature Structural & Molecular Biology*, *23*, 1117–1123.

Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J. H., Bantscheff, M., ... Kuster, B. (2014). Mass-spectrometry-based draft of the human proteome. *Nature*, *509*, 582–587.

Wojtowicz, W. M., Wu, W., Andre, I., Qian, B., Baker, D., & Zipursky, S. L. (2007). A vast repertoire of Dscam binding specificities arises from modular interactions of variable Ig domains. *Cell*, *130*, 1134–1145.

Wu, P., Pu, L., Deng, B., Li, Y., Chen, Z., & Liu, W. (2019). PASS: A proteomics alternative splicing screening pipeline. *Proteomics*, *19*, e1900041.

Yang, X., Coulombe-Huntington, J., Kang, S., Sheynkman, G. M. M., Hao, T., Richardson, A., Sun, S., Yang, F., Shen, Y. A. A., Murray, R. R. R., Spirohn, K., Begg, B. E., Duran-Frigola, M., MacWilliams, A., Pevzner, S. J., Zhong, Q., Trigg, S. A., Tam, S., Ghamsari, L., ... Vidal, M. (2016). Widespread expansion of protein interaction capabilities by alternative splicing. *Cell*, *164*, 805–817.

Yearim, A., Gelfman, S., Shayevitch, R., Melcer, S., Glaich, O., Mallm, J. P., Nissim-Rafinia, M., Cohen, A. H. S., Rippe, K., Meshorer, E., & Ast, G. (2015). HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Reports*, *10*, 1122–1134.

Yee, B. A., Pratt, G. A., Graveley, B. R., Van Nostrand, E. L., & Yeo, G. W. (2019). RBP-maps enables robust generation of splicing regulatory maps. *RNA*, *25*, 193–204.

Yoshikawa, H., Sundaramoorthy, R., Mariyappa, D., Jiang, H., & Lamond, A. I. (2021). Efficient and rapid analysis of Polysomes and ribosomal subunits in cells and tissues using Ribo mega-SEC. *Bio-Protocol*, *11*, e4106.

Zhan, X., Yan, C., Zhang, X., Lei, J., & Shi, Y. (2018). Structures of the human pre-catalytic spliceosome and its precursor spliceosome. *Cell Research*, *28*, 1129–1140.

Zhang, H., Wang, Y., Wu, X., Tang, X., Wu, C., & Lu, J. (2021). Determinants of genome-wide distribution and evolution of uORFs in eukaryotes. *Nature Communications*, *12*, 1076. https://pubmed.ncbi.nlm.nih.gov/33597535/

Zhang, S., Aibara, S., Vos, S. M., Agafonov, D. E., Lührmann, R., & Cramer, P. (2021). Structure of a transcribing RNA polymerase II-U1 snRNP complex. *Science (80-)*, *371*, 305–309.

Zhao, J., Qin, B., Nikolay, R., Spahn, C. M. T., & Zhang, G. (2019). Translatomics: The global view of translation. *International Journal of Molecular Sciences*, *20*, 212.