The physiology of alternative splicing

Luciano E. Marasco[1,2] *and Alberto R. Kornblihtt*[1] ✉

Abstract | Alternative splicing is a substantial contributor to the high complexity of transcriptomes of multicellular eukaryotes. In this Review, we discuss the accumulated evidence that most of this complexity is reflected at the protein level and fundamentally shapes the physiology and pathology of organisms. This notion is supported not only by genome-wide analyses but, mainly, by detailed studies showing that global and gene-specific modulations of alternative splicing regulate highly diverse processes such as tissue-specific and species-specific cell differentiation, thermal regulation, neuron self-avoidance, infrared sensing, the Warburg effect, maintenance of telomere length, cancer and autism spectrum disorders (ASD). We also discuss how mastering the control of alternative splicing paved the way to clinically approved therapies for hereditary diseases.

The first evidence of alternative splicing coincided with the discovery of splicing in adenovirus transcripts in 1977 (REFS.[1,2]). Due to the polycistronic nature of bacterial genes, molecular biologists before 1977 were used to the idea that a single gene could encode more than one polypeptide. However, the finding that eukaryotic genes were organized as a mosaic of exons, which are represented in the mature mRNA, and of introns, which are removed from the transcribed pre-mRNA with the concomitant covalent joining of exons, was as disruptive as fascinating. In his seminal commentary 'Why Genes in Pieces?'[3], Walter Gilbert immediately realized that splicing should give rise to differential splice site selection or alternative splicing, resulting in expansion of the coding capacity of eukaryotic genes. Whereas in bacteria each cistron may encode a polypeptide that is completely different in sequence and function from those of the other cistrons of the same gene, in eukaryotes the different polypeptides produced from a single gene through alternative splicing are usually similar but not identical, and include conserved and divergent regions giving rise to subtle or radical variations both at the mRNA and protein levels (FIG. 1a).

It should be noted that splicing and alternative splicing create primarily RNA variants (also referred to as isoforms) that may or may not give rise to protein variation. First, splicing is not restricted to genes encoding mRNAs, but also occurs in non-coding RNA[4–6]. Second, in the case of mRNAs, the first AUG — the first codon of a translated open reading frame — does not coincide with the beginning of the first exon and can be embedded farther downstream within the first or even the second exon, which also applies to non-AUG start codons that were shown to be more frequent than expected[7]. These facts constitute a big caveat in the oversimplified statement that exons encode protein and introns do not. Exons of genes that do not encode protein are as bona fide exons as those of genes that do encode protein. But most importantly, at mRNA 5′ and 3′ ends, exons harbour untranslated regions that may never encode an amino acid stretch. Therefore, exons more accurately should be considered portions of a gene that are represented in the mature RNA, whereas introns are portions that are absent from the mature RNA, as stated by the National Human Genome Research Institute, National Institutes of Health (NIH). Of course, due to the different modes of alternative splicing, in particular in the selection of alternative 5′ or 3′ splice sites, a portion of an intron may become part of an exon and vice versa.

Since the publication of the classical review by Smith and Valcárcel[8], many comprehensive reviews have been published on the mechanisms of splicing and the regulation and roles of alternative splicing[9–17]. In this Review, we aim not to revisit these subjects but to discuss the evidence supporting, and controversies about, the adaptive value and the physio-pathological roles of alternative pre-mRNA splicing, focusing on animals. During the first decades after the discovery of alternative splicing, many specific alternative splicing events were studied individually in depth, both functionally and mechanistically[18]. However, it was only following the revolution of next-generation sequencing and genome-wide methodologies that biologists realized that alternative splicing was more a rule than an exception in multicellular organisms, taking place in about 95% of human genes[19,20]. This compelling evidence posed the question of how much of alternative splicing, if not all, is just noise, or in other words, how much of alternative splicing contributes to proteomic diversity, cell fate and differentiation, organism homeostasis and disease. We discuss this question extensively in this Review. Following an overview of the basic mechanisms of alternative splicing, we focus on different aspects

[1]*Universidad de Buenos Aires (UBA), Facultad de Ciencias Exactas y Naturales, Departamento de Fisiología, Biología Molecular Celular and CONICET-UBA, Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE), Buenos Aires, Argentina.*

[2]*Present address: Sir William Dunn School of Pathology, University of Oxford, Oxford, UK.*

✉*e-mail: ark@fbmc.fcen.uba.ar*

https://doi.org/10.1038/s41580-022-00545-z

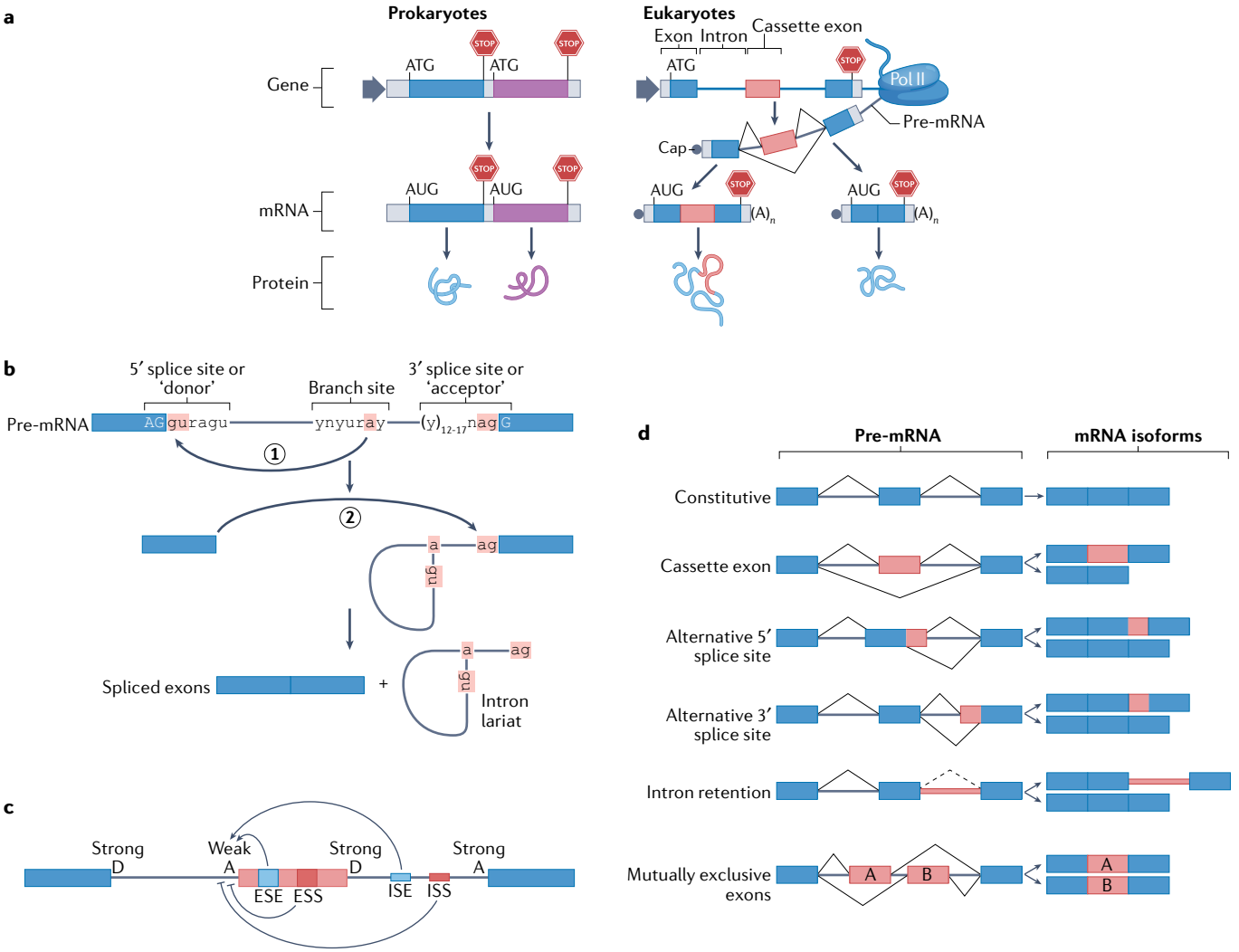


Fig. 1 | **Principles of alternative splicing. a** | Two different strategies to increase the coding capacity of genes. In prokaryotes (left), most genes encode a single mRNA molecule that includes more than one translation open reading frame (cistron). The different cistrons encode proteins with different amino acid sequences and functions, sometimes related to the same metabolic or regulatory pathway. In eukaryotes (right), most RNA polymerase II (Pol II)-transcribed genes are monocistronic, but most are able to encode more than one mRNA through alternative splicing and other mechanisms. The panel illustrates alternative splicing of a cassette exon, which gives rise to two similar but not identical mRNAs and proteins. **b** | Two-step splicing reaction, showing a pre-mRNA segment labelled with the consensus sequences of the splicing 'donor' (5′ splice site), 'acceptor' (3′ splice site) and branch sites. The first step (**1**) is a transesterification reaction, in which the 2′OH of the ribose of the adenine at the branch site performs a nucleophilic attack on the nucleotide at the 5′ end of the intron, forming a lariat intermediate through a 2′–5′ intra-intronic covalent bond. The second step (**2**) is another transesterification reaction, in which the 3′OH of the ribose of the released 5′ exon performs a nucleophilic attack on the nucleotide at the 3′ end of the intron with the subsequent covalent joining of the two exons and the release of the intron lariat. **c** | A pre-mRNA segment depicting *cis*-acting elements that control inclusion of a cassette exon with a weak or suboptimal 3′ splice site (weak acceptor, A). Promotion of the use of the weak acceptor will lead to inclusion of the alternative exon. Repression of the use of the acceptor will lead to inhibition of inclusion of the alternative exon. **d** | Constitutive splicing (top) and five different modes of alternative splicing. Constitutive splicing generally takes place when all splice sites are optimal. However, it may also occur in exons with suboptimal splice sites but with strong splicing enhancers[118]. $(A)_n$, polyadenylation tail; ATG, AUG, translation initiation codon; D, donor; ESE, exonic splicing enhancer; ESS, exonic spicing silencer; ISE, intronic splicing enhancer; ISS, intronic splicing silencer; STOP, translation termination codon.

of alternative splicing functionality and involvement in physio-pathological processes, and end with discussing novel successful splicing-correcting therapies.

Principles of alternative splicing

Splicing of eukaryotic RNA polymerase II (Pol II)-transcribed pre-mRNAs is carried out by the spliceosome, a megadalton complex that assembles along each pre-mRNA intron from small nuclear ribonucleoproteins and a set of auxiliary proteins[9]. Spliceosome components recognize and bind to consensus sequences located at the 5′ and 3′ ends — the 5′ splice site and the 3′ splice site — of each intron (FIG. 1b) and catalyse two consecutive transesterification reactions that end up removing the intron and covalently joining the adjacent exons. 'Strong' splice sites (that is, those more similar to the consensus sequence) are more efficiently recognized by spliceosome components and, therefore, typically lead

to constitutive splicing. When a splice site is 'weak' or suboptimal, alternative splicing often occurs, meaning that the weak site is partially used in intron removal and exon ligation reactions, and different proportions of mature mRNA variants arise. The use of splice sites is not only determined by their intrinsic sequences (optimal or suboptimal) acting in *cis* but also by proteins acting in *trans*, known as splicing factors, that recognize and bind to specific target sequences located in exons or introns. These regulatory sequences comprise exonic splicing enhancers and silencers (ESEs and ESSs) and intronic splicing enhancers and silencers (ISEs and ISSs) (FIG. 1c). In the case of cassette exons, binding of splicing factors such as serine/arginine-rich (SR) proteins to splicing enhancers can promote use of suboptimal splice sites and, consequently, exon inclusion, whereas binding of members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family to splicing silencers can promote exon skipping, although hnRNPs can also have the opposite effect depending on the location of their binding sites relative to the regulated splice sites, which greatly widens their regulatory potential. This kind of regulation applies to different modes of alternative splicing (FIG. 1d).

Alternative splicing is not only regulated by interactions of *trans*-acting splicing factors with *cis*-regulatory sequences present in the pre-mRNA but is also coupled to transcription by Pol II. Indeed, splicing factors are recruited to the pre-mRNA co-transcriptionally, which determines either the occurrence of co-transcriptional catalysis or the commitment to a post-transcriptional catalytic reaction. Factor recruitment to the nascent pre-mRNA can be modulated by phosphorylation of different residues of the carboxy-terminal domain of Pol II, by different histone post-translational modifications and by the elongation rate of Pol II along gene bodies, which may also control the secondary structure of the nascent pre-mRNA[21]. The Pol II speed depends on the degree of chromatin compaction, which in turn depends on specific histone post-translational modifications. A complex regulatory mechanism emerges, in which chromatin[22] and Pol II kinetics are superimposed on interactions of splicing factors with their target sequences in the pre-mRNA. According to the kinetic coupling model, slow elongation can promote either exon inclusion or skipping, depending on the type of exon and its associated sequence features[23–26]. In class I exons, slow elongation promotes inclusion by improving the recruitment of inclusion-enhancing splicing factors to the nascent pre-mRNA, whereas in class II exons slow elongation enhances the binding of inclusion-suppressing factors to their target splicing sites, resulting in exon skipping. Consequently, elongation-permissive histone modifications such as acetylated histone H3 Lys9 (H3K9Ac) at gene bodies promotes skipping of class I exons[27], whereas elongation-repressive histone modifications such as dimethylated histone H3 Lys9 (H3K9me2) promotes their inclusion[28,29]. Conversely, intragenic H3K9Ac promotes inclusion of class II exons, and treatments that inhibit elongation promote their skipping[24]. The links between chromatin conformation and alternative splicing have been recently strengthened by the demonstration that targeting of catalytically dead Cas9 fused to histone modifying enzymes induces predictable changes in alternative splicing[30].

Separating functional alternative splicing from noise

Noise has been defined as the stochastic effects in biochemical processes such as transcription and translation that contribute to phenotypic differences between cells[31]. A key question is how much of the alternative splicing-generated variation observed at the mRNA level is also represented at the protein level.

Proteomic analyses. Although genome-wide methodologies such as RNA sequencing (RNA-seq) and microarrays have clearly demonstrated the pervasiveness of alternative splicing at the transcriptomic level in cells from multicellular eukaryotes, alternative splicing was suggested to lack a significant role in generating proteomic complexity[32]. Evaluation of published data for liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) analyses revealed a surprisingly small number of alternatively spliced proteins, which led the authors of this report to claim that most protein-coding genes seem to produce a single main protein[32]. The authors provocatively concluded that most alternative exons are not under selective pressure and that most annotated mRNA alternative splicing variants are unlikely to have a function as proteins. A few years later, the same researchers seemed to have contradicted their previous statements when analysing tissue-specific alternative splicing at the protein level[33]. They manually curated a set of 255 alternative splicing events detected by proteomics and found that more than a third of them had significant tissue-specific differences. These alternative splicing events were especially abundant in the nervous system and in muscles, in genes encoding cytoskeleton or cell adhesion proteins. The authors concluded that conserved tissue-specific alternative splicing "may have" had a role in the development of the vertebrate brain and heart[33]. More recently, and also based on large proteomic data, the same researchers developed a bioinformatics tool able to predict the functional importance of splice isoforms, and reverted to asserting that 85% of alternative exons are unlikely to be relevant[34]. In any case, it seems advisable to be cautious when investigating the adaptive value of alternative splicing simply by analysing the correspondence between alternative splicing isoform abundance of mRNAs and their encoded proteins at the genome-wide level.

Absence of evidence is not evidence of absence. The above aphorism is the final line of a letter[35] that provides evidence supporting the contrary conclusion to that provided in REF.[32]. In particular, a major limitation of applying LC-MS/MS for the detection of protein splicing variants is its poor coverage and sensitivity, and an additional, compounding issue is that the models used[32] to compare levels of expected and observed protein variants did not take into account additional key technical limitations that contribute to the level of false negatives. One of these key technical issues is that, in complex samples, the

Transesterification
The exchange of the organic group R″ of an ester with the organic group R′ of an alcohol. In the second step of splicing, the uncleaved intron 3′ end is the (phosphodi)ester and the free exon is the alcohol. Following the reaction, the free intron becomes an alcohol and the joined exons form a phosphodiester bond.

Cassette exons
Exons that as a whole are either included in or skipped from the mature mRNA.

Class I exons
Exons whose inclusion in the mature mRNA is promoted by slow, and inhibited by fast, transcript elongation.

Class II exons
Exons whose inclusion in the mature mRNA is inhibited by slow, and promoted by fast, transcript elongation.

Nonsense-mediated decay (NMD). An mRNA degradation mechanism that is coupled to splicing and translation and is triggered by premature stop codons.

number of peptides exceeds the number of sequencing cycles provided by the mass spectrometer, which results in under-sampling and, subsequently, lack of sensitivity to peptides originating from alternative splicing protein isoforms that are variably expressed in different cells. Importantly, there is also evidence indicating that alternatively spliced mRNA isoforms are indeed translated. Using ribosome profiling, a sensitive method for the detection of ribosome-engaged mRNAs[36], about 75% of human mRNAs of medium to high abundance with skipped exons were found to be engaged by ribosomes[37]. It was also observed that low-abundance mRNAs that include alternatively spliced cassette exons and that do not appear in the ribosome profiling data also displayed intron retention, a mode of alternative splicing that is known to promote nuclear retention of the transcript or, if the mRNA is exported, degradation by nonsense-mediated decay (NMD)[38]. These findings are consistent with the detection of thousands of alternative splicing mRNA variants in polysome fractions[39,40].

We note, however, that alternative splicing mRNA isoforms related to cell fate decisions and differentiation are highly conserved and exist in parallel to high levels of species-specific alternative splicing. In fact, most mammalian alternative splicing events are species-specific, are subjected to relaxed selection pressure and correspond to mRNAs of low abundance in certain tissues but of high abundance in other tissues[41], which does not mean that such mRNAs are not translated or lack function. To exclude alternative splicing as a generator of protein diversity would require conducting careful mass spectrometry experiments in a panoply of different cell types to determine whether the major isoforms detected for each gene in each cell type differ from one another, even under low sensitivity. In conclusion, we think that a blind and non-contextualized interpretation of mass spectrometry data could lead to neglecting the importance of alternative splicing and other alternative mRNA maturation processes in shaping protein expression, as demonstrated by hundreds of studies of individual genes over the past four decades, ranging from sex determination in flies[42] to infrared sensing in bats[43]; from tissue-specific expression of two different proteins encoded by the calcitonin-related polypeptide-α (*CALCA*) gene[44] to switches from membrane-associated to secreted immunoglobulin[45]; and from neuron self-avoidance[46] to the generation of multiple functional mRNA variants encoding important proteins from a single HIV genome[47]. Obviously, none of these selected examples would exist if the splicing variants were not translated into protein, which in fact was proved for each of them.

Relevant global changes in alternative splicing patterns

A global analysis in steady-state conditions revealed that protein levels were largely determined by mRNA concentrations[48]. Therefore, it is not surprising that, as happens with transcription, specific patterns in alternative splicing become characteristic of specific cell and tissue phenotypes. In this section we discuss studies that linked global alternative splicing changes and cell proliferation, cancer, autism spectrum disorders (ASD) and thermal regulation. We also discuss alternative splicing patterns that are specific of tissues, organs and species.

Cell proliferation and cancer. The role of alternative splicing in the control of cell proliferation and cancer has been studied for decades. In fact, the term 'cancer spliceome' was added to other 'omes' such as transcriptome, proteome and interactome, to refer to alternative splicing patterns characteristic of cancer cells[49]. In the context of non-cancer cell proliferation, global alternative splicing patterns are altered during liver regeneration, which is mainly based on hepatocyte cell division. Liver regeneration-inducing toxins cause downregulation of epithelial splicing regulatory protein 2 (ESRP2), which in turn upregulates the production of a set of 'neonatal' splicing isoforms that attenuate the Hippo signalling pathway that regulates cell proliferation, differentiation, organ size and tissue homeostasis[50]. Similarly, overexpression in transgenic mice of serine/arginine-rich splicing factor 6 (SRSF6), which is frequently overexpressed in human skin cancer, causes changes in alternative splicing of hundreds of genes and induces hyperplasia of sensitized skin[51].

With respect to cancer, it is not our purpose to cover the vast literature on the link between alternative splicing and cancer, which has been already thoroughly reviewed[52–54], but to comment on some selected examples that involve global changes in alternative splicing patterns. One of the first and best characterized cases is that of the splicing factor SRSF1 (formerly known as ASF1 or SF2), whose experimental overexpression causes cell transformation in culture and which is overexpressed in many human tumour types[52,55,56]. Due to these observations, the *SRSF1* gene qualifies as a proto-oncogene with the particularity that, being a ubiquitously expressed splicing factor with a vast number of different mRNA targets, its overexpression causes massive pathogenic changes in alternative splicing patterns. Similar effects have been observed following overexpression of other SR proteins and of hnRNPs[52].

Mutations in genes encoding splicing factors are also an important source of global changes in alternative splicing associated with cancer. The most relevant examples are those related to myeloid malignancies, which include myelodysplastic syndromes and acute myeloid leukaemia. The mutation of Pro95 in SRSF2 (formerly known as SC35), which has high prevalence in chronic myelomonocytic leukaemia, changes its mRNA target site specificity and, therefore, alters alternative splicing of many transcripts, including those of the histone methyltransferase EZH2 and of the apoptosis repressor BCOR, proteins that have been implicated in cancer[57]. Although it is predictable that mutations in genes encoding regulatory splicing factors (splicing regulators) contribute to cancer, it is certainly more surprising to find mutations in genes encoding core spliceosome components, that is, proteins that are necessary for splicing of all genes in all cell types. The latter category includes mutations in U2 small nuclear RNA auxiliary factor 1 (U2AF1; also known as U2AF35), the small subunit of U2AF, which is

Detained introns
Introns that cause the mRNAs that harbour them to remain in the nucleus.

involved in recognition of the 3′ splice site, and splicing factor 3B subunit 1 (SF3B1), a subunit of the U2 small nuclear ribonucleoprotein. Even more surprising is that individuals with myelodysplastic syndrome carrying mutations in SRSF2, U2AF1 or SF3B1 showed similar alterations in alternative splicing patterns[54], revealing targets common to the three factors despite SRSF2 being a regulatory splicing factor and U2AF1 and SF3B1 being constitutive splicing factors. Why mutations in constitutive splicing factors do not affect splicing in general but, preferentially, cause changes in certain alternative splicing events is still puzzling and could be partially explained by redundancy with other factors in securing efficient constitutive slicing. Another explanation is based on the report of regulatory functions of core spliceosome components, which are related to the order and duration of their recruitment to the spliceosome[58]. The authors found that the timing of the recruitment of core spliceosome factors is important for their regulatory roles and that 'late'-recruited factors are predominant among spliceosome factors with regulatory function. Thus, owing to the sequential recruitment of splicing factors, limiting amounts of late-spliceosome components could favour alternative splicing at splice sites that already harbour early complexes that can more efficiently recruit late factors[58]. Similar evidence on regulatory roles of core spliceosome components has been reported not only in mammalian cells[59] but also in yeast[60].

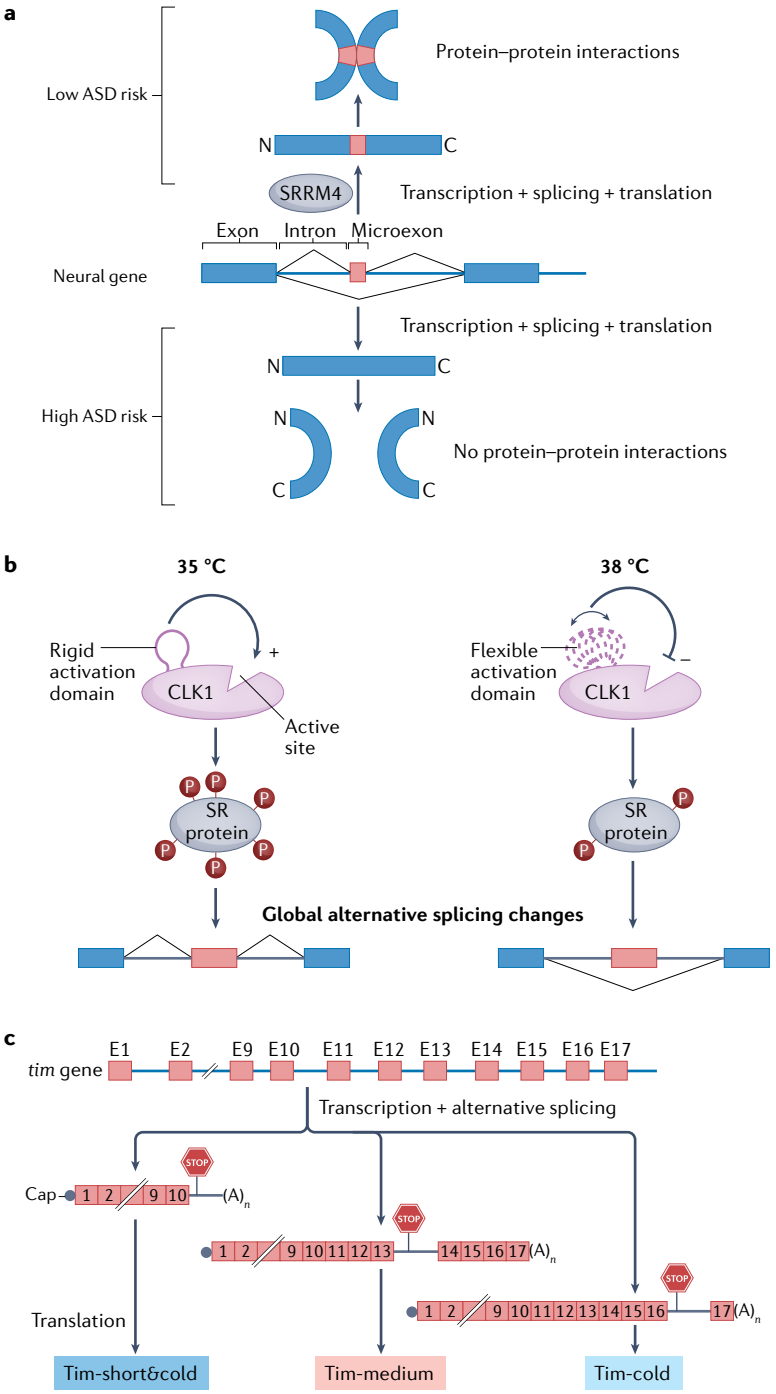
In addition to overexpression and mutation, splicing factor post-translational modifications may also contribute to malignancy. The catalytic activity of protein arginine *N*-methyltransferase 5 (PRMT5) is essential for the survival of acute myeloid leukaemia cells, because the function of SRSF1 depends on its methylation by PRMT5, and SRSF1 expression determines a pattern of alternative splicing of multiple genes that promote leukaemia cell survival, positioning PRMT5 inhibitors as potential therapeutic drugs[61]. In agreement with PRMT5 having a key role in the control of alternative splicing in cancer, PRMT5 depletion or inhibition suppresses glioblastoma tumours in vivo. PRMT5 deficiency prevents the removal of detained introns from pre-mRNAs of proliferation genes, causing cell cycle defects, cell senescence and/or apoptosis[62]. Finally, when assessing the roles in cancer of splicing factors such as SRSF1, it should be noted that SRSF1 shuttles between the nucleus and the cytoplasm and whereas in the nucleus it controls splicing, in the cytoplasm it regulates translation[63]. This translation function is so important that preventing SRSF1 export from the nucleus causes severe developmental defects[64].

Splicing factors that do not belong to the SR and hnRNP families and whose expression is cell type-specific were also implicated in cancer through global alternative splicing changes. The epithelial-specific splicing regulator ESRP2 is highly expressed in primary prostate cancer, whose progression is driven by androgens. These hormones cause the overexpression of ESPR2 and, thus, induce many alternative splicing events that drive an epithelial splicing programme characteristic of prostate cancer cells that correlates with cancer progression[65]. Androgen deprivation therapy represses ESRP2 expression and dampens the epithelial splicing programme, which is replaced by a normal, mesenchymal splicing programme[65].

Autism spectrum disorders. Microexons are 3–27 nucleotide-long exons whose lengths are usually multiples of three nucleotides, indicating that their inclusion in the mature mRNA does not disrupt the translation open reading frame. Through transcriptomic analyses of several species and organs, microexons were found to have the highest evolutionary conservation of sequence and of inclusion levels relative to other classes of alternative splicing elements[66]. In the nervous system, microexon inclusion is promoted by the splicing factor serine/arginine repetitive matrix protein 4 (SRRM4). Microexons of neural mRNAs are frequently excluded from their mRNAs in brains of individuals with ASD, which mirrors the reduced levels of SRRM4 found in these individuals. The peptide segments encoded by the neural microexons were consistently found to participate in protein–protein interactions of proteins involved in neurogenesis and synaptic functions, which suggests there is a cause–effect relationship between microexon misregulation and ASD[66] (FIG. 2a). Indeed, tissue-specific alternative exons of any length often encode structurally disordered protein regions that mediate protein–protein interactions[67,68].

More recently, a study of two microexons in eukaryotic translation initiation factor 4γ1 (eiF4G1) and in eiF4G3, whose inclusion is disrupted in autism, showed that their deletion from the corresponding genes specifically upregulates synaptic proteins that control neuron activity and plasticity, thereby mimicking the phenotype of activated neurons[69]. Mice lacking the eiF4G1 microexon, in particular, display behavioural and memory deficits. In another study, skipping of the neuronal-specific microexon 4 in the mRNA of cytoplasmic polyadenylation element binding protein 4 (CPEB4) is predominant in individuals with ASD. CPEB proteins bind to specific mRNAs and control the length of their poly(A) tails through deadenylation. In individuals with ASD, CPEB4 binds to mRNAs of most high-confidence ASD genes. The splicing isoform of CPEB4 lacking microexon 4 inhibits translation of the ASD mRNAs, thereby mimicking the effects of disruptive mutations in these ASD genes[70]. These eiF4G and CPEB4 examples illustrate how alternative splicing of microexons regulates global mRNA translation in neurons and controls cognitive functions.

Although mutations in various genes have been reported to increase the risk of autism, discussions of the genetic versus environmental aetiology of the condition are still alive. A surprising finding seems to reinforce the environmental aetiology hypothesis. The gut microbiome differs between individuals with ASD and typically developing individuals. Transplantation of gut microbiota from typically developing individuals or individuals with ASD to germ-free mice revealed that colonization with the ASD microbiota induced autistic behaviours in the mice[71]. Importantly, RNA-seq analysis of the mouse brains showed few changes in gene expression levels, but alteration of more than 500 alternative splicing events, 52 of which were previously characterized as strictly related to autism. How can bacteria that reside in the gut affect brain function? The prevailing



Protein/behaviour	18 °C	29 °C
Tim-short&cold	High	Low
Tim-cold	High	Low
Tim-medium	Low	High
Locomotor activity	High day/night ratio	Low day/night ratio
Midday siesta	No	Yes

Fig. 2 | **Examples of biologically relevant alternative splicing and its mode of regulation. a** | Autism spectrum disorders (ASD). High levels of the splicing factor serine/arginine repetitive matrix protein 4 (SRRM4) promote inclusion of microexons in various mRNAs expressed in the central nervous system, which is associated with low risk of ASD. Protein segments encoded by microexons participate in physiological protein–protein interactions. In individuals with ASD, low levels of SRRM4 cause skipping of microexons, thereby disrupting protein–protein interactions[66]. **b** | Body temperature. The serine/arginine-rich (SR) protein kinase CDC-like kinase 1 (CLK1) is a thermosensor that globally regulates alternative splicing. At 35 °C, the activation domain of CLK1 promotes its enzymatic activity, resulting in phosphorylation of different SR proteins, which in turn enhance inclusion of alternative exons in many genes. By contrast, at 38 °C the activation domain becomes flexible and loses its activating capacity, resulting in SR protein hypophosphorylation and global perturbation of alternative splicing[72]. **c** | Circadian rhythms in *Drosophila melanogaster*. The gene *tim* controls the circadian locomotor activity of fruit flies through the expression of three mRNA splicing isoforms encoding three protein variants: Tim-short&cold, Tim-cold and Tim-medium. The respective mRNA isoforms include introns 10, 16 and 13, which all contain a premature stop codon (STOP). When flies are at 18 °C, expression of Tim-short&cold and Tim-cold is predominant, which promotes a high day/night ratio of locomotor activity with no midday siesta. At 29 °C, expression of Tim-medium is predominant, which promotes a low day/night ratio of locomotor activity, including a siesta in the middle of the hot day[73]. $(A)_n$, polyadenylation tail; E1, exon 1.

hypothesis is that metabolites produced by the microbiota act on the nervous system as neurotransmitters. Two of these metabolites, 5-aminovaleric acid and taurine, act as γ-aminobutyric acid A ($GABA_A$)-receptor agonists and their levels are increased in the colons of mice transplanted with typically developing microbiota, but depleted in mice transplanted with ASD microbiota. Consistently, treatment of an ASD mouse model with these metabolites mitigates the ASD-like behavioural abnormalities and reduces neuronal excitability in the brain[71]. Whether these metabolites reverse brain changes in alternative splicing patterns remains to be investigated.

γ-Aminobutyric acid A ($GABA_A$). A small organic molecule synthesized from the amino acid glutamic acid that acts as a neurotransmitter whose receptors are part of ligand-gated ion channels.

Regulation of body temperature. SR protein activity is regulated by phosphorylation and dephosphorylation cycles. A set of key protein kinases that phosphorylate SR proteins belong to the family of CDC-like kinases (CLKs). A recent study discovered that the catalytic activity of CLK1 is highly responsive to physiological changes in body temperature of just 1–2 °C, and that CLKs may act as temperature sensors that globally control alternative splicing patterns[72]. CLK1 activity is more than fivefold higher at 35 °C than at 38 °C in human cells and in cell-free extracts, which is counter-intuitive to the thermodynamics of enzymatic activity. The reason is that, as revealed by molecular dynamics simulations of the known kinase structures at different temperatures, the 25 amino acid activation domain of CLK1 becomes more flexible at high temperatures, which reversibly inhibits the kinase activity[72] (FIG. 2b). Temperature does not affect the structure of the kinase domain, which indicates that the CLK activation domain is the actual cellular thermosensor. This conclusion was corroborated by domain-swapping experiments with SRSF protein kinases (SRPKs), the other family of protein kinases that phosphorylate SR proteins but are temperature insensitive: an SRPK with a CLK activation domain becomes

thermo-sensitive and a CLK with an SRPK activation domain does not respond to temperature changes[72]. The study not only shows that the CLK thermosensor is necessary for temperature control of global patterns of gene expression and alternative splicing but also that in poikilotherms — animals that adapt their body temperature to the ambient temperature — such as alligators, turtles and the fruitfly *Drosophila melanogaster*, the CLK homologues are switched off at physiologically relevant high temperatures as happens with the mammalian CLKs. In *D. melanogaster*, for example, a thermo-sensitive alternative splicing switch is part of locomotor activity control by the circadian clock. Although flies are diurnal animals, the circadian clock control allows them to reduce their locomotor activity on hot days, including a midday siesta, which is compensated by higher activity at the beginning and end of the night[73] (FIG. 2c).

Alternative splicing profiles characterize the identity and function of tissues and organs. A seminal paper compared organ transcriptomes from vertebrates covering 350 million years of evolution, and demonstrated that primates have the highest complexity in alternative splicing and that their alternative splicing profiles overall evolved such that they are more related to the identity of the species than to the type of organ[74]. In other words, the alternative splicing profile of a human liver resembles more alternative splicing profiles of other human organs than the alternative splicing profile of a chimpanzee liver. Similar conclusions were reached by another study[75]. Rapid changes in alternative splicing patterns among vertebrate species are not only due to DNA sequence changes in splicing regulatory sequences, but also owing to specific changes in splicing regulators, as demonstrated for the role of a mammalian-specific splicing isoform of the splicing regulator polypyrimidine tract-binding protein 1 (PTBP1) in shaping a mammalian-like alternative splicing profile when expressed in chickens[76]. These results highlight the importance of alternative splicing diversification in shaping molecular and likely phenotypic differences between species.

Nevertheless, conserved tissue-specific and organ-specific alternative splicing patterns coexist with species-specific patterns, and alterations of either may perturb tissue and organ homeostasis. For instance, loss-of-function mutations in the gene encoding RNA binding motif protein 20 (RBM20) cause dilated cardiomyopathy in humans. Depletion of rat *Rbm20* not only promotes a splicing isoform switch in the sarcomere protein titin, previously implicated in heart function, but also alters splicing patterns in a gene set that is enriched in genes linked to cardiomyopathy, ion homeostasis and sarcomere biology[77], which highlights the role of alternative splicing in the pathogenesis of heart failure. Other examples of tissue-specific and organ-specific alternative splicing include blood flow-driven changes in alternative splicing patterns in endothelial cells that are linked to vascular inflammation[78], changes in adipose tissue alternative splicing patterns caused by diet-induced obesity[79] and circadian rhythm-dependent patterns of alternative splicing in pancreatic cells that control peptidergic secretion and glucose homeostasis[80].

Peptidergic secretion
The secretion of small polypeptides with neurotransmitter activity.

Alternative splicing profiles can, in some cases, better distinguish between cell types than gene expression profiles, for example, in comparisons of data from different neuronal subtypes across developmental stages[81].

Biologically relevant gene-specific alternative splicing

In this section, we discuss selected examples of how gene-specific alternative splicing helps shaping various phenotypic features that do not involve global changes in alternative splicing patterns.

Two textbook examples. Sometimes it is necessary to go back to old literature to reaffirm the validity of a questioned concept such as the functional importance of protein variants arising from alternative splicing. The textbook case of the *CALCA* gene is an excellent example[44,82]. The gene encodes two proteins — calcitonin and calcitonin gene-related peptide 1 (CGRP1) — and is expressed in the neurons of the hypothalamus and in the parafollicular cells, also known as C cells, of the thyroid gland. A combination of alternative splicing and alternative cleavage and polyadenylation results in the production of the *CGRP1* mRNA in neurons and of the calcitonin mRNA in thyroid C cells (FIG. 3A). Although both mRNAs have common and different coding sequence segments, the polypeptides encoded by each of them are subjected to cell-specific proteolytic processes that produce two mature proteins, calcitonin and CGRP1, of completely different amino acid sequence. This case questions any simplification of the definition of a gene and confirms François Jacob's statement that natural selection is a tinkerer: calcitonin and CGRP1 are different proteins — they have different amino acid sequences, structures and functions and are expressed in different tissues. The only thing they have in common is that they are encoded by the same gene! No wonder the neuronal protein has been named 'calcitonin *gene*-related peptide' and not 'calcitonin-related peptide'.

Another textbook example is that of Down syndrome cell adhesion molecule (DSCAM) of *D. melanogaster*[46,83]. DSCAM is a neural cell adhesion membrane protein of the immunoglobulin superfamily. Through a complex mechanism of alternative splicing of mutually exclusive exons in four regions of the pre-mRNA, the *Dscam* gene can give rise to up to 38,000 different alternative splicing variants, half of which determine 19,000 different extracellular domains (FIG. 3Ba). Even if we did not know the underlying mechanism and function of this amazing number of variants, it is remarkable that the number of mRNA variants produced by this single gene is higher than the number of genes in the fly genome (ca. 15,000). Nevertheless, a theoretical model for the mechanism underlying the stochastic generation of the multiple *Dscam* mRNA variants has been proposed[84]. Regarding functionality, this extensive alternative splicing and the resulting variety of produced proteins is the molecular basis of neurite self-avoidance, that is, the prevention of neurites from a given neuron from establishing synapses within the same neuron, thereby imposing non-self connections that guarantee the establishment and function of proper neuronal networks. Self-avoidance is achieved

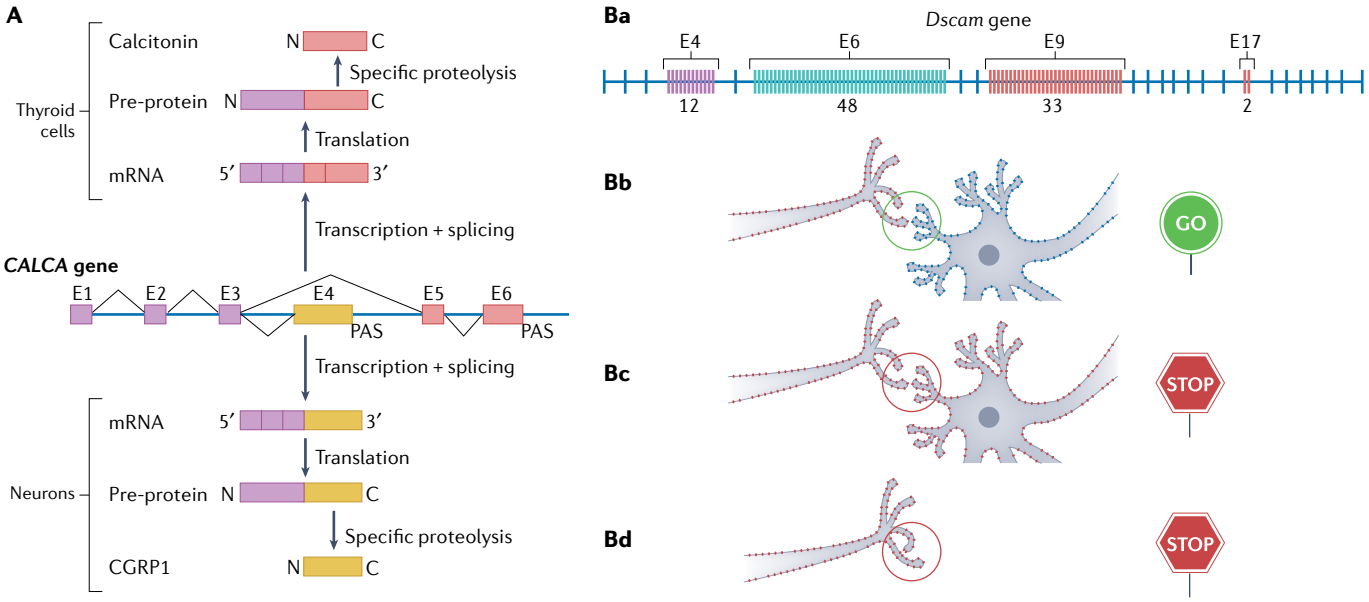


Fig. 3 | **Protein variants arising from alternative splicing. A** | One gene, two very different proteins. The calcitonin-related polypeptide-α (*CALCA*) gene encodes both calcitonin in thyroid cells and calcitonin gene-related peptide 1 (CGRP1) in neurons, through a combination of alternative splicing and specific proteolysis. Skipping of exon 4 (E4) in thyroid cells produces an mRNA variant containing E1, E2, E3, E5 and E6, the two latter exons encoding calcitonin. Translation generates a pre-calcitonin polypeptide that undergoes proteolytic processing giving rise to calcitonin. In neurons of the hypothalamus, E4 is included in the mRNA and owing to the presence of a cleavage and polyadenylation site (PAS) in the exon, E5 and E6 are excluded from the mRNA. Translation generates a pre-CGRP1 polypeptide, which following proteolytic processing gives rise to CGRP1 (REFS.[44,82]). **B** | One gene, thousands of mRNAs. **Ba** | Schematic diagram of the *Drosophila melanogaster* Down syndrome cell adhesion molecule (*Dscam*) gene, showing four regions of mutually exclusive alternatively spliced exons[83]. The number of possible variants is indicated below each region. For example, the E4 region comprises 12 exons ($E4_1$–$E4_{12}$) but owing to the mutually exclusive mode of splicing, if any of these 12 exons is included in the mature mRNA, the other 11 are excluded from it. Inclusion of a particular exon in each of the 4 regions is stochastic, which can give rise to 38,000 different mRNA variants, of which 19,000 correspond to different extracellular domains. **Bb**–**Bd** | Mechanism of neuron self-avoidance in *D. melanogaster*: two neurons having different sets of DSCAM variants in their membranes (simplified here with a single colour for each set) do not recognize each other as self and are allowed to establish synapses (part **Bb**); in an extremely unlikely situation of two neurons having the same DSCAM set in their membranes, synapses would not be established because the homophilic recognition between the identical extracellular domains triggers repulsion — however, owing to its unlikeliness, this case cannot provide the selective pressure required for the existence of the highly variable DSCAM system (part **Bc**); instead, the selective advantage of the DSCAM system is that it guarantees self-avoidance, that is, prevention of the establishment of a synapsis within the same neuron (part **Bd**).

because each neuronal cell precursor stochastically expresses a unique subset of few DSCAM variants in its membrane, which differs from the subsets of other neurons. Identical DSCAM extracellular domains bind to each other through homophilic recognition, which promotes repulsion, whereas different extracellular domains do not bind to each other and allow non-self interactions (FIG. 3Bb–Bd). The beauty of this sophisticated molecular strategy is that although self-avoidance requires neurons to express different sets of DSCAM isoforms, the identity of the isoforms they express is irrelevant. An analogous stochastic process that involves neither DSCAM homologues nor alternative splicing operates to guarantee self-avoidance in vertebrate neurons. The mechanism involves a random choice of alternative promoters of a set of clustered genes encoding up to 60 variants of the cell adhesion protein protocadherin[85]. The radically different *D. melanogaster* and vertebrate mechanisms also illustrate the nature of natural selection: it selects phenotypes (self-avoidance) no matter the genetics underlying them.

Binary splicing switches. Many single-gene, biologically relevant alternative splicing events involve two optional isoforms. Perhaps the most spectacular example is infrared radiation sensing by vampire bats, which are obligate blood feeders. Infrared sensing was studied in the vampire bat *Desmodus rotundus*[43], whose geographical distribution spans from Mexico to Argentina. Detecting infrared radiation through a pair of specialized pit organs on the face allows vampire bats to locate warm-blooded prey such as cattle and chickens. Vampire bats express the heat-sensitive, non-selective cation channel transient receptor potential cation channel subfamily V member 1 (TRPV1) in dorsal root ganglia, similar to all vertebrates, and in the thermal-sensitive fibres of their trigeminal ganglia, which innervate their facial pits. TRPV1 opens at temperatures above 43 °C, which allows all vertebrates to detect noxious heat. Alternative splicing of *TRPV1* mRNA occurring exclusively in the trigeminal ganglia produces a TRPV1 variant that lacks a segment of its C-terminal domain, thereby lowering its thermal activation threshold to 30 °C. This ganglion-specific splicing enables the bats to 'see' heat (prey) while maintaining the burning-prevention role of TRPV1 in the rest of their body, thereby evolutionarily avoiding the need for separate genes to guarantee both food seeking and burning prevention (FIG. 4a). Unfortunately, the molecular mechanism of the *TRPV1* splicing switch between dorsal root ganglia and trigeminal ganglia remains unknown.

A more recent example of a binary splicing switch involves telomere length maintenance by the telomerase complex. Telomerase is highly expressed in embryonic cells to compensate for telomere shortening owing to the substantial number of cell divisions they undergo. As differentiation proceeds and the number of cell divisions diminishes, telomerase expression is gradually silenced (to be resumed in cancer cells in support of their proliferation). Until recently, it was assumed that transcriptional control of the gene encoding the telomerase reverse transcriptase (*TERT*), the protein moiety of telomerase, was the source of developmental changes in telomerase expression. However, it was recently reported that alternative splicing of human *TERT* (*hTERT*) exon 2 (E2) is more crucial for this process than transcription. Skipping of E2 triggers decay of *hTERT* mRNA in differentiated cells and its retention supports telomerase function in pluripotent cells. The splicing cofactor SON is a positive regulator of E2 inclusion and an individual with a mutation in the *SON* gene was found to have insufficient telomerase and shorter telomeres[86] (FIG. 4b).

The methyltransferase G9a (also known as EHMT2) catalyses H3K9me2, a transcription-silencing modification of mammalian euchromatin. The main alternative splicing event at the *G9a* gene generates two mature mRNA isoforms, differing by the inclusion of E10. Our group has shown that E10 inclusion is upregulated during differentiation of mouse neuronal cell lines and in the developing mouse brain. Without having any clear binding or catalytic sites, the G9a segment encoded by E10 increases G9a nuclear localization and activity. The intrinsically unstructured nature of the protein segment encoded by E10 is predicted to promote higher exposure of a neighbouring constitutive nuclear localization signal. Furthermore, H3K9me2 promotes E10 inclusion, which defines a positive feedback loop that reinforces the cellular commitment to neuron differentiation[87].

Forkhead box protein P1 (FOXP1) is a member of the forkhead family of transcription factors, which recognize specific DNA sequences through their forkhead domain. A landmark study discovered that alternative splicing of the *FOXP1* mutually exclusive E18 and E18b was key to determining whether cells will behave as pluripotent embryonic stem cells or engage in a cell-differentiation programme[88]. The authors showed that the FOXP1 splicing variant that includes E18 binds preferentially to the DNA consensus sequence GTAAACA and activates transcription of differentiation genes, whereas the variant that includes E18b (named FOXP1-ES) prefers binding to sites with AATAAACA or CGATACAA sequences and stimulates the expression of transcription factor genes required for pluripotency, including OCT4, NANOG, NR5A2 and GDF3, while repressing transcription of the differentiation genes. So, thanks to its capacity to drive the expression of two transcription factors with different binding specificities, alternative splicing is paramount to the choice between proliferation and differentiation, which is readily related to cancer. However, in the words of the Roman poet Juvenal: *Quis custodiet ipsos custodes* (Who will guard the guards)? The answer is that upstream of FOXP1 alternative splicing are the RNA-binding proteins muscleblind-like protein 1 (MBNL1) and MBNL2, which inhibit E18b inclusion and repress the switch to the embryonic stem cell phenotype. Indeed, depletion of MBNL proteins activates expression of key pluripotency genes and the formation

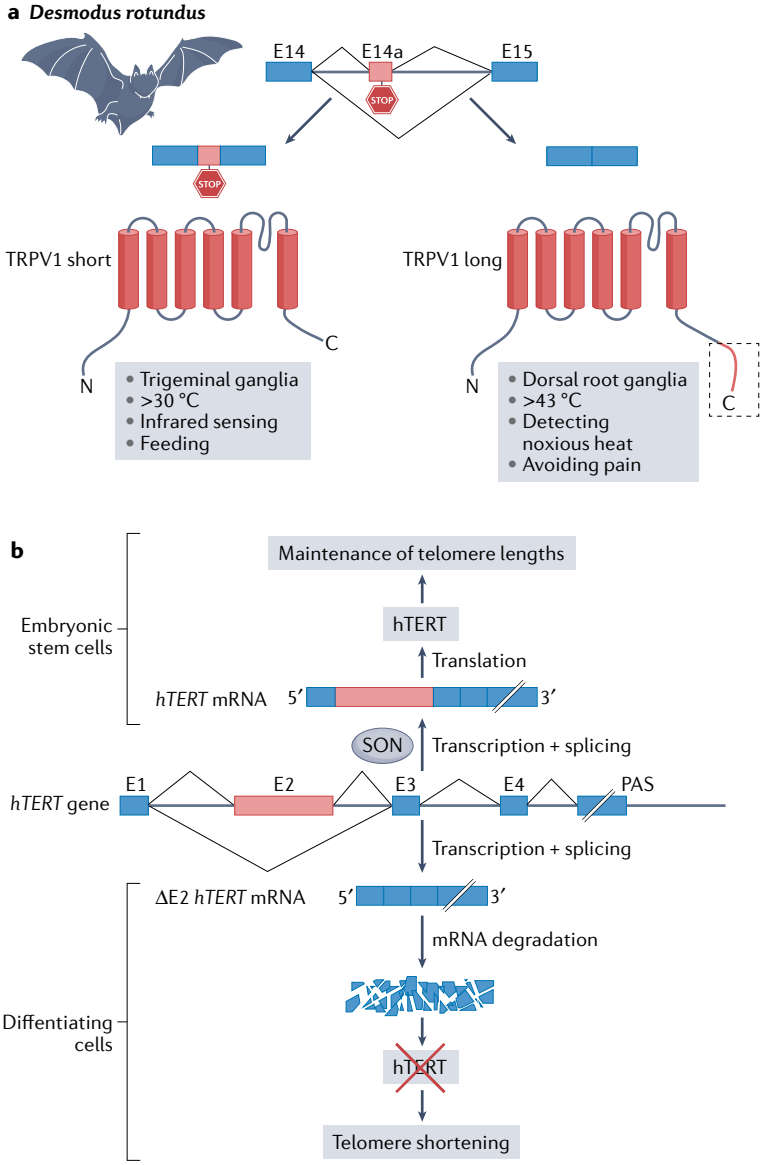


Fig. 4 | **Binary splicing switches. a** | Vampire bats 'see' heat projected from their prey thanks to alternative splicing. The non-selective cation channel transient receptor potential cation channel subfamily V member 1 (TRPV1) is heat sensitive. Terminals of neurons emerging from dorsal root ganglia of all vertebrates, including vampire bats, have TRPV1 channels that open at temperatures above 43°C, which allows the detection of noxious heat. Inclusion of exon 14a (E14a) in the *TRPV1* mRNA exclusively in neurons of the trigeminal ganglia of vampire bats introduces a premature translation stop codon (STOP) and gives rise to a protein that lacks a segment of its carboxy-terminal domain, thereby lowering its temperature threshold to 30 °C and allowing the vampire bats to 'see' their warm-blooded prey (in addition to retaining their ability to sense noxious heat through dorsal root ganglia)[43]. **b** | Telomere length maintenance. The human telomerase reverse transcriptase (*hTERT*) gene encodes the protein moiety of the telomerase complex. The levels of its full-length mRNA and, consequently, of its protein are not only controlled through transcription (not shown) but also by alternative splicing. Inclusion of *hTERT* E2 is stimulated by the splicing factor SON in dividing cells such as embryonic stem cells, to benefit from the maintenance of telomere lengths, which prevents proliferation arrest. In differentiating cells, low levels of SON result in E2 skipping with subsequent *hTERT* degradation, which results in telomere shortening[86]. PAS, cleavage and polyadenylation site.

of induced pluripotent stem cells from differentiated cells[89].

Another switch between mutually exclusive exons regulates energy production in cancer cells. In the presence of oxygen, normal cells convert glucose to pyruvate through glycolysis and pyruvate is taken by mitochondria to enter the Krebs cycle and produce big amounts of ATP through oxidative phosphorylation. In the absence of oxygen, these cells convert pyruvate to lactate in the cytosol through anaerobic glycolysis completed as lactic fermentation. Cancer cells, however, consume glucose more avidly by producing lactate through aerobic glycolysis, a process known as the Warburg effect after the German physiologist who discovered it. The Warburg effect depends on the glycolysis enzyme pyruvate kinase (PKM), which converts phosphoenol pyruvate into pyruvate. Mammals express two PKM splicing isoforms, PKM1 and PKM2, which differ in the segments encoded by the mutually exclusive E9 and E10 (both coding for 56 amino acids) of the *PKM* gene. Inclusion of E9 yields PKM1, which forms a tetramer and is constitutively active. Inclusion of E10 yields PKM2, which is more abundant in embryonic and cancer cells and forms both an active tetramer and a dimer whose basal activity is low but can be allosterically activated. Due to its lower activity, PKM2 provides cancer cells with an advantage compared with normal cells because the major amount of glucose present in the cell remains in the form of glycolysis intermediates, which provide the highly proliferating cells with building blocks for macromolecules such as nucleotides, amino acids and fatty acids. In other words, PKM2 expression enhances anabolic pathways needed for producing new cells[90]. *PKM* alternative splicing is controlled by three hnRNP splicing regulators — hnRNPA1, hnRNPA2 and PTBP1 (also known as hnRNPI) — which, by binding to target sequences flanking E9, promote E10 inclusion and increase the PKM2 to PKM1 expression ratio[91]. Interestingly, transcription of the genes encoding these splicing regulators is activated by MYC, a major transcription factor upregulated in cancer[90].

The list of physiologically important alternative splicing switches is vast and we chose to discuss just a few examples to highlight their functional diversity. Further reading on this topic may include the roles of two splicing isoforms of the enzyme hexokinase in hypoxia[92], of the calcium channel $Ca_V1.2$ in heart homeostasis[93] and of the transcription factor myocyte-specific enhancer factor 2 in muscle differentiation[94].

Therapeutic modulation of alternative splicing

Antisense oligonucleotides (ASOs) are synthetic single-stranded short RNA or DNA molecules, which through base-pairing with complementary sequences in mRNA or pre-mRNA targets affect gene expression. Several chemical modifications are usually introduced in ASOs, both in the ribose and in the backbone phosphate groups in order to improve their stability and half-life in body fluids (reviewed in REF.[95]). In this section we discuss only the clinically approved use of ASOs through the control of alternative splicing, in two hereditary diseases: spinal muscular atrophy (SMA) and Duchenne muscular dystrophy (DMD).

SMA is a motor-neuron disease caused by recessive mutations in the survival motor neuron 1 (*SMN1*) gene, which encodes the SMN protein. Humans have a paralog, *SMN2*, of approximately the same length (~30 kbp) resulting from duplication of a large part of chromosome 5. The sequences of *SMN1* and *SMN2* differ in only 11 nucleotides, including a crucial C > T base change located in *SMN2* E7, which increases its skipping[96]. The predominant *SMN2* mRNA variant lacks E7 and produces a truncated and inactive SMN protein. The *SMN2* variant that includes E7 produces the full-length, functional SMN protein but at low levels that are not sufficient to compensate for the lack of functional SMN protein in the absence of the wild type *SMN1* gene. Both *SMN1* and *SMN2* pre-mRNAs have a splicing silencer sequence in intron 7 that is the target site of the splicing silencing factors hnRNPA1 and hnRNPA2, which promote E7 skipping in *SMN2* (REF.[97]). Based on this knowledge, Krainer and colleagues developed nusinersen (Spinraza), a splicing-correcting ASO drug approved for clinical use by the USA Food and Drug Administration (FDA) in 2016 and by the European Medicines Agency (EMA) in 2017. Nusinersen targets the splicing silencer located in the *SMN2* intron 7 pre-mRNA and, by blocking the binding of hnRNPA1 and hnRNPA2, promotes E7 inclusion and increases the abundance of SMN. Nusinersen was the first drug approved for SMA therapy and also the first splicing-corrective drug[98–102].

We have recently added a new layer of regulation of *SMN2* E7 inclusion that may lead to a combined therapy for SMA[103]. We found that *SMN2* E7 is a class II alternative exon, because its inclusion is inhibited by slow, and stimulated by fast, transcript elongation. We found that by promoting transcript elongation, histone deacetylase (HDAC) inhibitors cooperate with a nusinersen-like ASO to upregulate E7 inclusion. Surprisingly, the ASO also elicits H3K9me2, a transcription-repressing modification, at the *SMN2* gene, thereby creating a roadblock to Pol II elongation and inhibiting E7 inclusion, which contradicts the positive effect of the ASO in the displacement of hnRNPA1 and hnRNPA2. By removing the roadblock, HDAC inhibition counteracts the undesired chromatin modification effects of the ASO, resulting in higher E7 inclusion[103] (FIG. 5). Combined systemic administration of HDAC inhibitors such as valproic acid (VPA) or trichostatin A with a nusinersen-like ASO in neonate SMA mice had strong synergistic effects on SMN expression in the brain, liver, kidney, muscle and spinal cord. Furthermore, the combined treatment improved growth, survival and neuromuscular functions in the SMA mice. Thus, HDAC inhibitors have the potential to increase the clinical efficacy of nusinersen without having large pleiotropic effects, as demonstrated by genome-wide RNA-seq analyses of gene expression and alternative splicing[103].

Another therapeutic treatment for SMA has been recently approved. It is not based on an oligonucleotide but on a small molecule, risdiplam, that affects snRNA–5′ splice site base-pairing in the *SMN2* pre-mRNA and promotes E7 inclusion with a subsequent increase in levels of functional SMN protein in the blood[104]. It would be interesting to know whether HDAC inhibitors also increase the clinical efficacy of risdiplam.

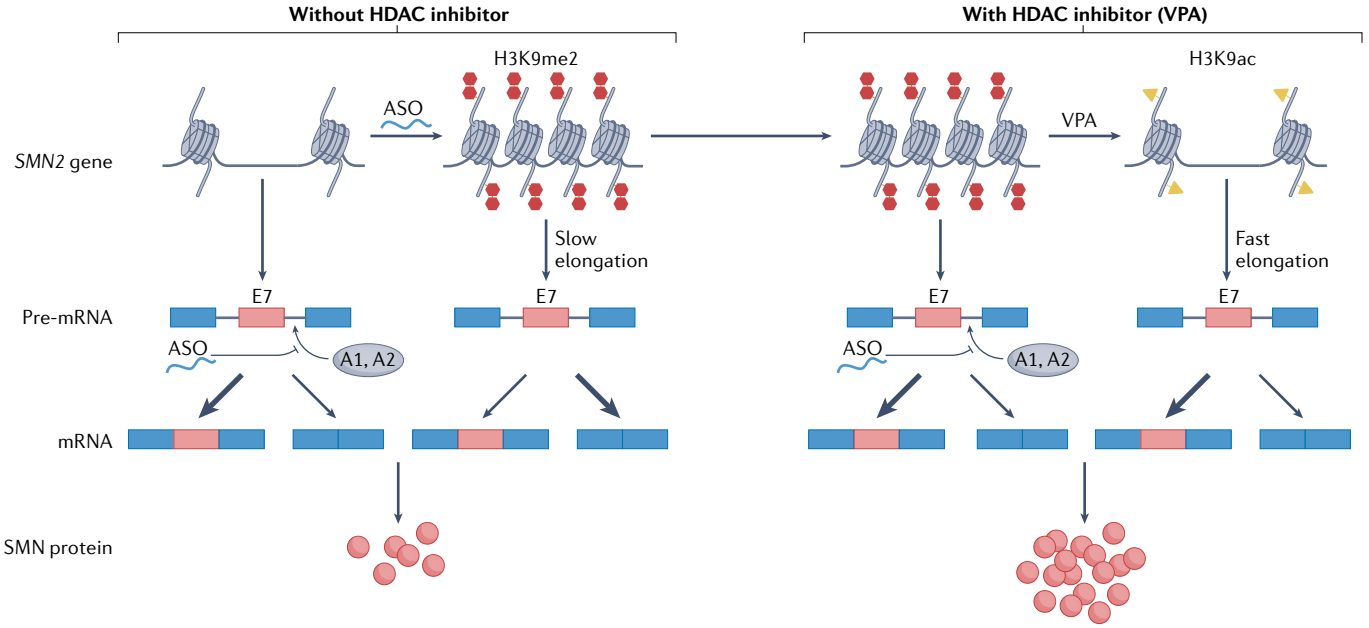


Fig. 5 | **Therapeutic modulation of alternative splicing.** Regulation of alternative splicing by splicing factors and chromatin modification towards a combined therapy for spinal muscular atrophy (SMA)[103]. Effects of the splicing-correcting antisense oligonucleotide (ASO) nusinersen (Spinraza) on exon 7 (E7) inclusion in the survival motor neuron 2 (*SMN2*) mRNA are illustrated. In the absence of histone deacetylase (HDAC) inhibitors (left), nusinersen has two opposing effects on E7 splicing. On the one hand, it targets an intronic splicing silencer (ISS) in intron 7, and displaces the splicing silencing factors heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and hnRNPA2 (A1, A2), thereby favouring E7 inclusion. On the other hand, nusinersen promotes dimethylation of histone H3 Lys9 (H3K9me2), which creates a roadblock to transcript elongation by RNA polymerase II (Pol II) (not shown), resulting in E7 exclusion. (The latter effect is due to E7 being a class II exon, whose inclusion is inhibited by slow elongation.) In the presence of HDAC inhibitors (right) such as valproic acid (VPA), the negative, H3K9me2-mediated effect of nusinersen on E7 inclusion is mitigated by their promotion of acetylation of histone H3 Lys9 (H3K9ac) and, thus, of faster Pol II elongation.

DMD is a sex-linked hereditary disease that causes loss of muscle function due to the lack of the protein dystrophin. Among the different mutations in the dystrophin gene, deletions of a segment encompassing E49 and E50 and their flanking introns account for 13% of DMD cases[105,106]. These deletions generate an mRNA in which E48 is joined to E51, which creates a translation frameshift and a premature stop codon in E51 that triggers degradation of the mutant mRNA by NMD. A variant of DMD with much milder symptoms, known as Becker muscular dystrophy[107], is caused by deletions in the *DMD* gene that maintain the open reading frame and result in a partially functional dystrophin protein with an internal truncation. Inspired by this protein, a 30-mer ASO, now known as eteplirsen, was designed that targets an ESE within E51 and displaces the splicing activators that bind to it, thereby inhibiting E51 inclusion[105]. The skipping of E51 restores the open reading frame by merging E48 directly to E52; this mRNA encodes an internally truncated version of dystrophin that is partially functional and mimics the phenotype of Becker muscular dystrophy in mice.

Conclusions

It is usually accepted that the two pillars that support a positive adaptive value of the exon–intron organization of eukaryotic genes are exon shuffling and alternative splicing. For many years after the discovery of alternative splicing, this mechanism was thought to affect a limited number of genes with biologically relevant consequences. The advent of the deep-sequencing technology applied to mRNAs revealed that most genes have many splicing variants in more than one region along their sequence. However, the proteomic studies discussed above could not find a correspondence between these mRNA variations and the encoded proteins. In our opinion, the combination of the extreme sensitivity of mRNA-seq with the low sensitivity of mass spectrometry is hampering the evaluation of the biological relevance of alternative splicing. It is likely that a considerable portion of the mRNA isoforms detected by RNA-seq are in such a low concentration that they have no functional, observable role in cells. At the same time, if we assume that, for example, 50% of the mRNA splicing variation detected by RNA-seq is functionally relevant, that variation should be seen in a proteomic study with high sensitivity. The use of short-read sequencing in traditional RNA-seq methods limits the assessment of full-length mRNA isoforms and may also explain the variations in mRNA and protein isoform detection. The third-generation sequencing technology, which produces long reads, appears to be a powerful tool to re-evaluate the apparent contradiction between transcriptomic and proteomic data, as recently reviewed elsewhere[108]. Another relevant issue is to what extent NMD contributes to the degradation of mRNA splicing isoforms. Nearly 15,000 human mRNA alternative splicing variants are annotated as NMD targets in the 2019 Ensembl database[109]. This connection is consistent with the discovery that coupling between alternative splicing

Exon shuffling
A mechanism for the formation of new genes in eukaryotes during evolution, usually through recombination between introns of different genes, yielding novel rearranged genes with altered functions, without elimination of the original genes.

r-Selection
A type of evolutionary selective pressure. r-Selected species, such as bacteria, are small organisms with short life cycles and fast maturation that are able to rapidly populate new environments.

K-Selection
A type of evolutionary selective pressure. K-Selected species, such as pluricellular organisms with specialized tissues and organs, are bigger, have longer life cycles and slow maturation, and their colonization of new environments depends on their ability to adapt physiologically.

and NMD is a widespread process that controls protein production at the post-transcriptional level[110,111].

Despite representing only the tip of an iceberg of the accumulated scientific literature, the multiple examples of alternative splicing discussed in this Review indicate that the regulation of alternative splicing is as important as the regulation of transcription in shaping cell phenotypes and physiology, at least in multicellular eukaryotes. Unicellular eukaryotes have smaller genomes than multicellular eukaryotes with few and shorter introns or no introns at all, which supports shorter life cycles. The fact that alternative splicing prevails in multicellular eukaryotes but not in unicellular ones cannot be neglected when assessing a fundamental role for alternative splicing in generating complexity. It has been proposed that ancestral eukaryotes had introns and that modern unicellular eukaryotes lost a substantial fraction of them[112] as an adaptation to an r-selection reproductive strategy. By contrast, in a K-selection strategy typical of most multicellular organisms, the alternative splicing contribution to complexity and physiological adaptation becomes advantageous. Herein, we focus on animals, but we should bear in mind that alternative splicing is also regulated in viruses[47] and in the cells infected by them[113]; in plants[114,115] in response to external cues[116]; and as part of circadian rhythms[117]. We wish to support the idea that, although several alternative splicing events in a given organism may not show a selective advantage, the existence of alternative splicing per se has a positive adaptive value because it provides a readily available pool of mRNA variability that can be co-opted during evolution to generate relevant functional variability without the need to eliminate existing gene functions or generate new genes. The above-discussed molecular strategy that allows infrared sensing in vampire bats is an outstanding example of the adaptive value of alternative splicing.

Published online 13 October 2022

1. Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. *Cell* **12**, 1–8 (1977).
2. Berget, S. M., Moore, C. & Sharp, P. A. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proc. Natl Acad. Sci. USA* **74**, 3171–3175 (1977).
3. Gilbert, W. Why genes in pieces? *Nature* **271**, 501 (1978).
4. Khan, M. R., Wellinger, R. J. & Laurent, B. Exploring the alternative splicing of long noncoding RNAs. *Trends Genet.* **37**, 695–698 (2021).
5. Rogers, S. O. Integrated evolution of ribosomal RNAs, introns, and intron nurseries. *Genetica* **147**, 103–119 (2019).
6. Sekulovski, S. et al. Assembly defects of human tRNA splicing endonuclease contribute to impaired pre-tRNA processing in pontocerebellar hypoplasia. *Nat. Commun.* **12**, 5610 (2021).
7. Kearse, M. G. & Wilusz, J. E. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.* **31**, 1717–1731 (2017).
8. Smith, C. W. & Valcárcel, J. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.* **25**, 381–388 (2000).
9. Kastner, B., Will, C. L., Stark, H. & Lührmann, R. Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harb. Perspect. Biol.* **11**, a032417 (2019).
10. Licatalosi, D. D. & Darnell, R. B. RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.* **11**, 75–87 (2010).
11. Irimia, M. & Blencowe, B. J. Alternative splicing: decoding an expansive regulatory layer. *Curr. Opin. Cell Biol.* **24**, 323–332 (2012).
12. Kornblihtt, A. R. et al. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.* **14**, 153–165 (2013).
13. Fu, X.-D. & Ares, M. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* **15**, 689–701 (2014).
14. Baralle, F. E. & Giudice, J. Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **18**, 437–451 (2017).
15. Ule, J. & Blencowe, B. J. Alternative splicing regulatory networks: functions, mechanisms, and evolution. *Mol. Cell* **76**, 329–345 (2019).
16. Shenasa, H. & Hertel, K. J. Combinatorial regulation of alternative splicing. *Biochim. Biophys. Acta Gene Regul. Mech.* **1862**, 194392 (2019).
17. Gordon, J. M., Phizicky, D. V. & Neugebauer, K. M. Nuclear mechanisms of gene expression control: pre-mRNA splicing as a life or death decision. *Curr. Opin. Genet. Dev.* **67**, 67–76 (2021).
18. Buratti, E., Baralle, M. & Baralle, F. E. From single splicing events to thousands: the ambiguous step forward in splicing research. *Brief. Funct. Genomics* **12**, 3–12 (2013).
19. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
20. Wang, E. T. et al. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
21. Saldi, T., Riemondy, K., Erickson, B. & Bentley, D. L. Alternative RNA structures formed during transcription depend on elongation rate and modify RNA processing. *Mol. Cell* **81**, 1789–1801.e5 (2021).
22. Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R. & Misteli, T. Epigenetics in alternative pre-mRNA splicing. *Cell* **144**, 16–26 (2011).
23. de la Mata, M. et al. A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* **12**, 525–532 (2003).
24. Dujardin, G. et al. How slow RNA polymerase II elongation favors alternative exon skipping. *Mol. Cell* **54**, 683–690 (2014).
25. Fong, N. et al. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev.* **28**, 2663–2676 (2014).
26. Maslon, M. M. et al. A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. *EMBO J.* **38**, e101244 (2019).
27. Schor, I. E., Rascovan, N., Pelisch, F., Alló, M. & Kornblihtt, A. R. Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc. Natl Acad. Sci. USA* **106**, 4325–4330 (2009).
28. Alló, M. et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* **16**, 717–724 (2009).
29. Schor, I. E., Fiszbein, A., Petrillo, E. & Kornblihtt, A. R. Intragenic epigenetic changes modulate NCAM alternative splicing in neuronal differentiation. *EMBO J.* **32**, 2264–2274 (2013).
30. Segelle, A. et al. Histone marks regulate the epithelial-to-mesenchymal transition via alternative splicing. *Cell Rep.* **38**, 110357 (2022).
31. Eling, N., Morgan, M. D. & Marioni, J. C. Challenges in measuring and understanding biological noise. *Nat. Rev. Genet.* **20**, 536–548 (2019).
32. Tress, M. L., Abascal, F. & Valencia, A. Most alternative isoforms are not functionally important. *Trends Biochem. Sci.* **42**, 98–110 (2017).
33. Rodriguez, J. M., Pozo, F., di Domenico, T., Vazquez, J. & Tress, M. L. An analysis of tissue-specific alternative splicing at the protein level. *PLoS Comput. Biol.* **16**, e1008287 (2020).
34. Pozo, F. et al. Assessing the functional relevance of splice isoforms. *NAR Genomics Bioinform* **3**, lqab044 (2021).
35. Blencowe, B. J. The relationship between alternative splicing and proteomic complexity. *Trends Biochem. Sci.* **42**, 407–408 (2017).
36. Ingolia, N. T. Ribosome footprint profiling of translation throughout the genome. *Cell* **165**, 22–33 (2016).
37. Weatheritt, R. J., Sterne-Weiler, T. & Blencowe, B. J. The ribosome-engaged landscape of alternative splicing. *Nat. Struct. Mol. Biol.* **23**, 1117–1123 (2016).
38. Kurosaki, T. & Maquat, L. E. Nonsense-mediated mRNA decay in humans at a glance. *J. Cell Sci.* **129**, 461–467 (2016).
39. Floor, S. N. & Doudna, J. A. Tunable protein synthesis by transcript isoforms in human cells. *eLife* **5**, e10921 (2016).
40. Sterne-Weiler, T. et al. Frac-seq reveals isoform-specific recruitment to polyribosomes. *Genome Res.* **23**, 1615–1623 (2013).
41. Xing, Y. & Lee, C. Alternative splicing and RNA selection pressure — evolutionary consequences for eukaryotic genomes. *Nat. Rev. Genet.* **7**, 499–509 (2006).
42. Bell, L. R., Horabin, J. I., Schedl, P. & Cline, T. W. Positive autoregulation of Sex-lethal by alternative splicing maintains the female determined state in *Drosophila*. *Cell* **65**, 229–239 (1991).
43. Gracheva, E. O. et al. Ganglion-specific splicing of TRPV1 underlies infrared sensation in vampire bats. *Nature* **476**, 88–91 (2011).
44. Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature* **298**, 240–244 (1982).
45. Peterson, M. L. Immunoglobulin heavy chain gene regulation through polyadenylation and splicing competition. *Wiley Interdiscip. Rev. RNA* **2**, 92–105 (2011).
46. Hattori, D. et al. Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms. *Nature* **461**, 644–648 (2009).
47. Emery, A. & Swanstrom, R. HIV-1: to splice or not to splice, that is the question. *Viruses* **13**, 181–190 (2021).
48. Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550 (2016).
49. El Marabti, E. & Younis, I. The cancer spliceome: reprograming of alternative splicing in cancer. *Front. Mol. Biosci.* **5**, 80 (2018).
50. Bangru, S. et al. Alternative splicing rewires Hippo signaling pathway in hepatocytes to promote liver regeneration. *Nat. Struct. Mol. Biol.* **25**, 928–939 (2018).
51. Jensen, M. A., Wilkinson, J. E. & Krainer, A. R. Splicing factor SRSF6 promotes hyperplasia of sensitized skin. *Nat. Struct. Mol. Biol.* **21**, 189–197 (2014).
52. Anczuków, O. & Krainer, A. R. Splicing-factor alterations in cancers. *RNA* **22**, 1285–1301 (2016).
53. Cherry, S. & Lynch, K. W. Alternative splicing and cancer: insights, opportunities, and challenges from an expanding view of the transcriptome. *Genes Dev.* **34**, 1005–1016 (2020).
54. Gurnari, C., Pagliuca, S. & Visconte, V. Alternative splicing in myeloid malignancies. *Biomedicines* **9**, 1844–1851 (2021).
55. Karni, R. et al. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.* **14**, 185–193 (2007).

56. Das, S., Anczuków, O., Akerman, M. & Krainer, A. R. Oncogenic splicing factor SRSF1 is a critical transcriptional target of MYC. *Cell Rep.* **1**, 110–117 (2012).
57. Rahman, M. A., Lin, K.-T., Bradley, R. K., Abdel-Wahab, O. & Krainer, A. R. Recurrent SRSF2 mutations in MDS affect both splicing and NMD. *Genes Dev.* **34**, 413–427 (2020).
58. Papasaikas, P., Tejedor, J. R., Vigevani, L. & Valcárcel, J. Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol. Cell* **57**, 7–22 (2015).
59. Saltzman, A. L., Pan, Q. & Blencowe, B. J. Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev.* **25**, 373–384 (2011).
60. Pleiss, J. A., Whitworth, G. B., Bergkessel, M. & Guthrie, C. Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol.* **5**, e90 (2007).
61. Radzisheuskaya, A. et al. PRMT5 methylome profiling uncovers a direct link to splicing regulation in acute myeloid leukemia. *Nat. Struct. Mol. Biol.* **26**, 992–1012 (2019).
62. Braun, C. J. et al. Coordinated splicing of regulatory detained introns within oncogenic transcripts creates an exploitable vulnerability in malignant glioma. *Cancer Cell* **32**, 411–426.e11 (2017).
63. Sanford, J. R., Gray, N. K., Beckmann, K. & Cáceres, J. F. A novel role for shuttling SR proteins in mRNA translation. *Genes Dev.* **18**, 755–768 (2004).
64. Haward, F. et al. Nucleo-cytoplasmic shuttling of splicing factor SRSF1 is required for development and cilia function. *eLife* **10**, e65104 (2021).
65. Munkley, J. et al. Androgen-regulated transcription of ESRP2 drives alternative splicing patterns in prostate cancer. *eLife* **8**, e47678 (2019).
66. Irimia, M. et al. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511–1523 (2014).
67. Ellis, J. D. et al. Tissue-specific alternative splicing remodels protein–protein interaction networks. *Mol. Cell* **46**, 884–892 (2012).
68. Buljan, M. et al. Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell* **46**, 871–883 (2012).
69. Gonatopoulos-Pournatzis, T. et al. Autism-misregulated eIF4G microexons control synaptic translation and higher order cognitive functions. *Mol. Cell* **77**, 1176–1192.e16 (2020).
70. Parras, A. et al. Autism-like phenotype and risk gene mRNA deadenylation by CPEB4 mis-splicing. *Nature* **560**, 441–446 (2018).
71. Sharon, G. et al. Human gut microbiota from autism spectrum disorder promote behavioral symptoms in mice. *Cell* **177**, 1600–1618 (2019).
72. Haltenhof, T. et al. A conserved kinase-based body-temperature sensor globally controls alternative splicing and gene expression. *Mol. Cell* **78**, 57–69 (2020).
73. Martin Anduaga, A. et al. Thermosensitive alternative splicing senses and mediates temperature adaptation in *Drosophila*. *eLife* **8**, e44642 (2019).
74. Barbosa-Morais, N. L. et al. The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**, 1587–1593 (2012).
75. Merkin, J., Russell, C., Chen, P. & Burge, C. B. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338**, 1593–1599 (2012).
76. Gueroussov, S. et al. An alternative splicing event amplifies evolutionary differences between vertebrates. *Science* **349**, 868–873 (2015).
77. Guo, W. et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med.* **18**, 766–773 (2012).
78. Murphy, P. A. et al. Alternative RNA splicing in the endothelium mediated in part by Rbfox2 regulates the arterial response to low flow. *eLife* **7**, e29494 (2018).
79. Vernia, S. et al. An alternative splicing program promotes adipose tissue thermogenesis. *eLife* **5**, e17672 (2016).
80. Marcheva, B. et al. A role for alternative splicing in circadian control of exocytosis and glucose homeostasis. *Genes Dev.* **34**, 1089–1105 (2020).
81. Ha, K. C. H., Sterne-Weiler, T., Morris, Q., Weatheritt, R. J. & Blencowe, B. J. Differential contribution of transcriptomic regulatory layers in the definition of neuronal identity. *Nat. Commun.* **12**, 335 (2021).
82. Leff, S. E., Evans, R. M. & Rosenfeld, M. G. Splice commitment dictates neuron-specific alternative RNA processing in calcitonin/CGRP gene expression. *Cell* **48**, 517–524 (1987).
83. Schmucker, D. et al. *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**, 671–684 (2000).
84. Graveley, B. R. Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell* **123**, 65–73 (2005).
85. Goodman, K. M. et al. How clustered protocadherin binding specificity is tuned for neuronal self-/nonself-recognition. *eLife* **11**, e72416 (2022).
86. Penev, A. et al. Alternative splicing is a developmental switch for hTERT expression. *Mol. Cell* **81**, 2349–2360.e6 (2021).
87. Fiszbein, A. et al. Alternative splicing of G9a regulates neuronal differentiation. *Cell Rep.* **14**, 2797–2808 (2016).
88. Gabut, M. et al. An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell* **147**, 132–146 (2011).
89. Han, H. et al. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* **498**, 241–245 (2013).
90. Choksi, A. et al. Tumor suppressor SMAR1 regulates PKM alternative splicing by HDAC6-mediated deacetylation of PTBP1. *Cancer Metab.* **9**, 16–21 (2021).
91. David, C. J., Chen, M., Assanah, M., Canoll, P. & Manley, J. L. hnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* **463**, 364–368 (2010).
92. Mirtschink, P. et al. HIF-driven SF3B1 induces KHK-C to enforce fructolysis and heart disease. *Nature* **522**, 444–449 (2015).
93. Li, G. et al. Exclusion of alternative exon 33 of Ca_v1.2 calcium channels in heart is proarrhythmogenic. *Proc. Natl Acad. Sci. USA* **114**, E4288–E4295 (2017).
94. Sebastian, S. et al. Tissue-specific splicing of a ubiquitously expressed transcription factor is essential for muscle differentiation. *Genes Dev.* **27**, 1247–1259 (2013).
95. Scharner, J. & Aznarez, I. Clinical applications of single-stranded oligonucleotides: current landscape of approved and in-development therapeutics. *Mol. Ther.* **29**, 540–554 (2021).
96. Monani, U. R. The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in $Smn^{-/-}$ mice and results in a mouse with spinal muscular atrophy. *Hum. Mol. Genet.* **9**, 333–339 (2000).
97. Singh, N. K., Singh, N. N., Androphy, E. J. & Singh, R. N. Splicing of a critical exon of human Survival Motor Neuron is regulated by a unique silencer element located in the last intron. *Mol. Cell. Biol.* **26**, 1333–1346 (2006).
98. Hua, Y., Vickers, T. A., Okunola, H. L., Bennett, C. F. & Krainer, A. R. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *Am. J. Hum. Genet.* **82**, 834–848 (2008).
99. Hua, Y. et al. Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature* **478**, 123–126 (2011).
100. Hua, Y. et al. Motor neuron cell-nonautonomous rescue of spinal muscular atrophy phenotypes in mild and severe transgenic mouse models. *Genes Dev.* **29**, 288–297 (2015).
101. Chiriboga, C. A. et al. Results from a phase 1 study of nusinersen (ISIS-SMN_{Rx}) in children with spinal muscular atrophy. *Neurology* **86**, 890–897 (2016).
102. Rigo, F. et al. Pharmacology of a central nervous system delivered 2′-*O*-methoxyethyl-modified survival of motor neuron splicing oligonucleotide in mice and nonhuman primates. *J. Pharmacol. Exp. Ther.* **350**, 46–55 (2014).
103. Marasco, L. E. et al. Counteracting chromatin effects of a splicing-correcting antisense oligonucleotide improves its therapeutic efficacy in spinal muscular atrophy. *Cell* **185**, 2057–2070 (2022).
104. Darras, B. T. et al. Risdiplam-treated infants with type 1 spinal muscular atrophy versus historical controls. *N. Engl. J. Med.* **385**, 427–435 (2021).
105. Arechavala-Gomeza, V. et al. Comparative analysis of antisense oligonucleotide sequences for targeted skipping of exon 51 during dystrophin pre-mRNA splicing in human muscle. *Hum. Gene Ther.* **18**, 798–810 (2007).
106. Aartsma-Rus, A. et al. Development of exon skipping therapies for Duchenne muscular dystrophy: a critical review and a perspective on the outstanding issues. *Nucleic Acid. Ther.* **27**, 251–259 (2017).
107. Koenig, M. et al. The molecular basis for duchenne versus becker muscular dystrophy: correlation of severity with type of deletion. *Tha Am. J. Hum. Genet.* **45**, 498–506 (1989).
108. Wright, C. J., Smith, C. W. J. & Jiggins, C. D. Alternative splicing as a source of phenotypic diversity. *Nat. Rev. Genet.* https://doi.org/10.1038/s41576-022-00514-4 (2022).
109. Cunningham, F. et al. Ensembl 2019. *Nucleic Acids Res.* **47**, D745–D751 (2019).
110. Lewis, B. P., Green, R. E. & Brenner, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl Acad. Sci. USA* **100**, 189–192 (2003).
111. Kovalak, C., Donovan, S., Bicknell, A. A., Metkar, M. & Moore, M. J. Deep sequencing of pre-translational mRNPs reveals hidden flux through evolutionarily conserved alternative splicing nonsense-mediated decay pathways. *Genome Biol.* **22**, 132 (2021).
112. Keren, H., Lev-Maor, G. & Ast, G. Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* **11**, 345–355 (2010).
113. Ashraf, U., Benoit-Pilven, C., Lacroix, V., Navratil, V. & Naffakh, N. Advances in analyzing virus-induced alterations of host cell splicing. *Trends Microbiol.* **27**, 268–281 (2019).
114. Szakonyi, D. & Duque, P. Alternative splicing as a regulator of early plant development. *Front. Plant. Sci.* **9**, 1174 (2018).
115. Syed, N. H., Kalyna, M., Marquez, Y., Barta, A. & Brown, J. W. S. Alternative splicing in plants — coming of age. *Trends Plant. Sci.* **17**, 616–623 (2012).
116. Godoy Herz, M. A. et al. Light regulates plant alternative splicing through the control of transcriptional elongation. *Mol. Cell* **73**, 1066–1074 (2019).
117. Sanchez, S. E. et al. A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature* **468**, 112–116 (2010).
118. Fairbrother, W. G., Yeh, R.-F., Sharp, P. A. & Burge, C. B. Predictive identification of exonic splicing enhancers in human genes. *Science* **297**, 1007–1013 (2002).

Acknowledgements
The authors apologize to those researchers whose work could not be cited owing to space constraints. They thank B. Blencowe for critical reading of the manuscript and L. Giono for the design of Fig. 3B. Work in the authors' laboratory was supported by a joint grant from Familias Atrofia Muscular Espinal (FAME, Argentina) and CureSMA (USA), and grants from the Lounsbery Foundation (USA), the Universidad de Buenos Aires (UBACYT 20020170100046BA), the Agencia Nacional de Promoción Científica y Tecnológica of Argentina (PICT-2019 862) and the CONICET (PUE 22920170100062CO).

Author contributions
Both authors researched data for the article, made substantial contribution to discussion of content and reviewed and/or edited the manuscript before submission. A.R.K. wrote the manuscript.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Molecular Cell Biology thanks Jimena Giudice, Juan Valcárcel, and Leilei Chen, who co-reviewed with Han Jian, for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

RELATED LINKS
National Human Genome Research Institute, National Institutes of Health (NIH): https://www.genome.gov/genetics-glossary/Exon