Special Issue: 40 Years of TiBS

Review

The Spliceosome:
The Ultimate RNA Chaperone
and SculptorPanagiotis Papasaikas^{1,2} and Juan Valcárcel^{1,2,3,*}

The spliceosome, one of the most complex machineries of eukaryotic cells, removes intronic sequences from primary transcripts to generate functional messenger and long noncoding RNAs (lncRNA). Genetic, biochemical, and structural data reveal that the spliceosome is an RNA-based enzyme. Striking mechanistic and structural similarities strongly argue that pre-mRNA introns originated from self-catalytic group II ribozymes. However, in the spliceosome, protein components organize and activate the catalytic-site RNAs, and recognize and pair together splice sites at intron boundaries. The spliceosome is a dynamic, reversible, and flexible machine that chaperones small nuclear (sn) RNAs and a variety of pre-mRNA sequences into conformations that enable intron removal. This malleability likely contributes to the regulation of alternative splicing, a prevalent process contributing to cell differentiation, homeostasis, and disease.

‘and deem not profitless those fleeting moods’ W. Wordsworth, *The Prelude*

In the Beginning Were Introns

Introns are intervening sequences that are removed from primary transcripts by the process of RNA splicing, which links together the flanking pieces (exons) to generate functional mature RNAs (Figure 1A) [1]. Out of the 16 h required to transcribe the human dystrophin gene, 15 h and 54 min are devoted to the seemingly useless synthesis of intronic RNA. Therefore, efficient removal of introns, with single nucleotide precision, is essential for eukaryotic cells to produce the correct complement of mRNA and lncRNA transcripts necessary to deploy genetic programs. Furthermore, ambiguity in defining certain introns, exons, and their precise boundaries allows alternative patterns of intron removal that greatly expand the diversity of transcripts generated per gene in multicellular organisms [2]. More than 90% of human genes are alternatively spliced, providing extensive opportunities for gene regulation during development, cell differentiation, and homeostasis [3]. For instance, alternative-splicing regulation is instrumental to control cellular and organ functions ranging from switches in glycolysis or cell division to building neural connectivity or modulating pain. Alterations in pre-mRNA splicing are also at the basis of a substantial number of genetic diseases, from spinal muscular atrophy, myotonic dystrophy, or familial dysautonomia to premature aging [4].

Alternatively spliced transcripts often differ in their coding capacity, stability, or translational efficiency [5]. A striking example is provided by the recent discovery of microexons encoding as few as one to five amino acids, whose inclusion is precisely regulated during neural differentiation and can modulate protein–protein interactions [6]. Alternatively spliced transcripts can also act

Trends

The spliceosome is a ribonucleoprotein complex that chaperones pre-mRNA and small nuclear (sn)RNAs into conformations resembling the RNA-based catalytic core of group II introns.

Proteins such as pre-mRNA-processing factor 8 (Prp8) organize and regulate the activation of the RNA-based catalytic core of the spliceosome.

The catalytic site is open to accommodate different intron lengths and sequences, and to interact with other cellular machineries.

Spliceosomal subcomplexes recognize and pair splice-site sequences in a highly dynamic and reversible fashion.

This architecture, together with iterative substrate recognition and proofreading, provide opportunities for alternative splicing regulation.

¹Centre de Regulació Genòmica, The Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain

²Universitat Pompeu-Fabra, Dr. Aiguader 88, 08003 Barcelona, Spain

³ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain

*Correspondence: juan.valcarcel@crg.eu (J. Valcárcel).

Figure 1. Chemical Mechanism of the Splicing Reaction. (A) RNA splicing entails two transesterification steps. In the first, the 2' hydroxyl group of an intronic adenosine residue (branch-point site, BPS) carries out a nucleophilic attack on the phosphate group between the 5' exon and the intron, generating a 2'-5' phosphodiester bond and, consequently, a lariat intermediate. In the second step, the free 3'OH of the 5' exon attacks the phosphate group between the intron and the 3' exon, splicing the two exons together and releasing the intron lariat. (B) Stabilization of transition states of the two phosphoryl-transfer reactions by Mg^{2+} . Productive transesterification relies on proper orientation and stabilization of the partial negative charges (δ^-) of the attacking and leaving groups in the transition state by two divalent magnesium ions (M1, M2), coordinated by nearby phosphates from the U6 small nuclear RNA (snRNA). In the two nucleophilic attacks, the two Mg^{2+} exchange roles in activating the attacking group and stabilizing the leaving oxyanion. After conclusion of the first step, remodeling of the spliceosome around the catalytic center shifts into place the pre-mRNA nucleotides (nt) that contribute the second step.

as drivers or modifiers of a variety of complex pathologies, including cancer, metabolic, and neurodegenerative diseases [7]. Therefore, understanding the splicing process and its selection rules is of fundamental importance for biology, medicine, and biotechnology. The possibility of harnessing splicing also holds the promise of novel gene modulation tools and therapeutic approaches [8].

The surprising organization of eukaryotic genes in introns and exons evolved in all likelihood through the invasion of mobile genetic elements known as Group II introns, which are currently present in bacterial, mitochondrial, and chloroplast genomes, but absent from eukaryotic nuclear genomes [9]. Group II introns act as RNA-based enzymes (ribozymes) that can self-splice and, therefore, eliminate themselves from the primary transcripts of their host genes. This, together with their ability to reverse splice and insert into new DNA sites, has facilitated their spread in genomes, which may have had a critical role in the origin and evolution of eukaryotes [10]. Overwhelming evidence argues for the common origin of group II and eukaryotic pre-mRNA introns. Both are removed by the same chemical mechanism, comprising two consecutive S_N2 -type transesterification reactions involving a lariat intermediate harboring a 2'-5' phosphodiester bond (Figure 1A). In addition, striking similarities exist between the 3D architecture of the RNA at the catalytic core of group II introns and at the recently reported catalytic site of a yeast spliceosome (Figure 2) [11]. A key common feature involves two Mg^{2+} ions, coordinated by equivalent RNA phosphate moieties, which activate the attacking group and stabilize the leaving group in each of the two phosphoryl-transfer steps of the splicing reaction (Figures 1B, 2C,D), a mechanism akin to phosphoryl-transfer reactions of various protein enzymes [12]. Elegant proof of the direct role of RNA in ion coordination and catalysis comes from metal rescue experiments. In these assays, substitution of phosphate by phosphorothioate at the precise sites that coordinate Mg^{2+} causes splicing defects. However, splicing can be restored by more thiophilic ions, such as Mn^{2+} or Cd^{2+} , which, in contrast to Mg^{2+} , can be coordinated by phosphorothioate moieties [13].

Trends in Biochemical Sciences

Figure 2. Structure of the RNA-Based Catalytic Core. (A) Arrangement of RNAs around the spliceosomal catalytic pocket at the conclusion of the first catalytic step. A U6 small nuclear (sn)RNA AGC triad (highlighted in orange) base pairs with U2-snRNA residues as well as with distant regions of U6-snRNA (discontinuous red lines), forming a triplex helix that brings in close proximity the nucleotides that coordinate the two Mg^{2+} at the catalytic center. Nucleotides participating in metal coordination (discontinuous black lines) are shaded darker. Positions of key nucleotides are indicated and colored following the code in part (C). Abbreviation: ISL, internal stem loop. (B) Analogous arrangements of sequences in group II introns shape a strikingly similar catalytic center. (C,D) 3D structure of the catalytic cores of the spliceosome C complex (C) and of a group II intron (D) after completion of the first step. Different nucleotide stretches in U6 and U2 snRNAs follow the same color code as in (A). Adapted from [11].

Nevertheless, a fundamental difference is that self-catalysis of group II introns is based upon the correct spatial positioning of the reactive chemical groups, precisely brought together by the elaborate 3D structure of the intron [9]. By contrast, the formidable diversity in length and sequence of pre-mRNA introns makes their removal dependent upon one of the most elaborate machineries of eukaryotic cells: the spliceosome [1]. In this review, we focus on how the molecular design of the spliceosome facilitates phosphoryl-transfer reactions by engaging the pre-mRNA splice sites of each intron with RNA moieties of the spliceosome to reconstruct the catalytic center of group II introns (Figure 2) [11]. Thus, the spliceosome can be envisioned as a highly sophisticated ribonucleoprotein chaperone that forces key sequences of the pre-mRNA into arrangements that enable the activation of an RNA catalytic center and consequent intron removal. This architecture is built from separate pieces and is subject to extensive, carefully timed compositional, and conformational remodeling that also displays surprising structural flexibility [1,14]. We argue that these features can facilitate regulation and, in multicellular organisms, allow chaperoning of alternative conformations of the same pre-mRNA to sculpt functionally different gene messages.

The Catalytic Spliceosome

The recent 3D structure of a *Schizosaccharomyces pombe* spliceosome bound to an intron lariat at 3.6-Å resolution offers important insights into splicing catalysis. It gratifyingly validates (and is validated by) a large body of exquisite inferences from 25 years of genetic, biochemical, structural, and conceptual work in multiple systems [11–13,15–20].

RNA Catalytic Core Embraced by a Protein Scaffold

After the first catalytic step, key intra- and intermolecular base pairing interactions involve three spliceosomal RNAs (U2, U5, and U6 snRNAs) and parts of the pre-mRNA (Figure 2A). The pre-mRNA is held close to the spliceosome by base pairing interactions between the 5' end of the intron and U6 snRNA, while the nucleotides flanking the branch point adenosine (but not the adenosine nucleotide itself) are base paired to U2 snRNA. These pre-mRNA–snRNA contacts, together with base pairing interactions between U2 and U6, provide the scaffold that holds the pre-mRNA onto the spliceosome and brings together the reacting groups. In addition, base pairing interactions between U2 and U6 configure the active site. A particular key set of contacts involves a triple helix between a U6 AGC triad, complementary nucleotides in U2 (forming Helix Ib), and nucleotides from distant regions of U6. As a result of this configuration, four phosphate residues are positioned such that they coordinate the two Mg^{2+} ions that facilitate phosphoryl-transfer reactions (Figures 1B, 2A–D). The structural details of the catalytic site are strikingly similar to those of the catalytic core of group II introns (Figure 2A–D), leaving little doubt about their common evolutionary origin.

The structure of this catalytic complex (known as complex C) is in fact an intermediate generated after the first step of the splicing reaction, and needs to be remodeled to allow the second reaction step. Remodeling brings the 3'-OH at the end of the 5' exon and the phosphate at the 3' end of the intron within the catalytic center (Figure 1). Genetic and molecular data indicate that the spliceosome toggles between distinct conformations associated with the two catalytic steps [21–24], with the two Mg^{2+} ions swapping roles in activating the donor and stabilizing the leaving groups for each reaction (Figure 1B) [13].

A key distinctive feature of the spliceosome is the fundamental role of proteins in configuring the ribozyme. In the ribosome, the other grand ribonucleoprotein ribozyme of the cell, RNA is responsible for the general organization of the complex (with proteins occupying peripheral positions) and single handedly configures its catalytic center. By contrast, spliceosomal proteins, in particular pre-mRNA-processing-splicing factor 8 (Prp8; Spp42 in *S. pombe*), the most conserved protein in the spliceosome, are instrumental in organizing the RNA-based catalytic site and in regulating the conformational changes needed to activate catalysis [11,15,19,23–26] (Figure 3). Interestingly, the arrangement of some domains of Prp8 resembles that of proteins encoded by certain group II introns, which facilitate RNA folding and ribozyme catalysis. This suggests that Prp8 has evolved from these RNA chaperones and accrued additional domains and protein partners to orchestrate spliceosome assembly [9,25,27]. As predicted by biochemical and structural data, the catalytic site docks on a cavity of Prp8, which stabilizes the structure (Figure 3) [11,15,19,20,25,26]. For instance, electrostatic interactions between a cluster of positively charged amino acids on the cavity surface and the internal stem loop (ISL) of U6 (Figure 2A) may stabilize an RNA conformation relevant for catalysis [11]. Determining the precise nature of the contacts contributing to the functional docking between Prp8 and the catalytic RNA core, as well as the relevance of alternative conformations [11,19,25] and their dynamics during the splicing process, can provide key insights into the catalytic capacities of ribonucleoprotein complexes. Furthermore, U5 snRNA, known to hold the first exon in place after the first catalytic step, is extensively contacted by Prp8 and other proteins, locking the extended helix and loops of U5 snRNA in a relatively rigid conformation and directing its terminal loop I to emerge close to the catalytic

Trends in Biochemical Sciences

Figure 3. Organization of Protein and RNA Moieties at the Spliceosome Center. The RNA-based catalytic center (orange circle, see Figure 2C, main text) docks in a cavity in the surface of the pre-mRNA-processing-splicing factor 8 (Prp8) protein (Spp42 in *Schizosaccharomyces pombe*). U5 small nuclear (sn)RNA stem-loops are also chaperoned by the protein into an extended conformation. Other contacts with Prp8 (Spp42) and Cwf19 (Cwf19l1 in humans) hold other RNA regions in an open configuration that helps to accommodate introns of different lengths and sequences. Adapted from [11].

site, thus providing a framework for holding the intermediates before the second step (Figure 3) [11].

An Open Catalytic Core

Prp8 and other proteins also contribute to another remarkable feature of the catalytic site. By interacting with other regions of the U2 and U6 snRNAs, these proteins ensure that these regions are kept away from the catalytic core. This open arrangement is likely to serve two distinct purposes. First, the open and peripheral location of the catalytic core makes it possible to accommodate the large variety of lengths (from <100 to >100 000 nucleotides) and sequences characteristic of pre-mRNA introns [11].

Second, we suggest that an open catalytic site is also relevant for the coupling between splicing and other gene expression steps. Recent work indicates that splicing intermediates can be identified near (within 50 nucleotides) the exit site of RNA polymerase II [28], indicating that splicing catalysis can take place shortly after the RNA emerges from the transcription complex. This implies that the machineries for RNA synthesis and splicing are very close in space and time, allowing multiple possibilities for reciprocal influences based upon interactions between (or even sharing of) their components. This is consistent with evidence that a significant fraction of splicing occurs co-transcriptionally (i.e., while the primary transcript is still associated with RNA polymerase [29–31]), and with evidence that links RNA polymerase II elongation rates with splicing efficiency [32] and splice-site selection [33–35]. Furthermore, transcription requires remodeling of chromatin architecture, including nucleosome repositioning and epigenetic modifications, and these have been also shown to influence splice-site recognition and regulation. Mechanistically, chromatin and epigenetic changes can influence splice-site choice through effects on transcription elongation or through recruitment of splicing regulators independent of transcription kinetics [35,36]. However, in contrast with these observations, a phosphorylation event in the core U2 small nuclear ribonucleoprotein (snRNP) SF3B1, which accompanies catalytic activation, is detected near interchromatin granules [37], indicating that, at least for some introns, splicing catalysis occurs post-transcriptionally.

In summary, the catalytic core of the spliceosome comprises RNA moieties that coordinate Mg^{2+} ions, which facilitate phosphoryl-transfer reactions. Proteins (in particular Prp8) establish a scaffold that stabilizes the RNA catalytic core and maintain an open arrangement that serves to accommodate different intron substrates and, possibly, facilitate the coupling between splicing and transcription.

Reconstructing the Broken Catalytic Heart

Another remarkable enzymatic feature of the spliceosome is that its catalytically active conformation is short lived and reconstructed from individual pieces for each splicing event [1]. This organization may respond to two related needs. First, intron boundaries are often located far from each other in the pre-mRNA and, therefore, the two substrates of both *trans*-esterification reactions need to be brought near each other in space and time. Efficient and accurate splice-site recognition, pairing, and regulation may be better accomplished by subcomplexes specialized in 5' and 3' splice-site identification that are subsequently bridged together, as we discuss in the next section. A second consideration for spliceosome design is that, in principle, $-OH$ and phosphate groups from any pair of nucleotides could undergo phosphoryl-transfer reactions and, therefore, it is important that the active site is generated only transiently, and only after the specific reacting groups have been identified and brought together. To achieve this, U6 snRNA is sequestered before catalysis by extensive base pairing with another spliceosomal snRNA (U4) (Figure 4A). Only after the U4:U6 structure is unwound can U6 base pair with U2 to form the catalytic site (Figure 2A).

Trends in Biochemical Sciences

Figure 4. Activation of the Catalytic Conformation. (A) Scheme of base pairing interactions between small nuclear (sn) RNAs and pre-mRNA in the precatalytic spliceosome. The 5' splice site and sequences flanking the branch-point adenosine are base paired to sequences at the 5' end of U1 and within U2 snRNAs, respectively. U6 snRNA sequences involved in catalysis, including the AGC triad, are locked by extensive base pairing with U4 snRNA. The sites of action of two DExD/H box helicases, pre-mRNA-processing-splicing factor 28 (Prp28), and Brr2, are indicated. (B) Dual regulation of Brr2 by Prp8 controls U4/6 unwinding and progression towards the catalytically active spliceosome. The helicase activity of Brr2 is inhibited by two parts of Prp8: the RNase H-like (RH) domain, which competes for binding to U4 snRNA, and the C-terminal tail of the Jab1 domain, which occludes the RNA-loading tunnel in Brr2. Unwinding of the U1 snRNA-5' splice-site interaction by Prp28 displaces Prp8 from U4 snRNA and a conformational change in its Jab1 domain not only allows U4 to enter the RNA tunnel in Brr2, but also enhances coupling between ATP hydrolysis and duplex unwinding, thus facilitating the disruption of the U4:U6 duplex and spliceosome activation.

Many conformational changes in the spliceosome, including this key transition, involve the rearrangement of base pairing interactions promoted by RNA helicases of the DExD/H box family [38]. The Brr2 helicase associates with U4 snRNA just 3' of the region of base pairing with U6. Recent results reveal a detailed mechanism for Brr2 regulation during spliceosome assembly, again involving a key orchestrating role for Prp8 [24,39–41] (Figure 4B). Before spliceosome activation, Brr2 is inhibited by a double-lock mechanism, whereby the RNase H-like domain of Prp8 competes with Brr2 for binding to U4 snRNA, while the C-terminal tail of the Prp8 Jab-1 domain occludes a tunnel in Brr2 that is occupied by U4 snRNA when the enzyme is actively involved in unwinding [39–41]. Remarkably, to achieve Brr2 activation, Prp8 not only releases this double repression, but also enhances Brr2 efficiency in coupling ATP hydrolysis with duplex unwinding (Figure 4B).

A striking observation is that particularly acute forms of retinitis pigmentosa, a retinal degeneration disease, are linked to mutations in the C-terminal tail of the Jab1 domain of Prp8, which prevent efficient Brr2 inhibition [41,42]. The reasons why mutations in a core component of the spliceosome, likely leading to premature U4/U6 unwinding, cause tissue-specific defects remain unclear. One possibility is that reduced splicing activity is particularly harmful for highly regenerative tissues that require abundant cell proliferation and an increased rate of transcription and RNA processing. This is consistent with the observation that the same mutations cause cell growth defects when engineered in yeast [43].

These and other observations suggest that interference with spliceosome function could also provide opportunities to limit the proliferation of cancer cells. Indeed, several families of natural products displaying antitumor properties target core splicing components [8] and the spliceosome appears to be a target of oncogenic stress in MYC-driven cancers [44,45]. More generally, alterations in alternative splicing contribute to the key hallmarks of cancer progression [8,46].

Identifying the Correct Phosphates for Splicing

In the yeast *Saccharomyces cerevisiae*, only 5% of protein-coding genes contain introns, typically a single one of relatively short length, defined by strongly conserved sequences at the exon–intron boundaries. By contrast, splice-site sequences in multicellular organisms are highly degenerate, despite the fact that their spliceosomes are confronted with the more daunting task of identifying multiple introns in each gene, commonly orders of magnitude longer than the flanking exons. For instance, it has been proposed that thousands of different sequences can function as 5' splice sites or as branch points, with only a few nucleotides being invariant [47,48]. Furthermore, sequences resembling consensus splice sites are often not used, and minimal sequence changes (e.g., mutations) can lead to cryptic splice-site activation, often with pathological consequences [4]. In addition to the splice sites, other sequences influence splice-site recognition [49]. Therefore, to develop a splicing code that predicts splice-site utilization under physiological or pathological conditions, hundreds of regulatory motifs may need to be integrated [50,51].

Before base pairing with U6 snRNA, 5' splice sites are initially recognized by U1 snRNP [1]. As is the case for the other small nuclear ribonucleoprotein particles, U1 comprises a snRNA (U1 snRNA) and proteins, a subset of which (the Sm core) is common to most spliceosomal snRNPs [1]. snRNP biogenesis involves multiple steps of RNA modification and protein assembly that occur in both the nucleus and the cytoplasm, and defects in this process are at the basis of spinal muscular atrophy, one of the most prevalent genetic diseases [52,53]. The recent determination of the structure of U1 snRNP bound to a 5' splice site at 3.3-Å resolution illustrates the principles of recognition and is consistent with a large body of molecular and genetic data (Figure 5A) [54]. Base pairing between the 5' end of U1 snRNA and the first six nucleotides of the intron, and up to three nucleotides of the exon, provides the main driving force for specific sequence recognition. The possibility of using different base pairing registers and of accommodating nucleotides bulged out from the base paired

Trends in Biochemical Sciences

Figure 5. Mechanisms of Early Splice-Site Recognition. (A) Structure of U1 small nuclear ribonucleic protein (snRNP) bound to a 5' splice-site RNA. Base pairing between the 5' end of U1 snRNA and the 5' splice-site sequence is assisted by contacts of the RNA backbone by the U1-C protein. Assembly of U1-C is facilitated by U1-70K, another U1-specific protein, whose binding depends, in turn, on the full assembly of complex of seven Sm proteins. Adapted from [54]. (A). (B,C) Tentative model for the initial sequential recognition of the 3' splice-site region. Cooperative recognition of the branch-point sequence by the KH domain of branch-point binding protein/splicing factor 1 (BBP/SF1), of the pyrimidine-rich tract by two RRM domains of U2AF65 and the AG dinucleotide at the 3' end of the intron by U2AF35 is facilitated by interactions between these proteins (B). U2 snRNP is subsequently assembled through a network of RNA:RNA [U2 snRNA:nucleotides flanking the branch-point base pairing, facilitated by U2AF65 arginine-serine (RS) domain], protein-protein (U2AF65 UHM domain/SF3B1) and RNA-protein (SF3B1/pre-mRNA) interactions.

region could allow the recognition of a variety of sequences [47,55]. Stabilization of such mismatches is promoted by hydrogen bonds and electrostatic contacts between a protein component of U1 (U1-C) and the RNA backbone at the 5' splice-site region [54]. Furthermore, regulatory factors can bind neighboring sequences to prevent or promote U1 snRNP binding by either sterically blocking snRNP assembly or stabilizing it, respectively [56]. For example, the T-cell-restricted intracellular antigen proteins (TIA1/TIAR) bind uridine-rich sequences 3' of the 5' splice site and their interaction with U1-C can promote U1 snRNP recruitment [57,58]. However, binding of U1 snRNP may not be the only determinant of 5' splice-site selection, because differences in the kinetics of splice-site pairing can also influence the selection between competing 5' splice sites [56,58,59].

After the 5' splice site is recognized by U1 snRNP, another DExD/H-box helicase, Prp28, disrupts the U1 snRNA-5' splice-site interaction to replace U1 with U6 [60]. This event is relevant for catalytic activation because base pairing between U6 and the 5' splice site allows (and may be assisted by) interactions between the RNase H-like domain of Prp8 and the 5' splice site [61,62]. This, in turn, releases Prp8 from its binding site in U4 snRNA and, thus, permits Brr2 function in U4:U6 unwinding and formation of the catalytic site (Figures 2, 4) [41]. This example

nicely illustrates how the ordered operation of DExD/H-box helicases and other activities [e.g., by arginine-serine (RS) domain-containing splicing factors [63]] can trigger RNA:RNA base pairing transitions that, in turn, mediate splice-site recognition, spliceosome assembly, and eventually catalytic activation and release of the spliced products from the spliceosome [38].

Similar to 5' splice-site recognition, 3' splice-site recognition also depends critically on base pairing interactions between pre-mRNA and snRNAs. In this case, U2 snRNA interacts with sequences flanking the branch point adenosine, which is bulged out as a necessary signal for catalysis [64]. Similar to 5' splice-site recognition, non-sequence-specific contacts between the protein components of U2 snRNP and the pre-mRNA help stabilize binding [65,66]. However, U2 snRNP recruitment is promoted by previous recognition of the branch site (including the branch adenosine) by branch-point binding protein/splicing factor 1 (BBP/SF1) and, in higher eukaryotes, by the heterodimer U2AF recognizing a pyrimidine-rich tract and the conserved AG dinucleotide at the 3' end of the intron (Figures 5B,C) [1]. Cooperative recognition of these adjacent sequences by the three proteins, along with conformational changes induced by RNA binding and protein–protein interactions with U2 snRNP components, help to define 3' splice-site regions despite their variable sequence arrangements [67–71]. Conformational switches also need to occur in U2 snRNA, involving the function of the DExD/H-box helicases Prp5 [72–74] and Prp16 [75]. Intriguingly, Prp5 function is also required for elongation of transcription complexes over intronic sequences, revealing a splicing-dependent transcriptional checkpoint associated with proper spliceosome assembly [76] and further highlighting the tight links between these two processes. Prp16 is involved in both catalytic steps by proofreading branch-point recognition and 5' splice-site cleavage, thus exerting multiple controls along the splicing process [38,77,78]. An important role of several of these DExD/H-box helicases is to provide a kinetic proofreading mechanism for key transitions using the rate of ATP hydrolysis as a molecular timer: if the transitions do not occur within this window of opportunity, the process is reversed or halted and the complexes discarded.

Stabilizing interactions between factors recognizing the 5' and 3' splice sites can occur either across the intron (intron definition) or across internal exons (exon definition) (Figure 6A), which in vertebrates are much shorter than the flanking introns. The exon definition complexes can be remarkably similar to those formed across introns, because they contain U4/6/5 tri-snRNP and are directly convertible into precatalytic complexes [79].

In summary, base pairing between snRNAs and splice sites, as well as other RNA–protein and protein–protein interactions, together define the correct phosphates that will undergo the splicing reaction. Multiple recognition and proofreading steps ensure precise splice-site identification despite the diversity of sequences present at intron–exon boundaries in metazoans.

Malleable Machinery

An emerging concept is that the spliceosome is a remarkably dynamic and flexible molecular machine [1,14,80]; its transitions are so malleable that the whole reaction can eventually be reversed to generate pre-mRNA from spliced products [81]. Directionality is fueled by helicase-driven ATP hydrolysis and by myriad other sequential activities that introduce post-translational protein modifications. Progression through the splicing reaction is accompanied by extensive remodeling of the protein composition of the subcomplexes associated with the different snRNA–pre-mRNA structural transitions (Figure 6) [1,82,83]. For example, catalytic activation is associated with the incorporation of approximately 35 proteins in splicing complexes, while a similar number dissociate from them. Few of the more than 150 proteins identified in purified spliceosomal complexes persist during the whole process [1,82].

Alternative assembly routes also exist, as illustrated by single molecule-imaging studies in yeast extracts, which revealed that association of every major spliceosomal subcomplex is reversible

Trends in Biochemical Sciences

Figure 6. Overview of Complexes Involved in Spliceosome Assembly and Catalysis. (A) Composition of U small nuclear ribonucleic proteins (snRNPs) (U1, U2, U4, U5, and U6) and additional select factors are indicated for each complex. Dark arrows indicate documented transitions between complexes; light arrows indicate transitions that likely exist. Alternative assembly pathways and reversible steps illustrate the highly dynamic and flexible architecture of splicing complexes. (B) Biochemical features of spliceosome assembly subcomplexes. The approximate number of different protein species (+, newly associated proteins; -, dissociated proteins) and DEAD/H box helicases involved in formation of each complex are indicated. Abbreviations: NTC, nineteen complex; Prp, processing-splicing factor; SR, SR proteins. Adapted from [1] (B).

and that the order of splice-site recognition depends on the pre-mRNA; that is, the substrate influences the assembly pathway [84–86] (Figure 6A). One example of this is assembly of the Prp19-associated multiprotein nineteen complex (NTC) [87]. The complex can be recruited at the final step of spliceosome assembly or through an interaction with U2AF65; the latter recruitment is stimulated co-transcriptionally by the phosphorylated carboxy terminal domain (CTD) of RNA polymerase II [88]. Once recruited, NTC has key roles in catalytic activation. Remarkably, the NTC complex, which has ubiquitin ligase activities, also has important roles in transcription elongation, mRNA export, and the DNA damage response, once again highlighting the extensive interplay between nuclear processes [89].

The levels of splicing factors can also have a role in the choice of assembly pathway. Thus, an excess of RS domain-containing splicing factors (SR proteins) can promote 5' splice-site recognition in the absence of U1 snRNP [90]. Adding to the concept of molecular malleability, spliceosomal proteins contain a significantly higher fraction of intrinsically disordered regions compared with the rest of the eukaryotic proteome (this fraction being even higher in peripheral components acting at earlier stages of assembly) [91,92]. These flexible regions may help to adapt spliceosomal complexes to different substrates or alternative assembly pathways and favor variable schemes of coupling with other processes or cellular status.

The flexibility and reversibility in assembly (or even catalysis) has important implications for splicing regulation, which could in principle exploit the full range of sophisticated transactions in spliceosome function to modulate splice-site choice [14,80]. In fact, genome-wide screens and functional network reconstruction reveal that variations in the activity of core splicing factors along the whole splicing pathway can result in pre-mRNA-specific regulation of splicing efficiency and alternative splicing [93–97].

Concluding Remarks

Introns likely originated from mobile genetic elements whose ability to self-excite from primary transcripts allowed them to spread through genomes without compromising the function of their host genes [9]. One view is that the progressive dependence of intron removal on auxiliary factors represents a purposeless evolutionary ratchet-trap that irreversibly renders increasingly complex spliceosomes [98]. A complementary perspective is that the spread of introns and the increasing sophistication of the machinery involved in their removal have been co-opted by cells to acquire unique opportunities to expand the information density of their genomes and to enrich the regulatory possibilities of their gene expression programs. Thus, each new intron or spliceosomal component may offer evolutionary playgrounds to explore novel mRNA arrangements, broadening the functional and regulatory repertoire of the organism. The highly modular and dynamic composition of the spliceosome, together with its malleable and even reversible function, allow alternative sequence constellations to work as splice sites and be sensitive to regulatory pathways, while maintaining the high precision required to generate functional mRNAs. Understanding the staggering complexity of regulatory sequences, factors, and their fleeting transactions remains a challenge for the next generation of splicing aficionados, which will likely require merging structural, biochemical, imaging, genomic, and computational approaches. These efforts have the potential to reveal the evolutionary and physiological logic behind gene-regulation programs, potentially allowing us to harness splicing for investigating gene function, and possibly correct splicing alterations in genetic and other diseases, including neurodegenerative disorders and cancer.

Acknowledgments

We sincerely apologize to many colleagues whose work could not be cited because of space constraints. We thank Roderic Guigó, Alberto Kornblihtt, Reinhard Lührmann, Michael Sattler, Josep Vilardell, and members of our group for comments on the manuscript, and Kiyoshi Nagai and Yigong Shi for kindly agreeing on the reproduction of images from published work.

Outstanding Questions

What is the function of the numerous protein components of the spliceosome that are characteristic of higher eukaryotes? Do they accommodate variation and regulation of splice-site sequences?

How do proteins contribute to remodeling of the catalytic site between the two chemical steps, and how do they chaperone transitions similar to those of group II introns?

How do intrinsically disordered regions contribute to the dynamics and plasticity of spliceosome assembly? How are spliceosomal transitions triggered and modulated by post-translational modifications?

Do modifications in pre-mRNA and snRNAs (e.g., methylation or pseudouridylation) have a role in splice-site recognition and modulation?

What are the mechanisms behind the pairing of splice sites across introns and exons? How do exon definition complexes transit to intron definition complexes?

Is there a simple mechanistic logic for the splicing code and for the effects of the numerous *trans*-acting factors, including RNA-packaging proteins, splicing regulators, core splicing factors, and components of the transcription, pre-mRNA decay, and mRNA transport machineries?

Why have the splice sites of introns processed by the minor (U11,U12) spliceosome not diverged in sequence as much as those processed by the major spliceosome?

What is the molecular basis for the seemingly target-specific effects of small compounds that target the spliceosome?

How does co-transcriptional spliceosome assembly occur under physiological conditions in different pre-mRNAs? How does the pathway vary in different cell types? Does it vary between individual molecules and in individual cells?

How do chromatin and DNA modifications influence splice-site recognition and what is the basis for the reciprocal effects of splicing on transcription and chromatin? Does the 3D organization of chromatin influence splice-site selection?

Work in our laboratory is supported by Fundación Botín, Banco de Santander through its Santander Universities Global Division, Consolider RNAREG, CellSyst, AGAUR, and the European Research Council. We acknowledge support of the Spanish Ministry of Economy and Competitiveness, 'Centro de Excelencia Severo Ochoa 2013-2017', SEV-2012-0208.

References

- Wahl, M.C. *et al.* (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–718
- Keren, H. *et al.* (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11, 345–355
- Braunschweig, U. *et al.* (2013) Dynamic integration of splicing within gene regulatory pathways. *Cell* 152, 1252–1269
- Ward, A.J. and Cooper, T.A. (2010) The pathobiology of splicing. *J. Pathol.* 220, 152–163
- Kelemen, O. *et al.* (2013) Function of alternative splicing. *Gene* 514, 1–30
- Irimia, M. *et al.* (2014) A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* 159, 1511–1523
- Singh, R.K. and Cooper, T.A. (2012) Pre-mRNA splicing in disease and therapeutics. *Trends Mol. Med.* 18, 472–482
- Bonnal, S. *et al.* (2012) The spliceosome as a target of novel antitumour drugs. *Nat. Rev. Drug Discov.* 11, 847–859
- Lambowitz, A.M. and Zimmerly, S. (2011) Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb. Perspect. Biol.* 3, a003616
- Martin, W. and Koonin, E.V. (2006) Introns and the origin of nucleus-cytosol compartmentalization. *Nature* 440, 41–45
- Hang, J. *et al.* (2015) Structural basis of pre-mRNA splicing. *Science* 349, 1191–1198
- Steitz, T.A. and Steitz, J.A. (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U.S.A.* 90, 6498–6502
- Fica, S.M. *et al.* (2013) RNA catalyses nuclear pre-mRNA splicing. *Nature* 503, 229–234
- Chen, W. and Moore, M.J. (2014) The spliceosome: disorder and dynamics defined. *Curr. Opin. Struct. Biol.* 24, 141–149
- Yan, C. *et al.* (2015) Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* 349, 1182–1191
- Madhani, H.D. and Guthrie, C. (1992) A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* 71, 803–817
- Wassarman, D.A. and Steitz, J.A. (1992) Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. *Science* 257, 1918–1925
- Fica, S.M. *et al.* (2014) Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nat. Struct. Mol. Biol.* 21, 464–471
- Montemayor, E.J. *et al.* (2014) Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. *Nat. Struct. Mol. Biol.* 21, 544–551
- Anokhina, M. *et al.* (2013) RNA structure analysis of human spliceosomes reveals a compact 3D arrangement of snRNAs at the catalytic core. *EMBO J.* 32, 2804–2818
- Moore, M.J. and Sharp, P.A. (1993) Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* 365, 364–368
- Konarska, M.M. *et al.* (2006) Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol. Cell* 21, 543–553
- Liu, L. *et al.* (2007) Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. *Nat. Struct. Mol. Biol.* 14, 519–526
- Hahn, D. *et al.* (2012) Brp2p-mediated conformational rearrangements in the spliceosome during activation and substrate repositioning. *Genes Dev.* 26, 2408–2421
- Galej, W.P. *et al.* (2013) Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature* 493, 638–643
- Nguyen, T.H. *et al.* (2015) The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* 523, 47–52
- Diakic, M. and Mushegian, A. (2011) Prp8, the pivotal protein of the spliceosomal catalytic center, evolved from a retroelement-encoded reverse transcriptase. *RNA* 17, 799–808
- Nojima, T. *et al.* (2015) Mammalian NET-Seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* 161, 526–540
- Gornemann, J. *et al.* (2005) Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol. Cell* 19, 53–63
- Tilgner, H. *et al.* (2009) Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* 16, 996–1001
- Khodor, Y.L. *et al.* (2011) Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev.* 25, 2502–2512
- Braberg, H. *et al.* (2013) From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. *Cell* 154, 775–788
- de la Mata, M. *et al.* (2003) A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* 12, 525–532
- Lin, S. *et al.* (2008) The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* 15, 819–826
- Naftelberg, S. *et al.* (2015) Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annu. Rev. Biochem.* 84, 165–198
- Luco, R.F. *et al.* (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* 144, 16–26
- Girard, C. *et al.* (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun.* 3, 994
- Liu, Y.C. and Cheng, S.C. (2015) Functional roles of DExD/H-box RNA helicases in Pre-mRNA splicing. *J. Biomed. Sci.* 22, 54
- Mozaffari-Jovin, S. *et al.* (2012) The Prp8 RNase H-like domain inhibits Brp2-mediated U4/U6 snRNA unwinding by blocking Brp2 loading onto the U4 snRNA. *Genes Dev.* 26, 2422–2434
- Mozaffari-Jovin, S. *et al.* (2013) Inhibition of RNA helicase Brp2 by the C-terminal tail of the spliceosomal protein Prp8. *Science* 341, 80–84
- Mozaffari-Jovin, S. *et al.* (2014) Novel regulatory principles of the spliceosomal Brp2 RNA helicase and links to retinal disease in humans. *RNA Biol.* 11, 298–312
- McKie, A.B. *et al.* (2001) Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Hum. Mol. Genet.* 10, 1555–1562
- Towns, K.V. *et al.* (2010) Prognosis for splicing factor PRPF8 retinitis pigmentosa, novel mutations and correlation between human and yeast phenotypes. *Hum. Mutat.* 31, E1361–E1376
- Hsu, T.Y. *et al.* (2015) The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* 525, 384–388
- Koh, C.M. *et al.* (2015) MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. *Nature* 523, 96–100
- Oltean, S. and Bates, D.O. (2014) Hallmarks of alternative splicing in cancer. *Oncogene* 33, 5311–5318
- Roca, X. and Krainer, A.R. (2009) Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA. *Nat. Struct. Mol. Biol.* 16, 176–182
- Mercer, T.R. *et al.* (2015) Genome-wide discovery of human splicing branchpoints. *Genome Res.* 25, 290–303
- Wang, Z. and Burge, C.B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* 14, 802–813
- Barash, Y. *et al.* (2010) Deciphering the splicing code. *Nature* 465, 53–59
- Xiong, H.Y. *et al.* (2015) RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* 347, 1254806
- Zhang, Z. *et al.* (2008) SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* 133, 585–600

Can ncRNAs, including endogenously produced small interfering RNAs (endo-siRNAs), Piwi-interacting RNA (piRNAs), and lncRNAs, influence splice-site recognition and selection?

53. Matera, A.G. and Wang, Z. (2014) A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* 15, 108–121
54. Kondo, Y. *et al.* (2015) Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *Elife* 4, 04986
55. Roca, X. *et al.* (2012) Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev.* 26, 1098–1109
56. Roca, X. *et al.* (2013) Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev.* 27, 129–144
57. Förch, P. *et al.* (2002) The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites. *EMBO J.* 21, 6882–6892
58. Wang, Z. *et al.* (2010) iCLIP predicts the dual splicing effects of TIA-RNA interactions. *PLoS Biol.* 8, e1000530
59. Yu, Y. *et al.* (2008) Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition. *Cell* 135, 1224–1236
60. Staley, J.P. and Guthrie, C. (1999) An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol. Cell* 3, 55–64
61. Reyes, J.L. *et al.* (1999) The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5' splice site. *RNA* 5, 167–179
62. Pena, V. *et al.* (2008) Structure and function of an RNase H domain at the heart of the spliceosome. *EMBO J.* 27, 2929–2940
63. Shen, H. and Green, M.R. (2006) RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes Dev.* 20, 1755–1765
64. Smith, D.J. *et al.* (2009) Insights into branch nucleophile positioning and activation from an orthogonal pre-mRNA splicing system in yeast. *Mol. Cell* 34, 333–343
65. Gozani, O. *et al.* (1996) Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev.* 10, 233–243
66. Corionero, A. *et al.* (2011) Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes Dev.* 25, 445–459
67. Sickmier, E.A. *et al.* (2006) Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. *Mol. Cell* 23, 49–59
68. Mackereth, C.D. *et al.* (2011) Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* 475, 408–411
69. Jenkins, J.L. *et al.* (2013) U2AF65 adapts to diverse pre-mRNA splice sites through conformational selection of specific and promiscuous RNA recognition motifs. *Nucleic Acids Res.* 41, 3859–3873
70. Thickman, K.R. *et al.* (2007) Alternative conformations at the RNA-binding surface of the N-terminal U2AF(65) RNA recognition motif. *J. Mol. Biol.* 366, 703–710
71. Corsini, L. *et al.* (2007) U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. *Nat. Struct. Mol. Biol.* 14, 620–629
72. Perriman, R.J. and Ares, M., Jr (2007) Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. *Genes Dev.* 21, 811–820
73. Perriman, R. and Ares, M., Jr (2010) Invariant U2 snRNA nucleotides form a stem loop to recognize the intron early in splicing. *Mol. Cell* 38, 416–427
74. Liang, W.W. and Cheng, S.C. (2015) A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes Dev.* 29, 81–93
75. Hilliker, A.K. *et al.* (2007) U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. *Genes Dev.* 21, 821–834
76. Chathoth, K.T. *et al.* (2014) A splicing-dependent transcriptional checkpoint associated with prespliceosome formation. *Mol. Cell* 53, 779–790
77. Koodathingal, P. *et al.* (2010) The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Mol. Cell* 39, 385–395
78. Tseng, C.K. *et al.* (2011) DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* 17, 145–154
79. Schneider, M. *et al.* (2010) Exon definition complexes contain the tri-snRNP and can be directly converted into B-like pre-catalytic splicing complexes. *Mol. Cell* 38, 223–235
80. Hoskins, A.A. and Moore, M.J. (2012) The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* 37, 179–188
81. Tseng, C.K. and Cheng, S.C. (2008) Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* 320, 1782–1784
82. Wahl, M.C. and Luhrmann, R. (2015) SnapShot: spliceosome dynamics I. *Cell* 161, 1474–e1
83. Wahl, M.C. and Luhrmann, R. (2015) SnapShot: spliceosome dynamics II. *Cell* 162, 456–e1
84. Abelson, J. *et al.* (2010) Conformational dynamics of single pre-mRNA molecules during in vitro splicing. *Nat. Struct. Mol. Biol.* 17, 504–512
85. Hoskins, A.A. *et al.* (2011) Ordered and dynamic assembly of single spliceosomes. *Science* 331, 1289–1295
86. Shcherbakova, I. *et al.* (2013) Alternative spliceosome assembly pathways revealed by single-molecule fluorescence microscopy. *Cell Rep.* 5, 151–165
87. Chan, S.P. *et al.* (2003) The Prp19p-associated complex in spliceosome activation. *Science* 302, 279–282
88. David, C.J. *et al.* (2011) The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.* 25, 972–983
89. Chanarat, S. and Strasser, K. (2013) Splicing and beyond: the many faces of the Prp19 complex. *Biochim. Biophys. Acta* 1833, 2126–2134
90. Crispino, J.D. *et al.* (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science* 265, 1866–1869
91. Korneta, I. and Bujnicki, J.M. (2012) Intrinsic disorder in the human spliceosomal proteome. *PLoS Comput. Biol.* 8, e1002641
92. Coelho Ribeiro Mde, L. *et al.* (2013) Malleable ribonucleoprotein machine: protein intrinsic disorder in the *Saccharomyces cerevisiae* spliceosome. *PeerJ* 1, e2
93. Clark, T.A. *et al.* (2002) Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* 296, 907–910
94. Pleiss, J.A. *et al.* (2007) Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol.* 5, e90
95. Park, J.W. and Graveley, B.R. (2005) Use of RNA interference to dissect the roles of trans-acting factors in alternative pre-mRNA splicing. *Methods* 37, 341–344
96. Tejedor, J.R. *et al.* (2015) Genome-wide identification of Fas/CD95 alternative splicing regulators reveals links with iron homeostasis. *Mol. Cell* 57, 23–38
97. Papasaikas, P. *et al.* (2015) Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol. Cell* 57, 7–22
98. Doolittle, F. and Gitschier, J. (2015) The philosophical approach: an interview with ford doolittle. *PLoS Genet.* 11, e1005173