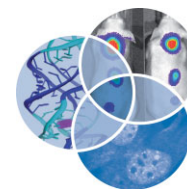


## Advanced Review



# The significant other: splicing by the minor spliceosome

Janne J. Turunen,<sup>†</sup> Elina H. Niemelä, Bhupendra Verma and Mikko J. Frilander\*

The removal of non-coding sequences, introns, from the mRNA precursors is an essential step in eukaryotic gene expression. U12-type introns are a minor subgroup of introns, distinct from the major or U2-type introns. U12-type introns are present in most eukaryotes but only account for less than 0.5% of all introns in any given genome. They are processed by a specific U12-dependent spliceosome, which is similar to, but distinct from, the major spliceosome. U12-type introns are spliced somewhat less efficiently than the major introns, and it is believed that this limits the expression of the genes containing such introns. Recent findings on the role of U12-dependent splicing in development and human disease have shown that it can also affect multiple cellular processes not directly related to the functions of the host genes of U12-type introns. At the same time, advances in understanding the regulation and phylogenetic distribution of the minor spliceosome are starting to shed light on how the U12-type introns and the minor spliceosome may have evolved. © 2012 John Wiley & Sons, Ltd.

**How to cite this article:**

*WIREs RNA* 2013, 4:61–76. doi: 10.1002/wrna.1141

## INTRODUCTION

U12-type introns were initially described as a handful of unusual introns containing non-consensus AT–AC termini and a high degree of conservation at the 5' splice site (5'ss).<sup>1</sup> These characteristics set them apart from most other introns that typically had GT–AG termini and relatively variable sequences at the 5'ss. However, the true significance of this finding did not become evident until a provocative hypothesis by Hall and Padgett,<sup>2</sup> who suggested that such introns may be recognized by the factors that are specific to this type of introns. This hypothesis set the stage for the discovery that the genomes of most eukaryotes actually harbor two different types of introns, termed U2 and U12 type, that are removed by two separate spliceosomes.

The significance of having two parallel intron types and two machineries for their removal was

initially perplexing, especially since the U12-type introns are present at very low frequencies. The function and significance of the U12-dependent spliceosome are still far from being fully understood, but there are several lines of evidence indicating that U12-type splicing has essential functions. These are related to the conservation of U12-type introns in distantly related organisms and their presence in specific types of genes, as well as to the slow kinetics of their removal, which may regulate the expression of the genes containing them. Here, we review the current knowledge on U12-type introns and the function of the U12-dependent spliceosome. In particular, we discuss the recent findings on the regulation of the U12-dependent splicing, its function in development and disease, and the impact of these findings on understanding the evolution of the two spliceosomes.

## U12-TYPE INTRON SEQUENCES: DEFINITION AND CHARACTERISTICS

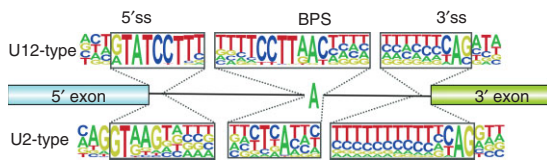
Although the U12-type introns were first identified on the basis of terminal AT–AC dinucleotides (therefore initially referred to as AT–AC or atac introns) it was soon realized that these termini are not exclusively present in U12-type introns.<sup>3</sup> Rather, the defining

<sup>†</sup>Present address: Department of Laboratory Medicine, Karolinska Institutet, Stockholm, Sweden

\*Correspondence to: Mikko.Frilander@Helsinki.Fi

Institute of Biotechnology, University of Helsinki, Helsinki, Finland

Re-use of this article is permitted in accordance with the Terms and Conditions set out at [http://wileyonlinelibrary.com/onlineopen#OnlineOpen\\_Terms](http://wileyonlinelibrary.com/onlineopen#OnlineOpen_Terms)



**FIGURE 1** | Consensus sequences of human U12- and U2-type introns. The height of the letters in each position indicates the relative frequency of individual nucleotides in that position. For frequency calculation the U12-type splice site sequences were obtained from U12DB,<sup>7</sup> and the corresponding 5'ss and 3'ss frequencies for U2-type introns from the Splice Rack database (Ref 8; <http://kathadhin.mssm.edu/splice/index.cgi?database=spliceNew2>) and the branch point sequences (BPS) data from Gao et al.,<sup>9</sup> The sequence logos were generated using the enoLOGOS web server.<sup>10</sup>

features of U12-type introns are the 5'ss and branch point sequences (BPS), which are more conserved than in most other introns<sup>4</sup> (Figure 1). GT-AG were also found to function as the terminal dinucleotides, and are actually the more common subtype.<sup>5</sup> Thus, a nomenclature was adopted in which the original spliceosomal introns are referred to as U2-type introns and the novel introns as U12-type introns. The introns are spliced by the U2- or U12-dependent spliceosomes, and are also often referred to as the 'major' and 'minor' spliceosomes, respectively.<sup>5,6</sup>

Naturally occurring U12-type introns mostly belong to either the AT-AC or GT-AG subtypes. A few natural examples of U12-type introns with other combinations of terminal residues have been reported, and non-canonical termini have also been shown to support U12-dependent splicing experimentally, although often with reduced efficiency.<sup>11-15</sup> However, combinations of the main subtypes (i.e., AT-AG or GT-AC) appear to be disfavored. In addition to their distinct 5'ss and BPS, U12-type introns are also characterized by the lack of a distinct polypyrimidine tract (PPT) that is typically found upstream of U2-type 3' splice sites. The distance from a U12-type BPS to the 3'ss has been shown to be an important factor for the recognition of U12-type introns, and is significantly shorter than in U2-type introns, with an optimum distance of 11–13 nt.<sup>11-13,16,17</sup>

## DISTRIBUTION OF U12-TYPE INTRONS

Currently, U12-type introns have been identified in all major eukaryotic taxa, including plants, fungi, and animals, as well as a few deep-branching single-celled eukaryotes, but are nonetheless absent in many species, including such common model organisms as *Caenorhabditis elegans* and *Saccharomyces cerevisiae*<sup>4,8,13,16,18,19</sup> (Figure 2). The putative U12-type introns can readily be recognized computationally on the basis of the 5'ss and the BPS that are conserved

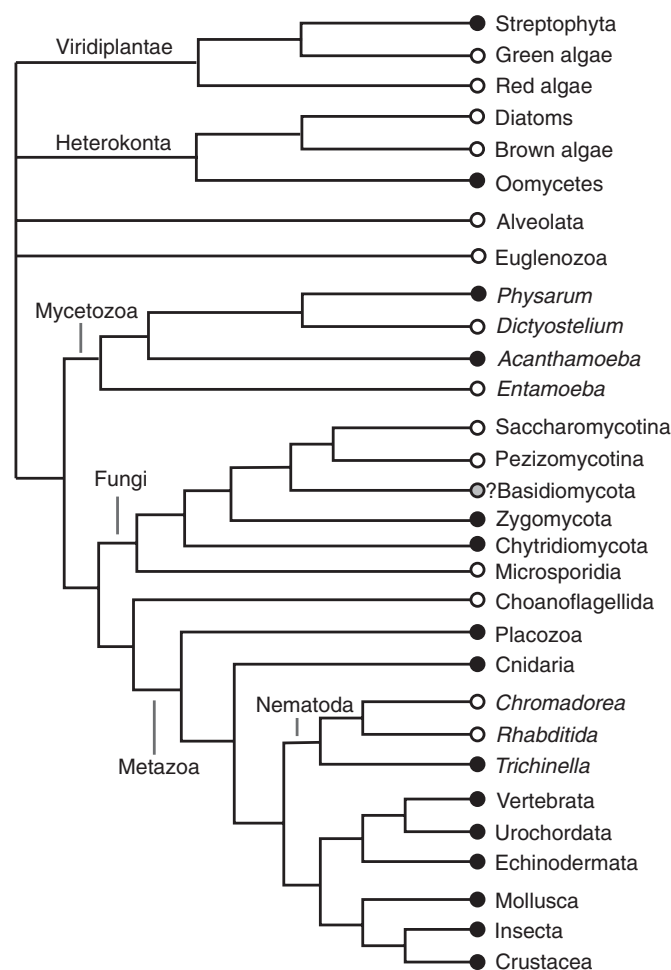
among different organisms. An examination of the U12-type introns in the Splice Rack database (Ref 8; see Figure 1) revealed approximately 700–800 putative genes carrying U12-type introns in human and mouse, and ca 300 in *Arabidopsis thaliana*, while in *Drosophila melanogaster* only 19 U12-type introns have been described.<sup>20,21</sup> However, these are estimates based on relatively stringent bioinformatic criteria, as there is little experimental evidence on the recognition of weak U12-type splice sites. Therefore, with less stringent criteria the number of putative U12-type introns could be significantly larger as suggested recently in *A. thaliana*.<sup>22</sup>

The genes containing U12-type introns are enriched in certain functional classes and pathways as described originally by Burge et al.<sup>4</sup> They are mainly present in genes related to 'information processing functions', such as DNA replication and repair, transcription, RNA processing, and translation, but can also be found in genes related to cytoskeletal organization, vesicular transport, and voltage-gated ion channel activity. In contrast, U12-type introns are almost absent in genes related to basic energy metabolism and biosynthetic pathways.<sup>4,24,25</sup> Typically only one U12-type intron is present in an individual gene. However, in humans, there are approximately 50 genes containing two U12-type introns and a few cases with three U12-type introns. This is most common in genes belonging to the voltage-gated ion channel superfamily.<sup>24,25</sup> Together, these observations have led to the suggestion that U12-type introns may have a role in regulating the expression of specific sets of genes.

## SPLICEOSOME ASSEMBLY AND CATALYSIS

### The Major Spliceosome

Introns are recognized and excised by a large molecular machine called the spliceosome that is composed of five small nuclear ribonucleoproteins (snRNPs) and additional non-snRNP protein components.<sup>26</sup> Each of the five snRNPs contains one small nuclear RNA (snRNA) and a number of protein components. The U2-type snRNPs are U1, U2, U4, U5, and U6, the latter three of which form a U4/U6.U5 tri-snRNP. The spliceosome is formed by the sequential interactions of the snRNAs and spliceosomal proteins with the pre-mRNA substrate and with one another. In U2-type introns, the 5'ss is initially recognized by the U1 snRNP, while the BPS, PPT, and 3'ss are recognized by the protein factors SF1, U2AF<sup>65</sup> and U2AF<sup>35</sup>, respectively, together forming the spliceosomal commitment (or E) complex. During



**FIGURE 2 |** Phylogenetic distribution of U12-type introns and splicing factors, with special emphasis on the recurrent loss of U12-dependent splicing in the fungi-metazoa lineage. Filled circles in the schematic tree indicate taxa in which U12-type introns and/or splicing factors have been identified, while open circles indicate taxa in which they have not been observed. Names of individual genera have been given in italics. The lengths of the branches do not indicate true phylogenetic distances. The tree is based on the data obtained from Refs 19 and 23.

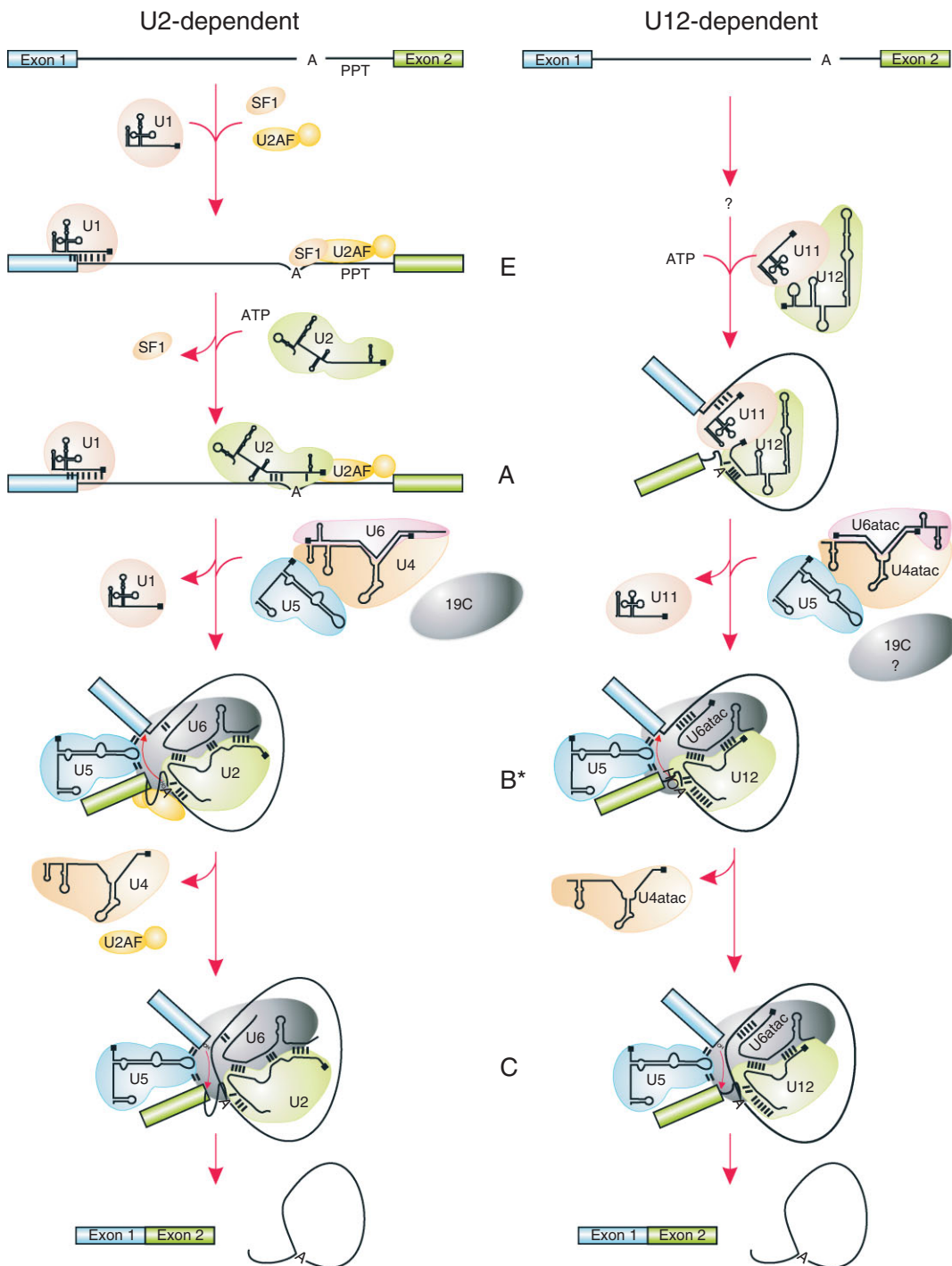
the formation of the pre-spliceosome, or A complex, U2 snRNP replaces SF1 at the BPS. At later stages, the U4/U6.U5 tri-snRNP stably associates with the spliceosome (B complex). Rearrangements in RNA and protein interactions lead to the formation of the catalytically active spliceosome (B\* complex) that catalyzes the first transesterification reaction, followed by the second transesterification in the C complex (Figure 3).

### Recognition of U12-Type Introns by the Minor Spliceosome

The U12-dependent spliceosome contains four specific snRNPs, U11, U12, U4atac, and U6atac, each of which contains a specific snRNA component that is equivalent to but distinct from its U2-type counterpart, i.e., U1, U2, U4, and U6, respectively. U5 snRNP is shared between the two spliceosomes. Although the sequences of the snRNAs with equivalent function are quite divergent in the two spliceosomes, they share a common overall secondary structure (Figure 4). U4atac, U6atac, and U5 associate into a tri-snRNP,

similar to the major U4/U6.U5, and the protein composition of the major and minor tri-snRNPs appears to be very similar, if not identical.<sup>28</sup> Moreover, it has also been shown that equivalent stem-loop structures of U4 and U4atac function as binding platforms for proteins that are required for the tri-snRNP formation.<sup>28,29</sup> In contrast, while the major U1 and U2 snRNPs are distinct snRNPs, their counterparts are present in the nucleus as a U11/U12 di-snRNP.<sup>30,31</sup> Interestingly, only seven protein components have thus far been reported to be specific to the minor spliceosomes (Table 1), and they are all located in the U11/U12 di-snRNP.<sup>32,33</sup> U11/U12 also lacks all the U1-specific proteins and some U2-associated proteins (Table 1), making it the most divergent component of the minor spliceosome, in comparison to its counterparts in the major spliceosome.

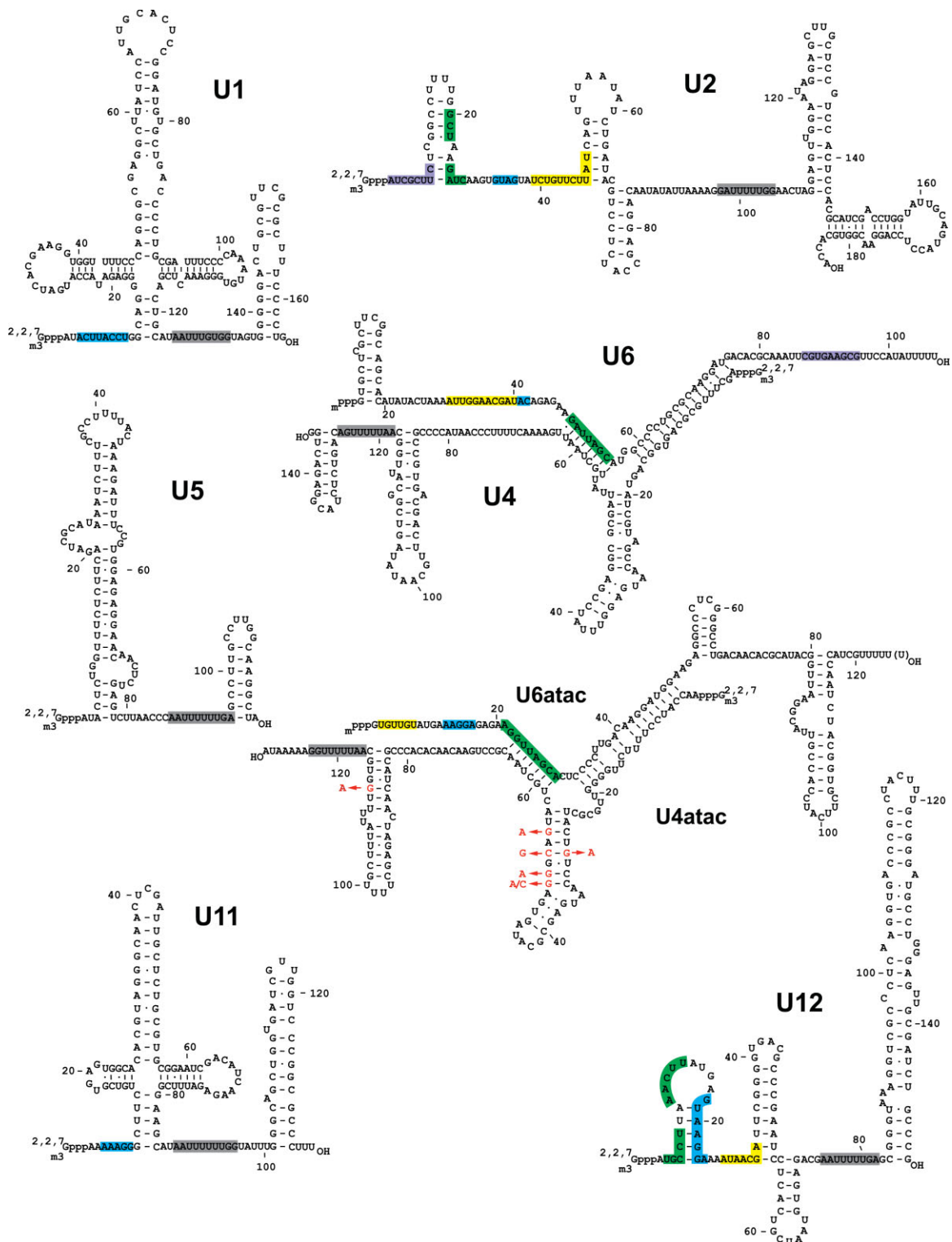
The overall assembly pathways of the two spliceosomes are similar, and the main difference is the absence of a separate commitment complex in the minor spliceosome. Instead, the preformed U11/U12 di-snRNP binds the intron as a unit, and the 5'ss and



**FIGURE 3** | Spliceosome assembly. The interactions of the spliceosomal snRNPs and some selected non-snRNP protein complexes at various stages of spliceosome assembly (complexes E, A, B\*, and C) are depicted schematically for both the U2- and U12-dependent spliceosomes. The Prp19/CDC5 complex is indicated by '19C'. Its association with the U12-dependent spliceosome is inferred from the major spliceosome and is therefore indicated with a question mark. (Adapted with permission from Ref 27. Copyright 2003 Macmillan Publishers Ltd)

BPS are recognized in a cooperative manner within the A complex, although U11/5'ss basepairing still precedes the formation of stable U12/BPS basepairing.<sup>47</sup> The initial basepairing interactions at the 5'ss are also

different in the two spliceosomes: in contrast to the U1, the U11 snRNA does not basepair across the exon-intron boundary or even with the first three nucleotides of the intron.<sup>2,48</sup> Instead, these nucleotides



**FIGURE 4** | The predicted secondary structures of the human spliceosomal snRNAs. The binding sites for Sm proteins are shaded in gray, and the sequences interacting with the 5'ss or BPS in cyan. Sequences involved in various U2/U6 or U12/U6atac interactions are indicated by green (helix I), purple (helix II), and yellow shading (helix III), similar to Figure 5. Nucleotide modifications are omitted. (Structures are based on data originally published in Ref 34 for U1, U2, and U5, Ref 18 for U11, Ref 35 for U12, and Ref 29 for U4, U6, U4atac, and U6atac). The locations and identities of the Taybi-Linder syndrome or microcephalic osteodysplastic primordial dwarfism type I (TALS/MOPD1) mutations in the U4atac snRNA are from Ref 36 and are indicated in red.

**TABLE 1** | Proteins of the U1, U2, and U11/U12 snRNPs

12S U1	17S U2	18S U11/U12	Functions (with Selected References)
Sm proteins <sup>1</sup>	Sm proteins <sup>1</sup>	Sm proteins <sup>1</sup>	snRNP core components <sup>26</sup>
U1-A (SNRPA)			Structural; RNA-binding <sup>37</sup>
U1-C (SNRPC)			5'ss recognition <sup>37</sup>
U1-70K (SNRNP70)			Structural; SR protein interactions <sup>37,38</sup>
	U2A' (SNRPA1)		Structural; RNA-binding
	U2B'' (SNRPB2)		Structural; RNA-binding
	SF3a complex <sup>2</sup>		BPS binding <sup>39</sup>
	SF3b complex <sup>2</sup>	SF3b complex <sup>2</sup>	BPS binding <sup>33,39</sup>
		20K (ZMAT5)	Unknown; homology to U1C <sup>32</sup>
		25K (SNRNP25)	Unknown
		31K (ZCRB1)	Unknown; RNA-binding <sup>40</sup>
		35K (SNRNP35)	SR protein interactions, homology to U1-70K <sup>33,41,42</sup>
		48K (SNRNP48)	5'ss recognition <sup>43</sup>
		59K (PDCD7)	Structural, binds 48K and 65K <sup>43,44</sup>
		65K (RNPC3)	Structural, binds U12 snRNA <sup>44</sup>
		Urp <sup>3</sup> (ZRSR2)	3'ss recognition <sup>45</sup>
	hPrp43 <sup>4</sup>	hPrp43 (DHX15)	
		Y Box-1 <sup>3</sup> (YBX1)	

The Hugo names of proteins have been provided in parentheses.

<sup>1</sup>Sm proteins B/B', D1, D2, D3, E, F, and G,

<sup>2</sup>Multi-subunit complexes.

<sup>3</sup>Also present in the major spliceosome, but not in U1 or U2 snRNPs.

<sup>4</sup>Almost stoichiometric presence in the 17S U2. Other proteins associated with U2 in substoichiometric amounts are omitted.<sup>46</sup>

are recognized specifically by the U11-48K protein, which likely also stabilizes the U11/5'ss helix.<sup>43,49</sup> Surprisingly, even though U11-48K does not share significant sequence homology with the U1-specific protein U1-C, both appear to stabilize the binding of their respective snRNAs in a similar manner through their zinc finger domains.<sup>37,49</sup>

Interaction of the U12 snRNA with the BPS is similar to the U2/BPS basepairing in the major spliceosome, resulting in the exclusion of the branch point (BP) adenosine from the U12/BPS helix.<sup>50,51</sup> There seems to be flexibility in the choice of the BP adenosine, as either one of the two A residues present in the consensus BPS can be used as a BP, depending on the intron in question.<sup>52</sup> BPS recognition is likely stabilized by proteins of the SF3b complex, which is present in both spliceosomes (Table 1), and is known to bind to the BPS in the major spliceosome, with the protein factor p14 shielding the BP adenosine from premature activation.<sup>39,53,54</sup> U12-type introns do not have PPTs, and U2AF is not required for the recognition of U12-type introns.<sup>55</sup> However, a U11/U12 di-snRNP component related to U2AF<sup>35</sup>, Urp,<sup>32</sup> is required for A complex formation and 3'ss recognition.<sup>45</sup> Urp is

not specific to the minor spliceosome, but its function is different than in the major spliceosome, where it apparently displaces U2AF from the 3'ss after the first catalytic step.<sup>45</sup>

In the major spliceosome the U1 and U2 snRNPs recognize the 5'ss and BPS independently, and non-snRNP proteins are required for their association to each other.<sup>56</sup> In contrast, the U12-type 5'ss and BPS are linked through the internal components of the U11/U12 di-snRNP already at the earliest phase of intron recognition. The protein factors U11-59K and U11/U12-65K interact at the interface of the two snRNPs.<sup>44</sup> U11-59K further interacts with the U11-48K protein,<sup>43</sup> while U11/U12-65K directly binds to the 3' terminal stem loop of U12 snRNA.<sup>44</sup> Owing to the compact structure of the di-snRNP,<sup>57</sup> further protein and/or RNA interactions are likely to contribute to the association of the two snRNPs. Furthermore, the binding sites for the 5'ss and BPS must be close to each other within the U11/U12 structure, as the very 5' end of U12 snRNA can be cross-linked to the pre-mRNA 2 nt upstream of the 5'ss already in the A complex, suggesting that the BPS and 5'ss must be within 50 Å of one another.<sup>58</sup> Thus, the 5' end of

U12 is already close to the position required for the formation of the catalytic core of the spliceosome.

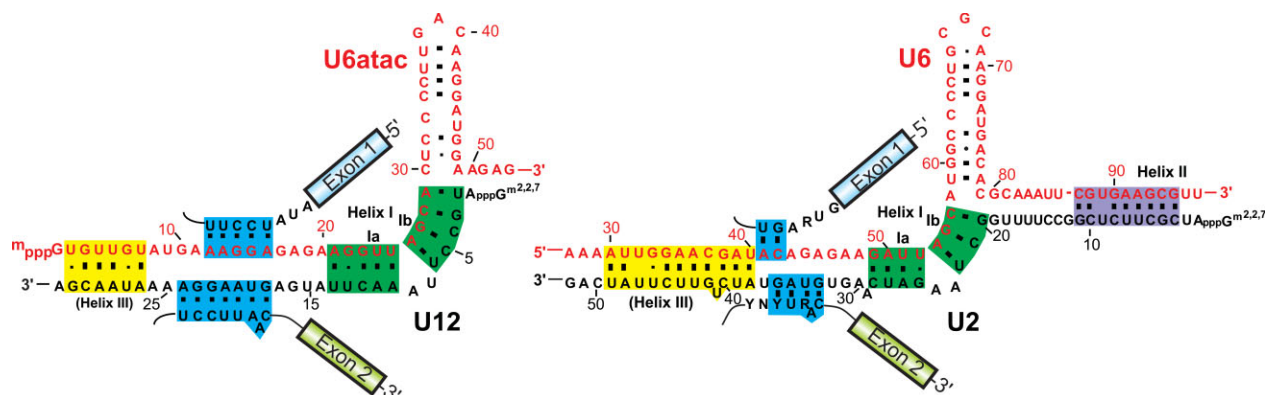
### Assembly of the Catalytic Core is Similar in the Two Spliceosomes

The formation of the catalytically active spliceosome is thought to follow a pathway similar to that of the major spliceosome. After initial recognition of the splice sites by the U11/U12 di-snRNP, the U4atac/U6atac.U5 tri-snRNP associates with pre-spliceosome to form complex B<sup>59</sup> (Figure 3). A terminal stem-loop structure in the U6atac snRNA (possibly together with U4atac snRNA sequences) contains a signal that directs the U4atac/U6atac.U5 tri-snRNP to the minor and not to the major pre-spliceosome.<sup>60</sup> A large number of structural rearrangements convert the pre-catalytic spliceosome (complex B) to the catalytic configuration (complexes B\* and C) in a manner similar to the major spliceosome. U6atac snRNA replaces U11 at the 5'ss, U4atac/U6atac structure is unwound, and U12 and U6atac basepair with each other to form the 'catalytic core' structure in which the reactive A residue at the BP and the 5'ss are juxtaposed for the first step of the catalysis.<sup>59,61–63</sup> During this process both the U11 and U4atac snRNAs are released from the spliceosome. In the major spliceosome, specific helicases drive forward spliceosome assembly and the transesterification reactions.<sup>64</sup> There is virtually no data showing whether helicases have similar activities in the minor spliceosome. However, given that most proteins are shared<sup>28,32</sup> and no helicases specific to the minor spliceosome have been identified, it is likely that the helicase activities are also the same during the assembly of the two spliceosomes.

Although the formation of the catalytic core appears to be somewhat more flexible in the U12-dependent spliceosome,<sup>61</sup> the structure and function of the core are likely to be highly similar. Indeed, both spliceosomes employ the same two-step transesterification mechanism for intron removal,<sup>51</sup> and U12-dependent splicing is even supported by the modified U6atac snRNAs in which the functional domain has been replaced by that of U6 snRNA.<sup>65</sup> The U12/U6atac interactions in the catalytic core of the U12-dependent spliceosome resemble those in the major spliceosome, although the helix II structure present in the U2/U6 complex cannot be formed in the minor spliceosome (Ref 59; Figure 5). Nevertheless, the similarity of snRNA structural domains suggests that, like the major spliceosome,<sup>66</sup> the minor spliceosome is likely to use RNA-based catalysis.

### Removal of U12-Type Introns Is Slow

Both intron types are spliced co-transcriptionally within the nucleus (Ref 68 and Box 1) and might be expected to display similar kinetics. However, the removal of U12-type introns appears to be significantly slower. Early *in vitro* splicing experiments documented a splicing rate for U12-type introns that was three- to fivefold slower than that of U2-type introns.<sup>47,51,69</sup> Similar observations were also made in *in vivo* experiments, where approximately twofold higher levels of unspliced U12-type introns were detected in the steady-state transcript pools isolated from insect and mammalian cells.<sup>69–71</sup> This is consistent with the observation that co-transcriptional splicing of U12-type introns is at least twofold slower than that of U2-type introns.<sup>68</sup> Thus far, the reason for the lower efficiency remains unknown. Minor



**FIGURE 5** | RNA–RNA interactions in the catalytic cores of the minor and major spliceosomes. Interactions between snRNAs and the 5'ss or BPS are indicated by cyan shading. U2/U6 or U12/U6atac interactions are indicated by green (helix I), purple (helix II), and yellow shading (helix III), as in Figure 4. The minor spliceosome structure is based on data published in Ref 59. The U12/U6atac helix III structure is controversial as it is not conserved in plants,<sup>67</sup> but mutations in U12 snRNA that weaken this structure reduce the splicing activity in mammals.<sup>35</sup>

## BOX 1

## NUCLEAR LOCALIZATION OF THE U12-DEPENDENT SPLICEOSOME

A provocative hypothesis by König et al. suggested that the U12-dependent spliceosome is located in the cytoplasm and that the splicing of U12-type introns also takes place there.<sup>73</sup> This suggestion was based on *in situ* hybridization studies, mostly with zebrafish, as well as cell fractionation and reverse transcription polymerase chain reaction (RT-PCR) analyses. The publication contradicted many of the earlier studies on minor spliceosome component localization and function (See Ref 74 and references therein) and spurred a lively debate on the localization of the U12-dependent spliceosome. Subsequent publications failed to reproduce the key findings of the paper and instead demonstrated nuclear localization for U12-type spliceosome snRNP and protein components in mammalian cells and tissues,<sup>75</sup> nuclear splicing of U12-type introns in *Xenopus* oocytes,<sup>76</sup> and co-transcriptional splicing of U12-type introns.<sup>68</sup> Together, these subsequent studies provide firm evidence for the nuclear localization of the splicing of U12-type introns.

spliceosome snRNPs are ca 100-fold less abundant than the major snRNPs,<sup>31,59</sup> and this could underlie the observed kinetic differences. However, the observation that a further 10-fold reduction in the levels of U4atac snRNA has no apparent effect on the efficiency of endogenous U12-dependent splicing argues against this simplistic explanation.<sup>70</sup> The slower rate could also be related to the kinetic effects caused by the less flexible recognition phase of U12-type introns or by the inability to form some of the structures present in the catalytic core of the major spliceosome. Regardless of the underlying mechanism, Patel et al.<sup>71</sup> have suggested that the slower rate of splicing could constitute a rate-limiting mechanism for the expression of genes containing U12-type introns. In this model, the transcripts containing unspliced U12-type introns would get trapped in the nucleus where they could be targeted by nuclear surveillance mechanisms.<sup>72</sup>

## EXON DEFINITION INTERACTIONS AND ALTERNATIVE SPLICING

Splicing enhancer and silencers sequences are important for defining splice sites both in constitutive and alternative splicing, and typically function by

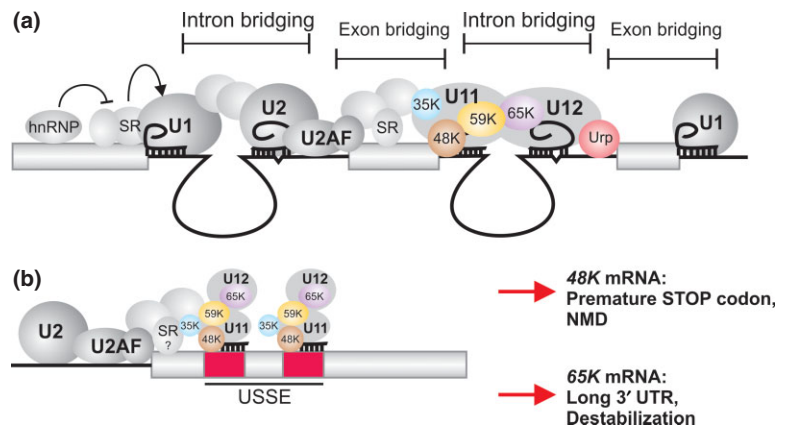
recruiting regulatory splicing factors, such as SR and hnRNP proteins. U12-type introns are not an exception in this regard, as U12-dependent splicing also responds to SR and hnRNP proteins that bind to splicing regulatory elements near the U12-type splice sites.<sup>17,77–79</sup> A prototypical case of SR protein-mediated enhancement in the major spliceosome is the interaction of SRSF1 with the U1-70K protein, which enhances 5'ss recognition.<sup>38</sup> The U11/U12 di-snRNP protein U11-35K is the putative paralog of U1-70K,<sup>33</sup> and may have a similar function in stabilizing U11/5'ss binding. This notion is indirectly supported by the findings that U11-35K interacts with SR proteins in plants, including SRSF1 homologs,<sup>41,42</sup> and that U12-type A complex formation is enhanced by SR proteins.<sup>77</sup> As in the major spliceosome, SR proteins likely enhance intron recognition in multiple ways and through several spliceosomal factors, and have also been shown to directly contact U12-type 5'ss and BPS.<sup>55,77</sup>

A particular problem related to the processing of nascent transcripts is that introns (especially those of vertebrates) can be extremely long, such that it can take an hour for an entire intron to be transcribed. Therefore, initial splice site definition typically takes place over the considerably shorter exons, with 3'ss-recognizing factors in the upstream intron interacting with the 5'ss-recognizing factors in the downstream intron (Figure 6(a)). Again, U12-dependent splicing has been shown to be enhanced by the presence of flanking U2-type introns,<sup>17,80</sup> an effect which is likely mediated by the SR proteins or other auxiliary factors. The most detailed evidence for exon definition interactions between components of the two spliceosomes comes from the alternative splicing (AS) of the transcripts coding for the minor spliceosome components U11-48K and U11/U12-65K.<sup>81</sup> In this system, an alternative U2-type 3'ss is activated by the U11/U12 di-snRNP bound at an atypical downstream splicing enhancer element consisting of a tandem repeat of U12-type 5'ss, designated as U11 snRNP-binding splicing enhancer (USSE; Figure 6(b)). The alternatively spliced transcripts are degraded by the nonsense-mediated decay (NMD) pathway or other RNA degradation machineries, and this system thus functions as a regulatory feedback mechanism for minor spliceosome components.

Given that AS is extremely prevalent in humans, and the extensive interactions of the minor spliceosomes with other splicing factors, it is somewhat surprising that AS events involving U12-type introns appear to be almost entirely absent. This may be because of the more rigid sequence and distance constraints on the U12-type 5'ss, BPS, and



**FIGURE 6** | Exon definition interactions and regulation of U12-type factors by alternative splicing nonsense-mediated decay (AS-NMD). (a) Exon definition interactions form between U2- and U12-dependent spliceosomes in both the upstream and downstream direction, aided by SR protein interactions.<sup>77,80</sup> Intron bridging differs in U12-dependent splicing as a consequence of the cooperative recognition by the U11/U12 di-snRNP.<sup>47</sup> (b) The di-snRNP activates alternative splicing by binding to the U11 snRNP-binding splicing enhancer (USSE) element in the U11-48K and U11/U12-65K transcripts and recruits U2-type splicing factors to the upstream 3'ss. Alternatively spliced products are degraded.<sup>81</sup>



3'ss. AS events involving the use of an alternative 3'ss have been observed, but it is not entirely clear whether they are the result of true regulated AS or of splicing errors, especially since the selection of the U12-type 3'ss has been shown to be prone to errors.<sup>13,16,82,83</sup> Exon skipping, which is the most common type of AS observed with major introns, has been reported only in a few instances involving U12-type introns, and the functional significance of these AS events is not known.<sup>14,16</sup> The absence of exon skipping is likely to be due to the general incompatibility of U12 and U2-type splice sites and the scarcity of genes containing more than one U12-type intron. Alternative usage of mutually exclusive U12-type and U2-type splice sites has been observed in at least three cases. In vertebrates, the members of the *JNK* gene family contain a hybrid intron with a U12-type 5'ss and U2-type BPS and 3'ss. These can be spliced in a mutually exclusive fashion to either the 3'ss of the downstream U12-type intron or to the 5'ss of the upstream U2-type intron, respectively,<sup>82</sup> and at least in mice these isoforms show tissue specificity, with the former isoform preferentially expressed in neurons. The *D. melanogaster* genes *prospero* and *dUrp* (which encodes the Urp protein present in both spliceosomes) contain overlapping U12 and U2-type introns or 'twintrons'.<sup>2,14,84</sup> The U12-dependent *dUrp* mRNA isoform is apparently destroyed by NMD,<sup>14</sup> while *prospero* alternative splicing results in distinct developmentally regulated protein isoforms with amino-acid differences in the homeodomain region.<sup>85,86</sup>

## PHYSIOLOGICAL SIGNIFICANCE OF THE MINOR SPLICEOSOME

### Minor Splicing Is Essential for Development

Knockdown of protein components specific to the minor spliceosome leads to reduced proliferation of cultured cells,<sup>32,43</sup> and thus minor splicing is

an essential process. Proper U12-dependent splicing is also required for development, as shown in humans<sup>87,88</sup> and in *D. melanogaster*,<sup>84</sup> despite the fact that the latter only contains a handful of U12-type introns.<sup>20,21</sup> Blocking minor splicing in zebrafish embryos resulted in developmental defects at a distinct stage, in comparison to blocking the major spliceosome.<sup>73</sup> Knockdown of the U11-31K protein in *A. thaliana* also led to abnormal growth of the plant, but only after the bolting stage.<sup>40</sup> These findings indicate that minor splicing has specific functions in regulating developmental processes. These functions may be mediated in part by alternative splicing, as with the *prospero* transcripts in *D. melanogaster*. However, it is likely that another, even more significant factor affecting developmental processes is the slower kinetics of minor splicing, which can limit not only the expression of the genes containing them<sup>68,70,71</sup> but also downstream pathways. Indeed, defects in minor splicing have been shown to perturb the expression of multiple metabolic genes (which themselves do not contain U12-type introns) in *D. melanogaster*, likely underlying the developmental arrest.<sup>84,89</sup> The developmental defect following the U11-31K knockdown in *A. thaliana* was also associated with downregulation of a specific set of genes related to gibberellic acid metabolism.<sup>40</sup> Thus, U12-dependent splicing is necessary for multiple processes required for the viability and development of multicellular organisms, and can also affect specific cellular pathways.

### Minor Splicing and Human Disease

From a medical point of view, it would be interesting to know to what extent the activity of the minor spliceosome contributes to regulating the expression of genes in different human tissues. This question remains mostly unanswered. Some protein components specific to the minor spliceosome (31K and 65K)

have been observed to be differentially expressed in different tissues,<sup>90,91</sup> and U12-type intron-containing genes in bone marrow CD34 positive cells and B lymphoblasts,<sup>24</sup> but the correlation between minor splicing and gene expression in different tissues has not been analyzed systematically. To date, only a few diseases have been linked to defects in minor splicing, some of them displaying tissue-specific symptoms, while others have a wide range of system-wide defects. Mutations in U12-type 5' ss in the *LKB1* and *SEDL* genes cause Peutz-Jeghers syndrome and spondyloepiphyseal dysplasia tarda, respectively,<sup>83,92,93</sup> and these diseases likely arise from the inactivity or insufficiency of the respective gene products.

A more complex class of diseases are those arising from defects in minor spliceosome components of which the TALS/MOPD1 (Taybi-Linder syndrome or microcephalic osteodysplastic primordial dwarfism type I) is a recent example. This severe developmental disorder is caused by recessive mutations in the U4atac gene (Refs 87 and 88; see Figure 4). Unlike their counterparts in the major spliceosome, each minor snRNA is expressed from a single locus in the genome,<sup>88,94,95</sup> and the mutations in such loci can therefore potentially disable the function of the snRNA in question. In the case of TALS/MOPD1, the mutations are mostly in the 5' stem loop of U4atac (shown in Figure 4) and are predicted to prevent the binding of tri-snRNP-specific 15.5K and 61K proteins which could lead to a defect in the formation of U4atac/U6atac.U5 tri-snRNP,<sup>36,96</sup> but experimental evidence for this is still lacking. This presumed defect in snRNP assembly decreases the activity of the U12-dependent spliceosome almost 10-fold,<sup>88</sup> but still allows correct splicing of U12-type introns in most genes, albeit at significantly reduced levels. This resembles a situation with *D. melanogaster* carrying mutations in U6atac snRNA.<sup>70</sup> An open question is whether the developmental defects observed in this disorder are due to a large-scale reduction in the expression of U12-type intron-containing genes or just a few affected genes in each tissue as has been suggested for *D. melanogaster*.<sup>89</sup> A comparable case is spinal muscular atrophy (SMA), which arises from the mutation of the gene coding for the general snRNP assembly factor SMN, and is characterized by multiple defects in mRNA processing, including widespread changes in alternative splicing,<sup>97,98</sup> but initially only affects spinal cord neurons. SMA has been reported to involve a decrease in the levels of the minor tri-snRNP<sup>99</sup> and possibly U11/U12 di-snRNP levels as well,<sup>98,100</sup> together with defective splicing of a few, but not all, U12-type introns.<sup>99</sup> While the direct link of these findings on the pathogeny

of SMA is not entirely clear, it is possible that the defect in the splicing of some U12-type introns in neurons either causes or contributes to the disease phenotype.

## EVOLUTION OF THE MINOR SPLICEOSOME

It is clear that the U12-dependent spliceosome is crucial for the viability and development of many multicellular organisms. However, intriguing questions about the minor spliceosome remain: Where did it originate? Why is it not present in all eukaryotes, and do its function and importance vary in different evolutionary lineages? We do not currently have clear answers to these questions, but recent advances in large-scale sequencing and bioinformatics are paving the way toward deeper understanding of the significance of the minor spliceosome.

### Origins and Conservation of the Minor Spliceosome

The two spliceosomes have been proposed to originate from group II self-splicing introns, but it is not clear whether they arose simultaneously or sequentially, and whether this took place in one ancestral eukaryote or separate lineages that later fused.<sup>4,101–103</sup> However, based on the distribution of introns and spliceosome components, it is evident that both U2- and U12-type introns and spliceosomes were present in the last common eukaryotic ancestor (Refs 18, 19, and 23; Figure 2). U12-type introns show remarkable conservation over long evolutionary distances: the rate of intron loss or conversion has been very low in all studied vertebrates, and while *D. melanogaster* has only 19 U12-type introns, almost all of them have orthologous U12-type introns in human.<sup>14</sup> Most remarkably, 20 U12-type introns are conserved in orthologous positions between human and *A. thaliana*.<sup>16,104</sup> Also, although the sequences of U12-type snRNAs are more diverged in distantly related organisms than their U2-type counterparts, their structures are highly conserved.<sup>18,20,41,67</sup>

### Repeated Loss of U12-Type Introns

Despite the high overall conservation of U12-type introns, they are entirely absent from many eukaryotes whose relatives nonetheless have them, suggesting repeated loss of U12-dependent splicing during evolution. The loss seems to have been most prevalent among diverse eukaryotic microbes, but some animals, including some but not all nematodes, have also lost

their U12-type introns (Refs 18 and 19; Figure 2). The loss of U12-type introns has also been extensive in *Diptera*, which typically only have 15–20 U12-type introns in total in their genomes, several fold less than in most other insects.<sup>14,21</sup> The precise reasons behind these losses remain enigmatic. However, it has been suggested that introns confer a slight deleterious effect for the fitness of the organism,<sup>101</sup> and could thus be removed more efficiently in organisms with large effective population sizes, such as microbes and small invertebrates.<sup>14</sup> Presumably, U12-type introns may be even slightly more deleterious, due to their slower excision and lower accuracy. In support of this, the loss of U12-type introns generally correlates with reduction in overall intron frequency, but U12-type introns still tend to be lost to a larger extent.<sup>14,21</sup>

In most cases, the mechanism of U12-type intron loss involves a simple deletion of the intronic sequence, apparently through the homologous recombination of reverse transcribed mRNAs.<sup>14</sup> Another pathway is the step-wise conversion of U12-type introns into U2-type introns. Natural examples as well as experimental approaches have also shown that simple point mutations in the 5' splice site of a U12-type intron, especially in the GT–AG subtype, suffice to turn it into a U2-type 5' splice site,<sup>4,5,14</sup> as U2-type introns have much more degenerate splicing signals. Subtype switches from a U12-type AT–AC intron to U12-type GT–AG are also observed.<sup>4,8,14</sup> However, such changes appear to be rare, possibly due to the poor splicing efficiency of the intermediate forms (e.g., AT–AG).<sup>11,17,21,105</sup> Sometimes intron conversion can also occur by the activation of a cryptic U2-type 5' splice site near the original U12-type 5' splice site.<sup>16,19</sup> Given that the chance of creating the highly stringent U12-type 5' splice site from a U2-type 5' splice site is low, conversion in the reverse direction is a highly unlikely event. In fact, only one instance of a novel U12-type intron has been observed (in the *dUrp* twintron).<sup>14</sup> Thus, without natural selection acting in favor of maintaining U12-type introns, they are likely to be converted into U2-type introns during evolution.

### Conservation of Regulatory Features

Many splicing factors are regulated through a feedback mechanism, in which they bind to regulatory elements within their own pre-mRNA, resulting in alternatively spliced, unstable mRNA isoforms. Interestingly, all canonical SR and hnRNP proteins, as well as a number of core spliceosomal factors, are regulated in this manner, and in many cases the regulatory elements are also evolutionarily

conserved.<sup>24,106–110</sup> While most of the splicing factors thus regulated are common to both spliceosomes, certain spliceosome-specific proteins have also been shown to be alternatively spliced to produce unstable mRNA isoforms. These include, e.g., U1-70K and U11/U12-25K, for which the AS events are also evolutionarily conserved between human and mouse.<sup>108</sup> The conservation of the AS events suggests that they have an important function in regulating the levels of these proteins, and may consequently affect the activity of the specific spliceosome in question. The feedback mechanism regulating the levels of minor spliceosome-specific proteins U11-48K and U11/U12-65K through the USSE element (Figure 6(b)) is conserved over extremely long evolutionary distances, in plants as well as in animals.<sup>81</sup> Interestingly, the USSE is most conserved in vertebrates,<sup>81</sup> which also show the most extensive conservation of U12-type introns,<sup>8,14,105</sup> and is absent in *Diptera*, where both the overall numbers and interspecies conservation of U12-type introns are low.<sup>21</sup> It is thus possible that this mode of regulation is important for the rate-limiting function of the minor spliceosome, without which the presence of U12-type introns would be mainly deleterious. Interestingly, the emergence of an AS-NMD-targeting U12-type intron in the *Drosophila dUrp*<sup>14</sup> suggests that it, too, may function as a feedback mechanism for the minor spliceosome. Thus, it is possible that mechanisms limiting the activity of the minor spliceosome are essential for its function in regulating the expression of U12-type intron-containing genes.

### CONCLUSION

Historically, insight into the function and significance of the U12-dependent spliceosome has come primarily from the biochemical characterization of splicing factors and spliceosome assembly, and has recently been complemented by the increasing number of bioinformatic analyses. However, our understanding of the significance of minor splicing is far from complete in both regards. The function of many of the U12-specific components is still unknown, and more components might still remain undetected. Importantly, although there are many independent lines of evidence indicating that U12-type introns are important for the regulation of gene expression, definitive large-scale evidence on the connection between the activity of the minor spliceosome and gene regulation is still lacking. Are there more regulatory elements that respond to U12-type factors? In what other ways is the minor spliceosome regulated,

and how does this correlate with the expression of U12-type intron-containing genes in different tissues? Could there be U12-type introns with more degenerate splice sites, and could they be involved in more extensive alternative splicing? An exciting way of studying

these questions is now offered by the rapidly developing high-throughput sequencing and cross-linking methods, and will hopefully deepen our understanding on the significance of minor splicing in the coming years.

## ACKNOWLEDGMENTS

This work was supported by the Academy of Finland and Sigrid Jusélius Foundation (MJF) and by the Helsinki Graduate Program in Biotechnology and Molecular Biology (EHN).

## REFERENCES

- Jackson IJ. A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res* 1991, 19:3795–3798.
- Hall SL, Padgett RA. Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice-sites. *J Mol Biol* 1994, 239:357–365.
- Wu Q, Krainer AR. Splicing of a divergent subclass of AT-AC introns requires the major class spliceosomal snRNAs. *RNA* 1997, 3:586–601.
- Burge CB, Padgett RA, Sharp PA. Evolutionary fates and origins of U12-type introns. *Mol Cell* 1998, 2:773–785.
- Dietrich RC, Ingorvaia R, Padgett RA. Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol Cell* 1997, 1:151–160.
- Sharp PA, Burge CB. Classification of introns: U2 type or U12 type. *Cell* 1997, 91:875–879.
- Alioto TS. U12DB: a database of orthologous U12-type spliceosomal introns. *Nucleic Acids Res* 2007, 35(suppl 1):D110–D115.
- Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res* 2006, 34:3955–3967.
- Gao K, Masuda A, Matsuura T, Ohno K. Human branch point consensus sequence is yUnAy. *Nucleic Acids Res* 2008, 36:2257–2267.
- Workman CT, Yin Y, Corcoran DL, Ideker T, Stormo GD, Benos PV. enoLOGOS: a versatile web tool for energy normalized sequence logos. *Nucleic Acids Res* 2005, 33(suppl 2):W389–W392.
- Dietrich RC, Fuller JD, Padgett RA. A mutational analysis of U12-dependent splice site dinucleotides. *RNA* 2005, 11:1430–1440.
- Dietrich RC, Peris MJ, Seyboldt AS, Padgett RA. Role of the 3' splice site in U12-dependent intron splicing. *Mol Cell Biol* 2001, 21:1942–1952.
- Levine A, Durbin R. A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res* 2001, 29:4006–4013.
- Lin C-F, Mount S, Jarmolowski A, Makalowski W. Evolutionary dynamics of U12-type spliceosomal introns. *BMC Evol Biol* 2010, 10:47.
- Brock JE, Dietrich RC, Padgett RA. Mutational analysis of the U12-dependent branch site consensus sequence. *RNA* 2008, 14:2430–2439.
- Zhu W, Brendel V. Identification, characterization and molecular phylogeny of U12-dependent introns in the *Arabidopsis thaliana* genome. *Nucleic Acids Res* 2003, 31:4561–4572.
- Lewandowska D, Simpson CG, Clark GP, Jennings NS, Barciszewska-Pacak M, Lin CF, Makalowski W, Brown JW, Jarmolowski A. Determinants of plant U12-dependent intron splicing efficiency. *Plant Cell* 2004, 16:1340–1352.
- Russell AG, Charette JM, Spencer DF, Gray MW. An early evolutionary origin for the minor spliceosome. *Nature* 2006, 443:863–866.
- Bartschat S, Samuelsson T. U12 type introns were lost at multiple occasions during evolution. *BMC Genomics* 2010, 11:106.
- Schneider C, Will CL, Brosius J, Frilander MJ, Lührmann R. Identification of an evolutionarily divergent U11 small nuclear ribonucleoprotein particle in *Drosophila*. *Proc Natl Acad Sci U S A* 2004, 101:9584–9589.
- Janice J, Pande A, Weiner J, Lin CF, Makalowski W. U12-type spliceosomal introns of *Insecta*. *Int J Biol Sci* 2012, 8:344–352.
- Marquez Y, Brown JWS, Simpson C, Barta A, Kalyna M. Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Res* 2012, 22:1184–1195.
- Dávila López M, Rosenblad MA, Samuelsson T. Computational screen for spliceosomal RNA genes aids in defining the phylogenetic distribution of major and minor spliceosomal components. *Nucleic Acids Res* 2008, 36:3001–3010.

24. Yeo GW, Van Nostrand EL, Liang TY. Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. *PLoS Genet* 2007, 3:e85.
25. Wu Q, Krainer AR. AT-AC pre-mRNA splicing mechanism and conservation of minor introns in voltage-gated ion channel genes. *Mol Cell Biol* 1999, 19:3225–3236.
26. Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harbor Perspect Biol* 2011, 3:a003707.
27. Patel AA, Steitz JA. Splicing double: Insights from the second spliceosome. *Nat Rev Mol Cell Biol* 2003, 4:960–970.
28. Schneider C, Will CL, Makarova OV, Makarov EM, Lührmann R. Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Mol Cell Biol* 2002, 22:3219–3229.
29. Nottrott S, Urlaub H, Lührmann R. Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J* 2002, 21:5527–5538.
30. Wassarman KM, Steitz JA. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol Cell Biol* 1992, 12:1276–1285.
31. Montzka KA, Steitz JA. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12 etc. *Proc Natl Acad Sci U S A* 1988, 85:8885–8889.
32. Will CL, Schneider C, Hossbach M, Urlaub H, Rauhut R, Elbashir S, Tuschl T, Lührmann R. The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* 2004, 10:929–941.
33. Will CL, Schneider C, Reed R, Lührmann R. Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science* 1999, 284:2003–2005.
34. Yu Y-T, Scharl E, Smith CM, Steitz JA. The growing world of small nuclear ribonucleoproteins. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world*. Cold Spring Harbour, NY: Cold Spring Harbour Laboratory Press; 1999.
35. Sikand K, Shukla GC. Functionally important structural elements of U12 snRNA. *Nucleic Acids Res* 2011, 39:8531–8543.
36. Padgett RA. New connections between splicing and human disease. *Trends Genet: TIG* 2012, 28: 147–154.
37. Pomeranz Krummel DA, Oubridge C, Leung AKW, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5Å resolution. *Nature* 2009, 458:475–480.
38. Cho S, Hoang A, Sinha R, Zhong XY, Fu XD, Krainer AR, Ghosh G. Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. *Proc Natl Acad Sci U S A* 2011, 108: 8233–8238.
39. Gozani O, Feld R, Reed R. 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* 1996, 10:233–243.
40. Kim WY, Jung HJ, Kwak KJ, Kim MK, Oh SH, Han YS, Kang H. The Arabidopsis U12-type spliceosomal protein U11/U12-31K is involved in U12 intron splicing via RNA chaperone activity and affects plant development. *Plant Cell* 2010, 22:3951–3962.
41. Lorković ZJ, Lehner R, Forstner R, Barta A. Evolutionary conservation of minor U12-type spliceosome between plants and humans. *RNA* 2005, 11: 1095–1107.
42. Lorković ZJ, Lopato S, Pexa M, Lehner R, Barta A. Interactions of Arabidopsis RS domain containing cyclophilins with SR proteins and U1 and U11 small nuclear ribonucleoprotein-specific proteins suggest their involvement in pre-mRNA splicing. *J Biol Chem* 2004, 279:33890–33898.
43. Turunen JJ, Will CL, Grote M, Lührmann R, Frilander MJ. The U11-48K protein contacts the 5' splice site of U12-type introns and the U11-59K protein. *Mol Cell Biol* 2008, 28:3548–3560.
44. Benecke H, Lührmann R, Will CL. The U11/U12 snRNP 65K protein acts as a molecular bridge, binding the U12 snRNA and U11-59K protein. *EMBO J* 2005, 24:3057–3069.
45. Shen H, Zheng X, Luecke S, Green MR. The U2AF35-related protein Urp contacts the 3' splice site to promote U12-type intron splicing and the second step of U2-type intron splicing. *Genes Dev* 2010, 24:2389–2394.
46. Will CL, Urlaub H, Achsel T, Gentzel M, Wilm M, Lührmann R. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J* 2002, 21:4978–4988.
47. Frilander MJ, Steitz JA. Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. *Genes Dev* 1999, 13: 851–863.
48. Kolossova I, Padgett RA. U11 snRNA interacts in vivo with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA* 1997, 3:227–233.
49. Tidow H, Andreeva A, Rutherford TJ, Fersht AR. Solution structure of the U11-48K CHHC zinc-finger domain that specifically binds the 5' splice site of U12-type introns. *Structure* 2009, 17:294–302.
50. Hall SL, Padgett RA. Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* 1996, 271:1716–1718.

51. Tarn W-Y, Steitz JA. A novel spliceosome containing U11, U12 and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. *Cell* 1996, 84:801–811.
52. McConnell TS, Cho SJ, Frilander MJ, Steitz JA. Branchpoint selection in the splicing of U12-dependent introns *in vitro*. *RNA* 2002, 8:579–586.
53. Will CL, Schneider C, MacMillan AM, Katopodis NF, Neubauer G, Wilm M, Lührmann R, Query CC. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J* 2001, 20:4536–4546.
54. Schellenberg MJ, Dul EL, MacMillan AM. Structural model of the p14/SF3b155.branch duplex complex. *RNA* 2011, 17:155–165.
55. Shen H, Green MR. RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. *Nat Struct Mol Biol* 2007, 14:597–603.
56. Shao W, Kim HS, Cao Y, Xu YZ, Query CC. A U1-U2 snRNP interaction network during intron definition. *Mol Cell Biol* 2012, 32:470–478.
57. Golas MM, Sander B, Will CL, Lührmann R, Stark H. Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. *Mol Cell* 2005, 17:869–883.
58. Frilander MJ, Meng X. Proximity of the U12 snRNA with both the 5' splice site and the branch point during early stages of spliceosome assembly. *Mol Cell Biol* 2005, 25:4813–4825.
59. Tarn W-Y, Steitz JA. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* 1996, 273:1824–1832.
60. Dietrich RC, Padgett RA, Shukla GC. The conserved 3' end domain of U6atac snRNA can direct U6 snRNA to the minor spliceosome. *RNA* 2009, 15:1198–1207.
61. Frilander MJ, Steitz JA. Dynamic exchanges of RNA interactions leading to catalytic core formation in the U12-dependent spliceosome. *Mol Cell* 2001, 7:217–226.
62. Incorvaia R, Padgett RA. Base pairing with U6atac snRNA is required for 5' splice site activation of U12-dependent introns *in vivo*. *RNA* 1998, 4:709–718.
63. Yu Y-T, Steitz JA. Site-specific crosslinking of mammalian U11 and U6atac to the 5' splice site of an AT-AC intron. *Proc Natl Acad Sci U S A* 1997, 94:6030–6035.
64. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009, 136:701–718.
65. Shukla GC, Padgett RA. The intramolecular stem-loop structure of U6 snRNA can functionally replace the U6atac snRNA stem-loop. *RNA* 2001, 7:94–105.
66. Valadkhan S, Mohammadi A, Jaladat Y, Geisler S. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci U S A* 2009, 106:11901–11906.
67. Shukla GC, Padgett RA. Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. *RNA* 1999, 5:525–538.
68. Singh J, Padgett RA. Rates of *in situ* transcription and splicing in large human genes. *Nat Struct Mol Biol* 2009, 16:1128–1133.
69. Santoro B, De Gregorio E, Caffarelli E, Bozzoni I. RNA-protein interactions in the nuclei of *Xenopus* oocytes: complex formation and processing activity on the regulatory intron of ribosomal protein gene L1. *Mol Cell Biol* 1994, 14:6975–6982.
70. Pessa HK, Ruokolainen A, Frilander MJ. The abundance of the spliceosomal snRNPs is not limiting the splicing of U12-type introns. *RNA* 2006, 12:1883–1892.
71. Patel AA, McCarthy M, Steitz JA. The splicing of U12-type introns can be a rate-limiting step in gene expression. *EMBO J* 2002, 21:3804–3815.
72. Doma MK, Parker R. RNA quality control in eukaryotes. *Cell* 2007, 131:660–668.
73. König H, Matter N, Bader R, Thiele W, Müller F. Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation. *Cell* 2007, 131:718–729.
74. Steitz JA, Dreyfuss G, Krainer AR, Lamond AI, Matera AG, Padgett RA. Where in the cell is the minor spliceosome? *Proc Natl Acad Sci U S A* 2008, 105:8485–8486.
75. Pessa HKJ, Will CL, Meng X, Schneider C, Watkins NJ, Perälä N, Nymark M, Turunen JJ, Lührmann R, Frilander MJ. Minor spliceosome components are predominantly localized in the nucleus. *Proc Natl Acad Sci U S A* 2008, 105:8655–8660.
76. Friend K, Kolev NG, Shu M-D, Steitz JA. Minor-class splicing occurs in the nucleus of the *Xenopus* oocyte. *RNA* 2008, 14:1459–1462.
77. Hastings ML, Krainer AR. Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. *RNA* 2001, 7:471–482.
78. Wu Q, Krainer AR. Purine-rich enhancers function in the AT-AC pre-mRNA splicing pathway and do so independently of intact U1 snRNP. *RNA* 1998, 4:1664–1673.
79. McNally LM, Yee L, McNally MT. Heterogeneous nuclear ribonucleoprotein H is required for optimal U11 small nuclear ribonucleoprotein binding to a retroviral RNA-processing control element: implications for U12-dependent RNA splicing. *J Biol Chem* 2006, 281:2478–2488.
80. Wu Q, Krainer AR. U1-mediated exon definition interactions between AT-AC and GT-AG introns. *Science* 1996, 274:1005–1008.
81. Verbeeren J, Niemelä EH, Turunen JJ, Will CL, Ravantti JJ, Lührmann R, Frilander MJ. An ancient mechanism for splicing control: U11 snRNP as an

- activator of alternative splicing. *Mol Cell* 2010, 37: 821–833.
82. Chang W-C, Chen Y-C, Lee K-M, Tarn W-Y. Alternative splicing and bioinformatic analysis of human U12-type introns. *Nucleic Acids Res* 2007, 35: 1833–1841.
  83. Hastings ML, Resta N, Traum D, Stella A, Guanti G, Krainer AR. An LKB1 AT-AC intron mutation causes Peutz-Jeghers syndrome via splicing at non-canonical cryptic splice sites. *Nat Struct Mol Biol* 2005, 12:54–59.
  84. Otake LR, Scamborova P, Hashimoto C, Steitz JA. The divergent U12-type spliceosome is required for pre-mRNA splicing and is essential for development in *Drosophila*. *Mol Cell* 2002, 9:439–446.
  85. Borah S, Wong AC, Steitz JA. *Drosophila* hnRNP A1 homologs Hrp36/Hrp38 enhance U2-type versus U12-type splicing to regulate alternative splicing of the prospero twintron. *Proc Natl Acad Sci U S A* 2009, 106:2577–2582.
  86. Scamborova P, Wong A, Steitz JA. An intronic enhancer regulates splicing of the twintron of *Drosophila melanogaster prospero* pre-mRNA by two different spliceosomes. *Mol Cell Biol* 2004, 24: 1855–1869.
  87. Edery P, Marcaillou C, Sahbatou M, Labalme A, Chastang J, Touraine R, Tubacher E, Senni F, Bober MB, Nampoothiri S, et al. Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* 2011, 332: 240–243.
  88. He H, Liyanarachchi S, Akagi K, Nagy R, Li J, Dietrich RC, Li W, Sebastian N, Wen B, Xin B, et al. Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* 2011, 332:238–240.
  89. Pessa HKJ, Greco D, Kvist J, Wahlström G, Heino TI, Auvinen P, Frilander MJ. Gene expression profiling of U12-type spliceosome mutant *Drosophila* reveals widespread changes in metabolic pathways. *PLoS ONE* 2010, 5:e13215.
  90. Wang H, Gao MX, Li L, Wang B, Hori N, Sato K. Isolation, expression, and characterization of the human ZCRB1 gene mapped to 12q12. *Genomics* 2007, 89:59–69.
  91. Zhao E, Li J, Xie Y, Jin W, Zhang Z, et al. Cloning and identification of a novel human RNPC3 gene that encodes a protein with two RRM domains and is expressed in the cell nucleus. *Biochem Genet* 2003, 41:315–323.
  92. Xiong F, Gao J, Li J, Liu Y, Feng G, Fang W, Chang H, Xie J, Zheng H, Li T, He L. Noncanonical and canonical splice sites: a novel mutation at the rare noncanonical splice-donor cut site (IVS4+1A>G) of *SEDL* causes variable splicing isoforms in X-linked spondyloepiphyseal dysplasia tarda. *Eur J Hum Genet* 2009, 17:510–516.
  93. Shaw MA, Brunetti-Pierrri N, Kádasi L, Kováčová V, Van Maldergem L, De Brasi D, Salerno M, Géczy J. Identification of three novel *SEDL* mutations, including mutation in the rare, non-canonical splice site of exon 4. *Clin Genet* 2003, 64:235–242.
  94. Jacobs EY, Frey MR, Wu W, Ingledue TC, Gebuhr TC, Gao L, Marzluff WF, Matera AG. Coiled bodies preferentially associate with U4, U11, and U12 small nuclear RNA genes in interphase HeLa cells but not with U6 and U7 genes. *Mol Biol Cell* 1999, 10:1653–1663.
  95. Tarn W-Y, Yario TA, Steitz JA. U12 snRNA in vertebrates: evolutionary conservation of 5' sequences implicated in splicing of pre-mRNAs containing a minor class of introns. *RNA* 1995, 1:644–656.
  96. Pessa HKJ, Frilander MJ. Minor splicing, disrupted. *Science* 2011, 332:184–185.
  97. Cooper TA, Wan L, Dreyfuss G. RNA and disease. *Cell* 2009, 136:777–793.
  98. Zhang Z, Lotti F, Dittmar K, Younis I, Wan L, Kasim M, Dreyfuss G. SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* 2008, 133: 585–600.
  99. Boulisfane N, Choleza M, Rage F, Neel H, Soret J, Bordonné R. Impaired minor tri-snRNP assembly generates differential splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. *Hum Mol Genet* 2011, 20:641–648.
  100. Gabanella F, Butchbach MER, Saieva L, Carissimi C, Burghes AHM, Pellizzoni L. Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. *PLoS ONE* 2007, 2:e921.
  101. Lynch M, Richardson AO. The evolution of spliceosomal introns. *Curr Opin Genet Dev* 2002, 12:701–710.
  102. Basu MK, Rogozin IB, Koonin EV. Primordial spliceosomal introns were probably U2-type. *Trends Genet* 2008, 24:525–528.
  103. Roy SW, Irimia M. Splicing in the eukaryotic ancestor: form, function and dysfunction. *Trends Ecol Evol* 2009, 24:447–455.
  104. Basu MK, Makalowski W, Rogozin IB, Koonin EV. U12 intron positions are more strongly conserved between animals and plants than U2 intron positions. *Biol Direct* 2008, 3:19.
  105. Abril JF, Castelo R, Guigo R. Comparison of splice sites in mammals and chicken. *Genome Res* 2005, 15:111–119.
  106. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* 2007, 446:926–929.
  107. Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares

- M Jr. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* 2007, 21:708–718.
108. Saltzman AL, Kim YK, Pan Q, Fagnani MM, Maquat LE, Blencowe BJ. Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol Cell Biol* 2008, 28:4320–4330.
109. Saltzman AL, Pan Q, Blencowe BJ. Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev* 2011, 25:373–384.
110. Huelga SC, Vu AQ, Arnold JD, Liang TY, Liu PP, Yan BY, Donohue JP, Shiue L, Hoon S, Brenner S, et al. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep* 2012, 1:167–178.