Cell PRESS

Widespread and subtle: alternative splicing at short-distance tandem sites

Michael Hiller[1] and Matthias Platzer[2]

[1] Bioinformatics Group, Albert-Ludwigs-University Freiburg, 79110 Freiburg, Germany
[2] Genome Analysis, Leibniz Institute for Age Research–Fritz Lipmann Institute, 07745 Jena, Germany

Alternative splicing at donor or acceptor sites located just a few nucleotides apart is widespread in many species. It results in subtle changes in the transcripts and often in the encoded proteins. Several of these tandem splice events contribute to the repertoire of functionally different proteins, whereas many are neutral or deleterious. Remarkably, some of the functional events are differentially spliced in tissues or developmental stages, whereas others exhibit constant splicing ratios, indicating that function is not always associated with differential splicing. Stochastic splice site selection seems to play a major role in these processes. Here, we review recent progress in understanding functional and evolutionary aspects as well as the mechanism of splicing at short-distance tandem sites.

Alternative splicing as a source of proteome complexity

A fundamental difference between prokaryotic and eukaryotic gene structures is that the typical eukaryotic gene contains introns [1]. In the splicing process, these introns are excised from the primary transcript while the remaining sequences (exons) are joined together (Box 1). Alternative splicing enables one gene to produce multiple mature transcripts with different sequences. Large-scale characterization of eukaryotic transcriptomes using expressed sequence tags (ESTs), cDNAs, and splicing sensitive microarrays revealed that alternative splicing is frequent [2], with 75% of the human genes estimated to undergo alternative splicing [3]. Alternative splicing (along with other mechanisms such as alternative transcriptional initiation, alternative polyadenylation, RNA editing) is believed to be a major mechanism to bridge the gap between the surprisingly small gene number and the much higher transcript and protein number in higher eukaryotes [4].

Alternative splice variants can have biological roles by giving rise to functionally different proteins [5], and the regulation of alternative splicing is important for diverse biological processes. For example, the brain-specific splicing factor Nova regulates alternative splicing of several genes with neuronal functions [6]. In *Drosophila*, alternative acceptor use of the *tra* gene is the beginning of a regulatory cascade that leads to male or female development [7]. Furthermore, alternative splicing can regulate gene expression at a posttranscriptional level by producing transcripts that are subject to the surveillance mechanism nonsense-mediated decay (NMD) that degrades mRNAs with stop codons >50 nucleotides (nt) upstream of the last exon–exon junction [8,9]. Defects in splicing are associated with cancer [10,11] and other human diseases [12].

There are several types of alternative splice events, which vary in frequency. In vertebrates, the inclusion or skipping of entire exons is the most frequent alternative splice event, whereas in plants, the retention of entire introns is most prevalent [13,14]. Intensive research has focused on exon skipping, and progress has been made in understanding the function, evolution, and characteristics of such exons and in distinguishing them from constitutive exons. Another major mode of alternative splicing is the selection of two or sometimes more alternative splice sites. In humans and mice, alternative donors and acceptors each make up >25% of all alternative splice events [15]. In contrast to exon skipping or intron retention, such splice events are roughly equally frequent in vertebrates, invertebrates and plants [14].

In recent years, several studies showed that many alternative donor and acceptor sites are close to each other, thus producing mRNAs that differ by a few nucleotides [2,16–19]. Such splice site pairs are called tandem sites (Figure 1; please note that we use the notation 'tandem site/donor/acceptor' for alternatively spliced sites and the notation 'tandem motif' for those that currently lack experimental evidence for the use of both splice sites). Here, we discuss recent progress in characterizing the frequency, the splicing mechanism and the functional and evolutionary implications of tandem sites. It should be noted that U12-dependent introns also produce subtle splice events [20], but we focus on U2-dependent introns because they represent the majority of introns. For reviews on other aspects of alternative splicing, please see Refs [21–24].

Widespread subtle alternative splicing

Although alternative splicing at tandem sites was observed only a few years after the discovery of introns, the extent of this phenomenon was only recently appreciated. Computational analysis using large EST databases found that one third of all alternative donor and acceptor pairs are located 2–10 nt apart [2,19]. Such short-distance alternative acceptor events are about twice as frequent as alternative donor events [16,18,19]. Of all alternative acceptors, Δ3 (nucleotide distance between two splice sites is denoted with Δ nt; Figure 1) sites having a NAGNAG motif (N = A,C,G,T) are by far most frequent in plants [13]

Corresponding author: Hiller, M. (hiller@informatik.uni-freiburg.de).

 0168-9525/$ – see front matter © 2008 Elsevier Ltd. All rights reserved. doi:10.1016/j.tig.2008.03.003 Available online 3 April 2008

Box 1. Introns and splicing

Introns possess several sequence signals required for splicing: the 5′ or donor splice site, the branch point, the polypyrimidine tract and the 3′ or acceptor splice site (Figure I). We prefer here the notation donor and acceptor because 5′ and 3′ splice site can be ambiguous if there are alternative sites. Apart from rare exceptions [75,76], the 5′ terminal intronic dinucleotide is GT or sometimes GC (<1%) and the 3′ terminal dinucleotide is always AG.

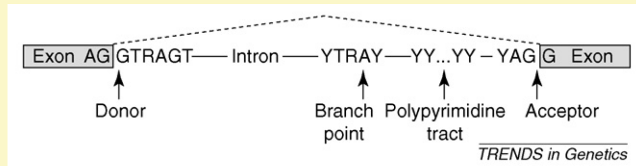

Figure I. Schematic illustration of basic splicing signals. Splicing is carried out by the spliceosome, which consists of several small nuclear (sn)RNAs (U1, U2, U4, U5 and U6) and numerous proteins. Splicing proceeds in two steps. First, the mRNA is cleaved at the donor splice site and the 5′ intron end is attached to the branch point adenosine, forming a lariat structure. Second, the mRNA is cleaved at the acceptor site, both exons are ligated, and the intron is released. During the first step, the U1 snRNA initially binds to the donor site by base pairings to the last three exonic and first six intronic positions and is later displaced by the U6 snRNA [77]. The 3′ intron end is recognized by a heterodimer U2AF. The U2AF65 subunit binds to the polypyrimidine tract and the U2AF35 component binds to the YAG (Y = C,T). The U2 snRNA binds to the branch point by base pairings, and this process is aided by protein factors. The recognition of these basic splicing signals is influenced by splicing factors such as serine-arginine–rich proteins and heterogeneous nuclear ribonucleoproteins that bind to sequence motifs in exons and introns [23]. These splicing factors can have enhancing or silencing effects and are essential for alternative and constitutive splicing.

and animals [2,17–19,25], with almost 2000 alternatively spliced acceptors observed in humans [26]. At the donor site, alternative Δ4 donor events are most frequent [13,19,27]. This is because of the intronic donor consensus sequence GTRAGT (R = A,G) that provides a second GT 4 nt from the donor site in ~40% of introns [27]. Accurate identification of such subtle splice events from noisy EST data requires stringent filtering of the EST-to-genome alignments [2,17]. Experimental validation using RT–PCR succeeded in identifying alternative splicing in 92% of 220 tested NAGNAG acceptors [25], suggesting that alternative splicing at tandem sites can be accurately identified using ESTs.

The frequency of donor and acceptor events decreases with increasing distance between the two splice sites. At short distances, Δ4 and Δ5 events outnumber the Δ6 cases, even though the frameshifted transcripts arising by Δ4 and Δ5 events are less likely to be detected, because NMD reduces their abundance [19]. This shows that potential splice sites at a close distance to the dominant site are very prone to alternative splicing [16,19]. At larger distances, the greater fraction of frame-preserving events is likely caused by the suppression of frameshifted transcripts by NMD [19].

Tandem splice events are found in species in which alternative splicing is frequent, including mammals, chicken, zebrafish, *Drosophila* and plants [13,17,28]. The only exception found thus far is the nematode *Caenorhabditis elegans*, where alternative splicing at NAGNAG motifs is very rare. In *C. elegans*, the intron–proximal acceptor is exclusively used in almost all NAGNAG motifs, and only 33 alternative cases (constituting <1% of all NAGNAG motifs) are found [17,28]. This seems to be caused by a unique acceptor selection mechanism, because the *C. elegans* U2AF dimer binds to a tightly restricted TTTTCAG motif, and deviations from this consensus affect binding [29]. This restricted acceptor motif may prevent recognition of an alternative intron–distal AG, indicating that flexibility in U2AF binding is a crucial step for alternative splicing at tandem acceptors with the potential for regulation. In contrast to NAGNAG sites, alternative splicing at Δ3 donors was observed in *C. elegans* with a frequency similar to other species [28].

Insights into the tandem splice mechanism

Experimental studies confirmed that alternative donor or acceptor splice events are the cause of the observed subtle splice variants. Disruption of one of the two splice sites by a targeted single nucleotide mutation [25,30,31] or a naturally occurring polymorphism [32–35] leads to the disappearance of the respective splice variant in all investigated cases. This excludes very short exons or highly similar mutually exclusive exons being a frequent cause of subtle alternative transcripts.

The two splice sites in alternative tandems generally resemble normal splice sites in terms of their sequence preference. That is, both alternative donor sites often have a reasonable base pairing potential to the U1 small nuclear RNA (snRNA) and both alternative acceptors mostly have YAG motifs (where Y = C,T), which are preferred U2AF35 binding sites (Box 1) [19,28,36]. For example, one half of all CAGCAG acceptor motifs in humans are known to be alternatively spliced [17]. Furthermore, basically all of the 16 possible NAGNAG motifs allow alternative splicing except for those having a GAG acceptor [17,35], consistent with experimental observations that GAG is a very inefficient acceptor [37]. This indicates that the capability of alternative splicing at tandem sites is an intrinsic property of the spliceosome.

We know from computational studies that the NAG sequence and its local sequence context are the most important features to distinguish alternative from constitutive NAGNAG acceptors [16,38]. Features derived from the local context (e.g. strength of the splice sites, the polypyrimidine tract and the branch point, as well as the distance between them) enable the prediction of alternative splice sites using machine learning methods [38,39]. The relative strength of both splice sites in a tandem is also predictive for the predominantly used site [16,19,27,28,40–42]. However, for NAGNAG acceptors, experiments have shown that the ratio of splicing at the intron–proximal versus the distal site also depends on the intronic sequence between the branch point and the NAGNAG acceptor [35]. Tsai *et al.* [35] switched the intronic context between NAGNAG acceptors that are predominantly spliced at the proximal and at the distal acceptor. The recombinant NAGNAG tandem with the heterologous intron context resembled the splicing ratio of the NAGNAG tandem from which the intron sequence was taken. However, neither the sequence motifs responsible for a specific ratio nor the splicing factors involved are known. At larger distances, the location and strength of the polypyrimidine tract is important for alternative acceptors [38]. This is

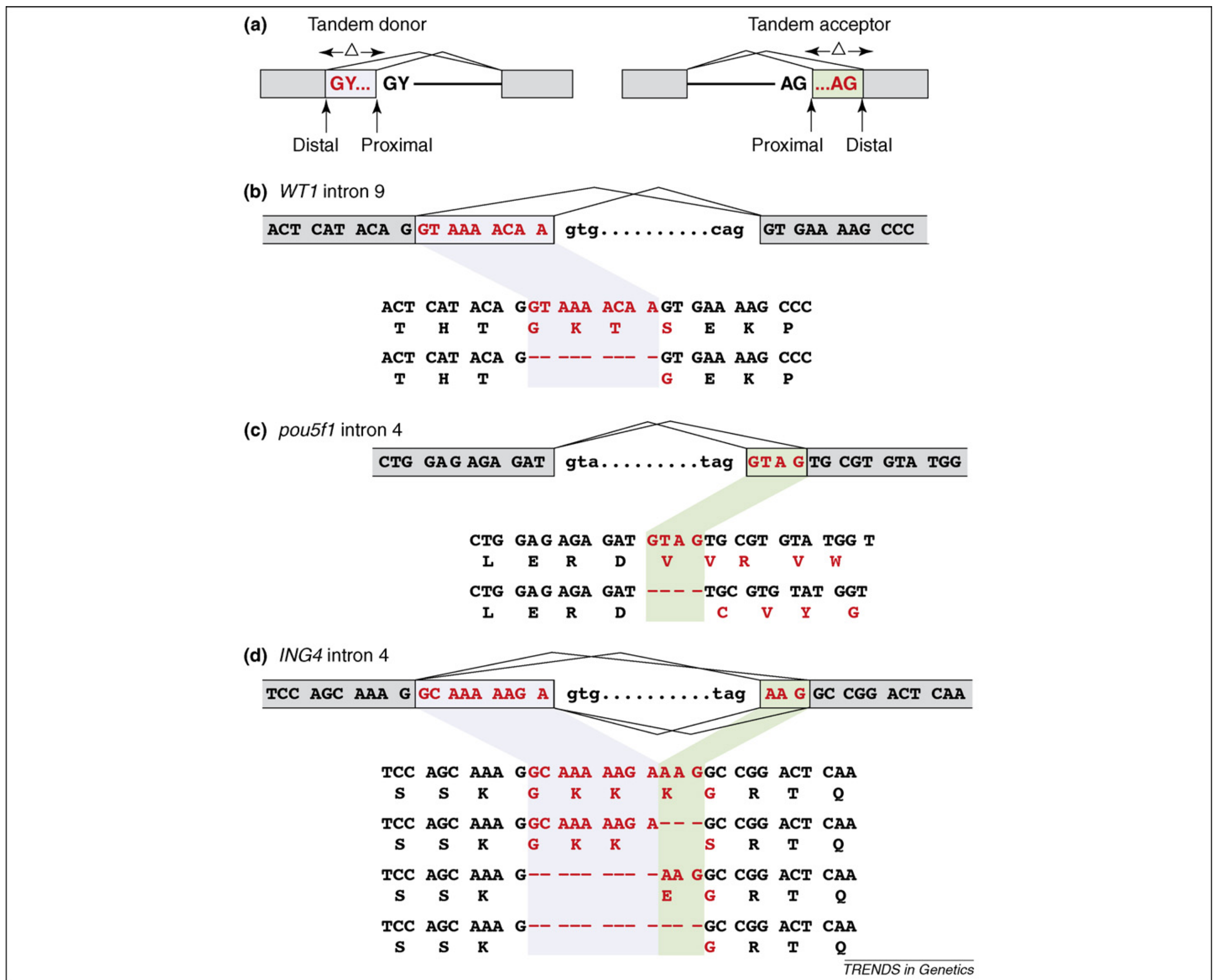

Figure 1. Examples of alternative splicing at short-distance tandem sites. **(a)** Schematic illustration of tandem donors and acceptors. **(b)** A Δ9 donor in human *WT1* results in the indel of three amino acids (KTS). **(c)** A Δ4 acceptor in the last intron of zebrafish *pou5f1* leads to a frameshift and a different C terminus of the protein. **(d)** Intron 4 in the human *ING4* gene exhibits a Δ9 donor and a Δ3 acceptor [31,34]. The combination of both tandem sites leads to four different transcript and protein variants. The partial exon–intron structure is shown in boxes; missing residues in the transcript sequences are indicated by dashes.

consistent with the finding that at distances of >8 nt, the sequence between the two AGs resembles a second polypyrimidine tract [19]. Thus, at shorter distances, only one polypyrimidine tract seems to be sufficient, whereas at larger distances, two polypyrimidine tracts are needed.

Tandem splicing ratios: constant and variable

Alternative splicing is a highly regulated process with the potential to express a splice variant only under specific spatio-temporal conditions. For example, alternative acceptor use of the *Drosophila tra* gene is controlled in a sex-specific manner [7]. Experimental studies showed that the ratio between two splice variants at short-distance tandem sites may differ between tissues or cell types [17,25,43,44], between developmental stages [45–47] or with respect to external stimuli [47]. Splicing factors such as serine-arginine–rich (SR) proteins and heterogeneous nuclear ribonucleoprotein (hnRNPs) regulate alternative splicing [23], and these proteins are candidates to regulate tandem splicing by promoting intron–proximal or distal splice sites, respectively (Box 2). Although the short distance between two tandem sites largely excludes the possibility that splicing regulatory motifs are placed between them [48], the up- and downstream regions can contain potential splicing factor binding sites, as shown for Δ3 tandem sites [28,36]. In case of a Δ3 acceptor involving a noncanonical TG dinucleotide in the GNAS complex locus gene, SF2/ASF and hnRNP A1 were shown to affect the ratio between the proximal and distal site [49]. It was also suggested that alternative donor/acceptor use is associated with splicing motifs that are distinct from those that regulate exon skipping [42]. Because the position of mRNA cleavage at the acceptor site is ultimately determined in the second splicing step (Box 2), variations in the concentration of second-step splicing factors can lead to variations in the splicing ratio, as shown for hSlu7 [50,51]. However, the details of the underlying mechanism(s) for these variable splicing ratios are currently unclear.

Box 2. Mechanisms of alternative donor and acceptor selection

Donor selection

A widely accepted mechanism for alternative donor splicing is the differential binding of the U1 snRNA to one of the potential donor sites. According to this splice site competition model, alternative splicing happens when one donor is sufficiently good to compete with the other donor for U1 binding. Constitutive splicing at a tandem motif (exclusive selection of only one donor) occurs when either donor is much stronger and consequently outcompetes the other.

Apart from the intrinsic strength of donor sites, splicing factors such as serine-arginine–rich (SR) proteins and and heterogeneous nuclear ribonucleoproteins (hnRNPs) affect splice site selection. SF2/ASF and other SR proteins promote splicing at the intron–proximal donor site, whereas hnRNP A1 promotes the distal site [78]. The relative concentration of SR proteins and hnRNPs affects donor selection, and tissue-specific variations in this ratio might lead to tissue-specific splicing patterns [78].

U1 snRNA binding is almost always required for splicing. One exception is an alternative donor site in *FGFR1*, where the U6 snRNA is responsible for splicing at a noncanonical GA donor site [79]. Also, *in vitro* experiments have shown that splicing can proceed without U1 when the donor site allows a sufficiently stable U6 snRNA binding [77]. However, these cases seem to be rare because computational studies found no evidence that U6 snRNA activates alternative donor sites [19].

Acceptor selection

In vitro experiments found evidence for different modes of acceptor AG selection that depend on the distance of the AG to the branch point. If the branch point is more than ~20–35 nucleotides (nt) away from the AG, the AG selection occurs by a scanning mechanism that starts from the branch point and usually selects the intron–proximal AG [37,80]. The proximal AG can be bypassed if it is too close to the branch point or if it is in competition with a more distal AG. This competition can lead to alternative splicing and depends on (i) the distance between the AGs (shorter distances lead to a higher competition), (ii) the nucleotide upstream of the AGs (C and T are preferred over A and especially over G) and (iii) the sequence between both AGs [19,37,80,81]. Scanning does not occur if the distance to the branch point is short (<20 nt) [80]. In these cases, a distal AG can efficiently compete with a proximal AG given the distance between both AGs <6 nt [81].

However, initial selection of one AG does not ultimately determine the site of exon ligation. It was shown that a proximal AG can be necessary for mRNA cleavage at a distal AG in the second splicing step [81] and vice versa [82]. Splicing step II is subject to proofreading mechanisms, and several proteins are involved in selecting the ultimate AG. For example, changing the concentration of hSlu7 affects the splicing ratio of two acceptors [50,51]. The second step splicing factor SPF45 promotes the use of a proximal AG despite the presence of a polypyrimidine tract and an AG located downstream, which are bound by the U2AF heterodimer [82].

Similar to donor selection, SR proteins were shown to promote proximal acceptor splice sites, whereas hnRNP A1 promotes distal sites [83]. Thus, although alternative splicing is often regulated at the early splicing step, alternative acceptor selection can be regulated at the early and the late step [19,82].

In other experimentally well-documented cases, the tandem splicing ratio is rather constant in different tissues [25,31,34,52–54], developmental stages [25,30,34,55–57], as a response to external stimuli [31,53], and in different disease states [34,55] or with respect to other factors such as sex and ethnic background [55]. For example, the four splice variants of the *ING4* (inhibitor of growth family, member 4) gene that arise by a Δ9 donor and a Δ3 acceptor (Figure 1d) have a similar ratio in different tissues and cell lines, which is unaffected by serum starvation, DNA damage or overexpression of SF2/ASF [31,34]. Likewise, the ratio of the Δ9 tandem donor variants of the human and mouse *WT1* (Wilms tumor 1) gene (Figure 1b) are constant in tissues, cell lines, and during development [30,58].

The mechanism behind such constant splicing ratios is also not well understood. The splice site competition model (Box 2) suggests that two donors compete for U1 snRNA binding and two acceptors for U2AF35 binding, which leads to interference between the overlapping sites. If U1 snRNA and U2AF are critical factors that enable tandem splicing, this might explain constant splicing ratios because both factors are ubiquitously expressed. This competition model is supported by the finding that the splice site with the higher affinity (inferred from the splice site sequence) is generally the predominantly used site [16,19,27,28,40–42]. For NAGNAG tandems, the splice site strength is predictive for alternative or constitutive splicing and also correlated with the preferred acceptor [16]. Further support for this competition model comes from the observed independence of splicing at NAGNAG acceptors in two adjacent introns in the *MLLT4* (myeloid/lymphoid or mixed-lineage leukemia) gene [33] or at tandem donor and acceptor sites within the same intron of *ING4* (Figure 1d) [31]. Thus, stochastic splice site selection based on spliceosomal core components might represent a simple mechanism for producing constant splicing ratios.

However, this competition model has some limitations because it does not explain observed tissue- and/or developmental-specific variations in the splicing ratio or cases where a weaker splice site is predominantly used. It is conceivable that additional factors (Box 2) bias the outcome of an underlying stochastic mechanism of splice site selection under specific conditions.

Functional implications of short-distance tandem sites

The impact of tandem sites affecting the protein coding sequence depends on whether they are frameshifting (Δ2, 4, 5, 7, 8, etc.) or frame-preserving (Δ3, 6, 9, etc.). Frameshifting tandems have severe effects for protein function because they result in proteins with a different C terminus or mRNA degradation through the NMD pathway [16,19,27] (Figure 2). By contrast, the protein effects of frame-preserving tandems are mostly subtle but highly diverse. Most of them lead to the insertion or deletion (indel) of a few amino acids, but the exchange of an oligopeptide for an unrelated residue is also observed (Figure 2). This increases the proteome diversity by producing highly similar protein isoforms of many genes [17]. Further diversity emerges from the independent usage of a tandem donor and acceptor in one intron [28], as described for *ING4* [31,34], or the existence of more than one tandem site in one gene [17]. Apart from subtle effects, a few frame-preserving tandems have a drastic effect by creating a stop codon [17,28].

Several experimental studies have shown that subtle splice events can yield functionally different proteins. The functional implications range from differences in protein, DNA or RNA binding, modifying subcellular localization to modulating ion channel and receptor properties (Table 1).

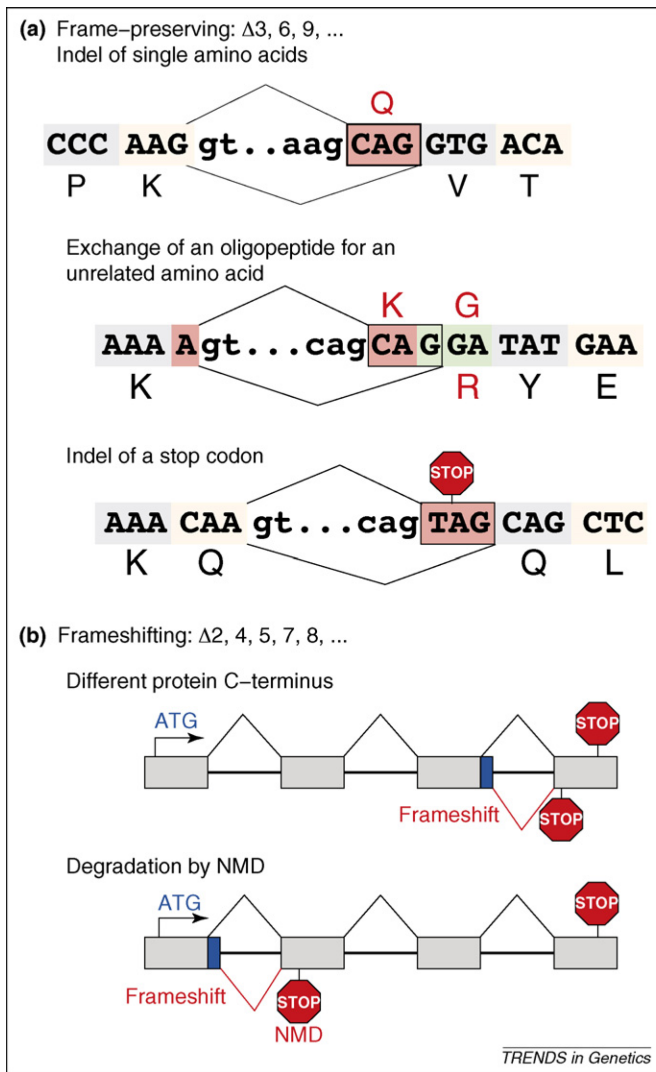

Figure 2. Effects of alternative tandem splicing on proteins. Depending on the distance between the two splice sites, one can distinguish between **(a)** frame-preserving and **(b)** frameshifting tandem sites. The protein effects of frame-preserving tandems is illustrated for NAGNAG acceptors and make up the indel of a single amino acid [shown for the insulin-like growth factor 1 receptor (*IGF1R*) gene, intron 13], the exchange of an oligopeptide for another amino acid [paired box 3 (*PAX3*) gene, intron 2] and the indel of a stop codon [protein phosphatase, EF-hand calcium binding domain 2 (*PPEF2*) gene, intron 5]. Frameshifting tandems can lead to a protein with a different C terminus or mRNA degradation by nonsense-mediated decay if the premature stop codon occurs upstream of the last exon–exon junction. It should be mentioned that the original reading frame can be rescued by a second frameshifting splice event, but such events are rare.

For example, the four different protein variants of *ING4* differ in several functional aspects such as protein binding, transcriptional activation and the ability to suppress cell growth, cell spreading and migration [31]. Furthermore, these splice events affect a nuclear localization signal and alternative splicing results in a switch from nuclear to cytoplasmic protein localization. The biological role of these alternative splice events seems to be conserved because the tandem donor and acceptor splice variants exist in mouse, rat and pig [31].

Another striking example is a Δ9 tandem donor in the transcription factor *WT1* that leads to the indel of three amino acids, Lys-Thr-Ser (KTS). This tandem donor together with its flanking regions is almost perfectly conserved in vertebrates, with only a single mutation found in fugu [30]. Both protein variants differ in their nuclear localization pattern and their affinity to bind DNA, RNA and proteins, which has consequences for transcriptional regulation [58]. Experiments with hetero- or homozygous knockout mouse mutants, where either donor site was destroyed, demonstrated distinct functions for both splice variants in kidney and gonad development, as well as male sex determination [30]. Homozygous mouse mutants died within 1 d after birth.

The human ectodysplasin A (*EDA*) gene provides an example where the indel of two amino acids caused by splicing at a conserved Δ6 donor results in distinct receptor binding specificity [46]. These two amino acids lead to a remodeled receptor-binding site, such that the longer protein binds only to the ectodysplasin A receptor, whereas the shorter variant binds exclusively to the ectodysplasin A2 receptor. This affinity switch is of interest because both receptors have distinct functional domains and diverge in their signaling pathways. Apart from frame-preserving tandems, a function for a frameshifting Δ4 acceptor was proposed for the zebrafish transcription factor *pou5f1* [45] (Figure 1c). Furthermore, tandem splice sites in the 5′ untranslated region can have a function by affecting the translational efficiency [59]. As for other alternative splice events, many details about the biological role of splicing at short-distance tandems in cells and organisms remain to be determined.

By contrast, for several tandem splice variants, functional assays revealed no differences [47,53,60]. This shows that small variations can leave protein function largely unaffected, making it likely that such events are evolutionary neutral and thus simply tolerated by the cell.

Subtle splice events can also be associated with diseases. In the *WT1* gene, human mutations that disrupt the proximal donor site lead to a deviation in the defined +KTS/−KTS variant ratio and are associated with Frasier syndrome causing glomerulosclerosis, abnormal gonad development and XY sex reversal [61]. Knockout mice where the proximal donor site is destroyed resemble the human phenotype of Frasier syndrome [30]. This shows that deviations in the +KTS/−KTS ratio are highly deleterious. In another case, a mutation in the *ABCA4* gene that changes the acceptor sequence from TAGGAG to TAGCAG has much higher frequency in patients with macular degeneration (Stargardt disease 1) [62]. The effect of this mutation is twofold: (i) it changes a Gly to an Ala and (ii) it activates the distal acceptor resulting in the lack of the CAG in ~50% of the patient's transcripts. Together with other more severe *ABCA4* mutations, this mutation may cause Stargardt disease [62]. Finally, it is noteworthy that the frequency of alternative donor and acceptor events is increased in cancer tissues, whereas the frequency of exon skipping events is reduced [11].

Functional tandem splice events with constant ratios

Differential regulation of alternative splice events is often considered to be an indication for function. In this respect, it is striking that several of the tandem sites with clear functional implications exhibit remarkably constant splicing ratios (Table 1) [25,31,52,54–57]. In fact, for some splice events, a deviation in the splicing ratio is deleterious

Table 1. Examples of short-distance tandem splice sites with functional implications

Species	Gene[a]	Splice site distance (Δ)	Splice site pattern (| = intron–exon boundary)	Functional consequences of alternative splicing	Variation in splicing ratio observed	Refs
Tandem donor						
Human	*NR3C1*	3	|GTA|GTG	Transcriptional activation	No	[54]
Mouse	*ALDH18A1*	6	|GTAAAT|GTG	Sensitivity versus insensitivity to ornithine inhibition	Yes	[43]
Human	*EDA*	6	|GTAGAA|GTG	EDAR versus XEDAR receptor binding specificity	Yes	[46]
Mouse	*FGFR1*	6	|GTAACA|GAA[b]	Signaling through the Ras/MAPK pathway by modifying the binding of adaptor protein FRS2	No	[52]
Human, mouse	*WT1*	9	|GTAAAACAA|GTG	Localization within the nucleus; binding to U2AF65, RNA, DNA; transcriptional regulation of *SRY*, *NR5A1*, *BCL2*, *NR0B1*, *AMH*; kidney and gonad development and male sex determination	No	[58][e]
Tandem acceptor						
Human	*ATN1*	3	CAG|CAG|	Nuclear versus cytoplasmic localization	No	[25]
Human	*IGF1R*	3	CAG|CAG|	Signaling activity and the receptor internalization rate	no	[84]
Mouse	*GGT1*	3	CAG|CAG|	Translational efficiency	Unknown	[59]
Arabidopsis	AT2G43370	3	TAG|CAG|	Binding to SR proteins and U11 snRNA	Unknown	[85]
Mouse	*PAX3*	3	AAG|CAG|	DNA binding and transcriptional activation	No	[57]
Human	*SCN5A*	3	CAG|CAG|	Current density and recovery time of the sodium channel	No	[55]
Human	*NUDT4*	3	AAG|CAG|	Catalytic activity toward diphosphoinositol and diadenosine polyphosphates	Unknown	[86]
Rat	*CACNA1A*	3	TTG|TTG|[c]	Inactivation kinetics of calcium channel	Yes	[87]
Zebrafish	*pou5f1*	4	TAG|GTAG|	Protein reading frame which leads to loss of DNA binding caused by a truncated homeo domain and developmental arrest in overexpression experiments	Yes	[45]
Drosophila	*acj6* (CG9151)	6	AAG|AAAAAG|	DNA binding, binding to Cf1-a, switch between transcriptional activator and inhibitor	No	[56]
Mouse	*IRF2*	6	AAG|TTGTAG|	Transcriptional regulation	Yes	[44]
Tandem donor...acceptor[d]						
Human	*ING4*	9...3	|GCAAAAAGT|GTG...TAG|AAG|	Nuclear versus cytoplasmic localization; binding to Liprin α1 and G3BP2; transcriptional activation of the p21^{WAF1} promoter; cell growth, spreading and migration	No	[31]

[a]Full names for the genes are as follows: *NR3C1*, nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor); *ALDH18A1*, aldehyde dehydrogenase 18 family, member A1; *EDA*, ectodysplasin A; *FGFR1*, fibroblast growth factor receptor 1; *WT1*, Wilms tumor 1; *ATN1*, atrophin 1; *IGF1R*, insulin-like growth factor 1 receptor; *GGT1*, gamma-glutamyltransferase 1; *AT2G43370*, U1 small nuclear ribonucleoprotein 70 kDa; *PAX3*, paired box 3; *SCN5A*, sodium channel, voltage-gated, type V, α subunit; *NUDT*, nudix (nucleoside diphosphate linked moiety X)-type motif 4; *CACNA1A*, calcium channel, voltage-dependent, P/Q type, α 1A subunit; *pou5f1*, POU domain, class 5, transcription factor 1; *acj6*, abnormal chemosensory jump 6; *IRF2*, interferon regulatory factor 2; *ING4*, inhibitor of growth family, member 4.

[b]Use of a noncanonical GA donor.

[c]Use of noncanonical TG acceptors.

[d]Tandem donor and acceptor in the same intron.

[e]And references therein.

and leads to disease [12], exemplified by the *WT1* Δ9 donor (see above) [30]. These constant splicing ratios can be advantageous in situations where alternative splicing simply increases the functional diversity of proteins and where both functional protein isoforms are required ubiquitously. As mentioned above, it is tempting to speculate that expression of two functionally important splice variants in a constant ratio might be achieved just by a stochastic mechanism in which the two splice sites compete for binding to spliceosomal core components. If so, these cases would represent further examples of 'cultivated noise' in biology [63]. In this context, it is noteworthy that stochastic splicing of 48 mutually exclusive exons in *Drosophila DSCAM* is needed for proper axon guidance [64], and stochasticity in gene expression is involved in cell fate divergence and cell individualization [63].

Furthermore, some studies found no function for alternative splice events that are produced in variable ratios. For example, assays showed no functional difference between Δ9 acceptor variants in *Arabidopsis PIMT2*, despite changes in the splicing ratio during development and under hormone treatment [47]. Similarly, tissue- and development-specific skipping of a NAGNAG acceptor containing exon in mouse *Psap* has no detectable consequences in a knockout mouse lacking this alternative exon [65]. Although it remains possible that these events have a yet undetected function, these findings suggest that variable splicing ratios do not always imply functional importance. In analogy, the tissue-specific expression of a gene does not always imply a tissue-specific function [66].

Selection for and against tandem sites in evolution

Several computational studies investigated conservation of tandem sites in other species because conservation implies purifying (negative) selection and thus function [16–18,27,28,36]. For NAGNAG acceptors, initial studies reached conflicting conclusions [16,17,67] that are most likely caused by an unbalanced distribution of NAGNAG motifs together with different conservation levels for specific motifs [68]. Controlling for these biases, a fraction of tandem sites was found to be under purifying selection, even for large evolutionary distances such as between human and

chicken or fish. Remarkably, both frame-preserving and frameshifting tandems are under selection [68]. A hallmark of conserved alternative exons is that their upstream and downstream intron flanks are more conserved than the intron flanks around constitutive exons. Similarly, for exons with a tandem site, only the intron flank that is associated with the alternative splice event exhibits a higher conservation [40]. With respect to short-distance tandems, this finding is currently controversial because some studies found a higher intronic conservation for alternatively spliced tandems in comparison to constitutively spliced sites [28,36,68], whereas others did not [40]. However, differences in datasets and methods do not enable a direct comparison of these results. In summary, although it is mostly unknown which tandem sites are functionally important, the conservation of tandem splice events between distant species (as observed for *EDA*, *WT1* or *PAX3*; Table 1) and/or strong conservation of the intron flank (such as in *WT1*) seem to be reasonable criteria to select promising candidates for further experimental efforts.

Although many tandem splice events might have little or no effect on the function of the transcript and protein and are probably tolerated, other events can be deleterious and therefore expected to be eliminated by natural selection. Consistent with the latter assumption, Δ3 tandem sites are significantly underrepresented in protein coding compared with untranslated regions [16,69]. This might be caused by selection against mutations that create a NAGNAG motif that allows alternative splicing. Indeed, ≥60% of these mutations are thought to have been eliminated in coding regions [69], suggesting that even the indel of a single triplet into the coding sequence can be deleterious to protein function.

Furthermore, the strength of selection in the coding sequence is not uniform, because protein regions that fold into a fixed three-dimensional structure show a stronger avoidance of frame-preserving tandem acceptor events [69]. This agrees with observations that tandem acceptors are preferentially located at the protein surface and avoided within protein secondary structure elements such as α-helices or β-sheets. Similar results were also reported for other alternative splice events [70,71]. A preferred location of tandem acceptors at the protein surface is also in line with the enrichment of polar amino acids in their flanking protein regions [17,69]. Moreover, the protein effects of Δ3 tandems are enriched in single amino acid indels [17,25,28], indicating that such events are often more compatible with protein function than events that affect dipeptides (Figure 2).

Emergence of tandem sites in evolution

One can basically distinguish two main mechanisms for the emergence of alternative splicing at a tandem site. The first mechanism involves the emergence of a novel splice site around a constitutive exon [40,42]. For example, a single nucleotide exchange can result in a NAGNAG

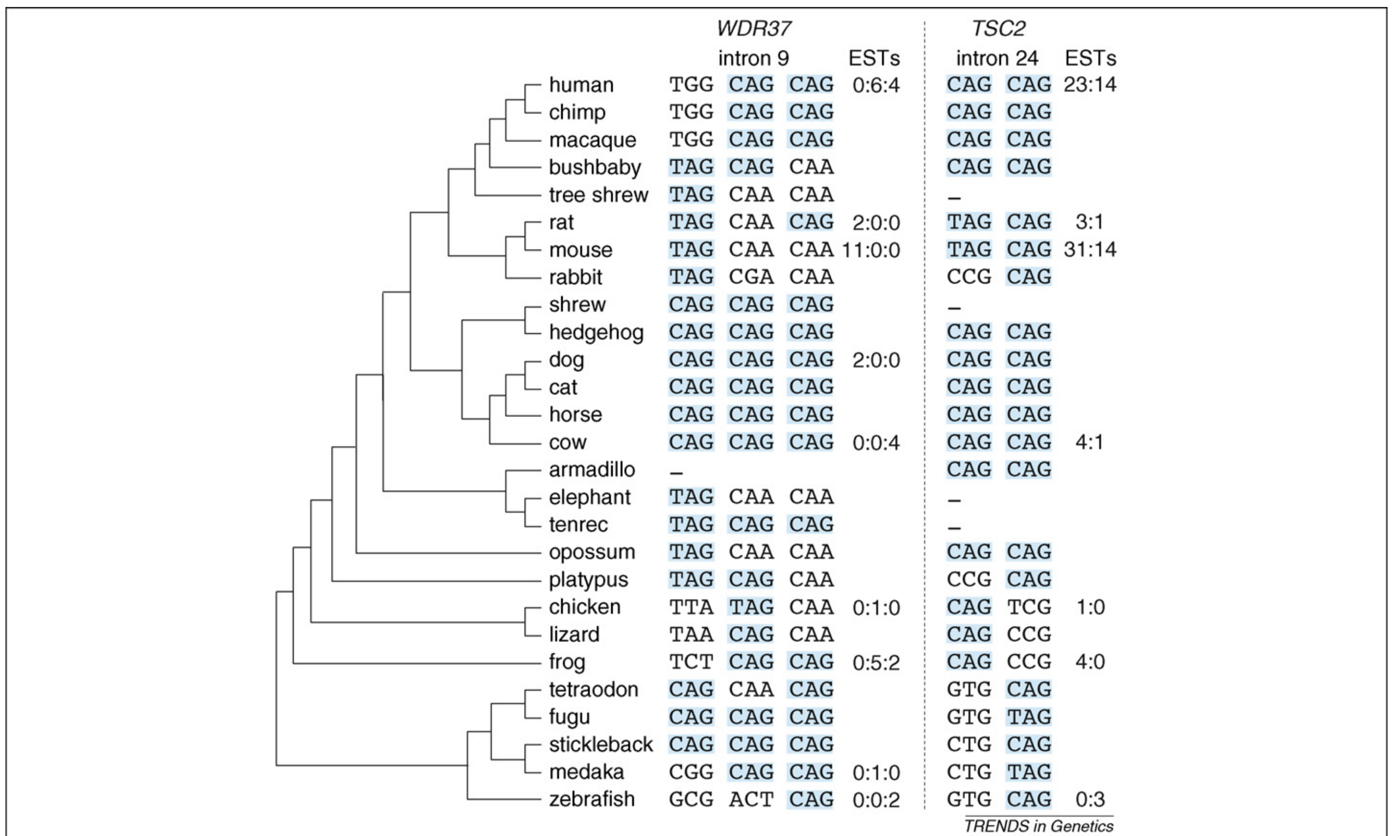

Figure 3. Multiple and independent gains or losses of short-distance tandem acceptor NAG motifs in evolution. The phylogenetic vertebrate tree shows that the acceptor AGs were created or lost several times in independent evolutionary lineages [such as in mouse and opossum for *WDR37* (WD repeat domain 37) or in rabbit, platypus and fish for *TSC2* (tuberous sclerosis 2)]. These acceptor variations result in a change of the exon length of 3 nucleotides (nt) (*TSC2*) or up to 6 nt (*WDR37*) as indicated by available expressed sequence tag counts. In case of *WDR37*, all eight possible combinations of the three acceptor NAG motifs are observed with alternative splicing observed in human and frog. NAG triplets that may function as an acceptor site are highlighted in blue. Branch lengths are not drawn to scale.

acceptor that is alternatively spliced [32]. Often, the ancestral splice site remains the preferred site, presumably because of purifying selection maintaining the ancestral transcript and protein isoform and because the sequence context of the novel site is initially not optimized for efficient splice site selection [40]. If the new splice site is located in the intron, the region that extends the exon shows rapid evolution of the protein coding sequence [40], consistent with the notion that alternative splicing enables a rather neutral or even directional evolution, allowing this region to potentially acquire a new function [72].

A second mechanism does not directly create new splice site motifs. Here, existing donor or acceptor motifs can function as splice sites if the strength of an adjacent constitutive splice site is reduced by a mutation. In the wild-type context, these so-called cryptic sites are never used, indicating that they are only competitive to the weakened but not to the original splice site [73]. This agrees with the finding that both splice sites in a tandem are weaker than the splice sites of alternative and constitutive exons [40,73]. It is also conceivable that mutations in splicing regulatory sequences enable alternative tandem splicing, for example, by destroying silencer sequences that inhibit the use of intron–proximal splice sites [48].

Most orthologous mammalian exons have identical lengths (90% for human and mouse). By contrast, some NAGNAG acceptors show an exceptional flexibility in that one of the two acceptors is destroyed or created multiple times in independent lineages in evolution (Figure 3). It is conceivable that such tandem sites represent intermediate stages during an evolutionary process, leading to shorter or longer exon lengths. Because this mechanism requires as little as a single mutation to change the exon length, it might be more frequent than exon length evolution through genomic indels. Furthermore, this mechanism might counteract the slow sequence evolution at exon boundaries that is caused by constraints on splicing enhancer motifs and the protein coding sequence [74].

Concluding remarks

Splice events at short-distance tandem sites are widespread and contribute to transcriptome and proteome complexity. From the viewpoint of evolution, alternative splice events can be roughly divided into being deleterious (maybe causing a disease), neutral (being tolerated), or advantageous (having functional implications). This classification also holds for short-distance tandem splice events because experimental and computational studies found evidence for all three aspects. A remarkable feature of tandem sites is that events with functional implications can have constant splicing ratios. This suggests that the lack of detectable variation in alternative splicing does not rule out a function for the event. Moreover, stochastic splice site selection is likely important for both variable and constant splicing ratios and functional and nonfunctional events.

Further studies are needed to provide more insights into the tandem splicing mechanism, for example by investigating which splicing factors are involved, which motifs they bind and in which splicing steps they act. This will increase our understanding of how variable and constant splicing ratios at tandem sites arise and how these processes are regulated. Detailed characterization of further tandem splice events should reveal more examples of events with functional consequences, contributing to the question of what fraction of the tandem site events is deleterious, neutral and advantageous. Finally, mutations that create or destroy a tandem site or that change its splicing pattern might be candidates to investigate their potential involvement in dysfunction and disease.

Acknowledgements

We thank Rileen Sinha, Anke Busch, Klaus Huse and Karol Szafranski for helpful discussions. This work was supported by grants from the German Ministry of Education and Research (01GR0504, 0313652D) and the Deutsche Forschungsgemeinschaft (SFB604–02).

References

1 Sharp, P.A. (2005) The discovery of split genes and RNA splicing. *Trends Biochem. Sci.* 30, 279–281
2 Zavolan, M. *et al.* (2003) Impact of alternative initiation, splicing, and termination on the diversity of the mRNA transcripts encoded by the mouse transcriptome. *Genome Res.* 13, 1290–1300
3 Johnson, J.M. *et al.* (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302, 2141–2144
4 Graveley, B.R. (2001) Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* 17, 100–107
5 Stamm, S. *et al.* (2005) Function of alternative splicing. *Gene* 344, 1–20
6 Ule, J. *et al.* (2005) Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* 37, 844–852
7 Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* 72, 291–336
8 Lareau, L.F. *et al.* (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* 446, 926–929
9 Ni, J.Z. *et al.* (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* 21, 708–718
10 Karni, R. *et al.* (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.* 14, 185–193
11 Kim, E. *et al.* (2008) Insights into the connection between cancer and alternative splicing. *Trends Genet.* 24, 7–10
12 Wang, G.S. and Cooper, T.A. (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.* 8, 749–761
13 Campbell, M.A. *et al.* (2006) Comprehensive analysis of alternative splicing in rice and comparative analyses with *Arabidopsis*. *BMC Genomics* 7, 327
14 Kim, E. *et al.* (2007) Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res.* 35, 125–131
15 Holste, D. *et al.* (2006) HOLLYWOOD: a comparative relational database of alternative splicing. *Nucleic Acids Res.* 34, D56–D62
16 Chern, T.M. *et al.* (2006) A simple physical model predicts small exon length variations. *PLoS Genet.* 2, e45
17 Hiller, M. *et al.* (2004) Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity. *Nat. Genet.* 36, 1255–1257
18 Sugnet, C.W. *et al.* (2004) Transcriptome and genome conservation of alternative splicing events in humans and mice. *Pac. Symp. Biocomput.* 66–77
19 Dou, Y. *et al.* (2006) Genomic splice-site analysis reveals frequent alternative splicing close to the dominant splice site. *RNA* 12, 2047–2056
20 Levine, A. and Durbin, R. (2001) A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res.* 29, 4006–4013
21 Artamonova, I.I. and Gelfand, M.S. (2007) Comparative genomics and evolution of alternative splicing: the pessimists' science. *Chem. Rev.* 107, 3407–3430
22 Blencowe, B.J. (2006) Alternative splicing: new insights from global analyses. *Cell* 126, 37–47

23 Matlin, A.J. *et al.* (2005) Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol.* 6, 386–398
24 Xing, Y. and Lee, C. (2006) Alternative splicing and RNA selection pressure - evolutionary consequences for eukaryotic genomes. *Nat. Rev. Genet.* 7, 499–509
25 Tadokoro, K. *et al.* (2005) Frequent occurrence of protein isoforms with or without a single amino acid residue by subtle alternative splicing: the case of Gln in DRPLA affects subcellular localization of the products. *J. Hum. Genet.* 50, 382–394
26 Hiller, M. *et al.* (2007) TassDB: a database of alternative tandem splice sites. *Nucleic Acids Res.* 35, D188–D192
27 Ermakova, E.O. *et al.* (2007) Overlapping alternative donor splice sites in the human genome. *J. Bioinform. Comput. Biol.* 5, 991–1004
28 Hiller, M. *et al.* (2006) Phylogenetically widespread alternative splicing at unusual GYNGYN donors. *Genome Biol.* 7, R65
29 Hollins, C. *et al.* (2005) U2AF binding selects for the high conservation of the *C. elegans* 3′ splice site. *RNA* 11, 248–253
30 Hammes, A. *et al.* (2001) Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell* 106, 319–329
31 Unoki, M. *et al.* (2006) Novel splice variants of ING4 and their possible roles in the regulation of cell growth and motility. *J. Biol. Chem.* 281, 34677–34686
32 Hiller, M. *et al.* (2006) Single-nucleotide polymorphisms in NAGNAG acceptors are highly predictive for variations of alternative splicing. *Am. J. Hum. Genet.* 78, 291–302
33 Lai, C.H. *et al.* (2006) Single amino-acid InDel variants generated by alternative tandem splice-donor and -acceptor selection. *Biochem. Biophys. Res. Commun.* 342, 197–205
34 Tsai, K.W. and Lin, W.C. (2006) Quantitative analysis of wobble splicing indicates that it is not tissue specific. *Genomics* 88, 855–864
35 Tsai, K.W. *et al.* (2007) Wobble splicing reveals the role of the branch point sequence-to-NAGNAG region in 3′ tandem splice site selection. *Mol. Cell. Biol.* 27, 5835–5848
36 Akerman, M. and Mandel-Gutfreund, Y. (2006) Alternative splicing regulation at tandem 3′ splice sites. *Nucleic Acids Res.* 34, 23–31
37 Smith, C.W. *et al.* (1993) Scanning and competition between AGs are involved in 3′ splice site selection in mammalian introns. *Mol. Cell. Biol.* 13, 4939–4952
38 Akerman, M. and Mandel-Gutfreund, Y. (2007) Does distance matter? Variations in alternative 3′ splicing regulation. *Nucleic Acids Res.* 35, 5487–5498
39 Xia, H. *et al.* (2006) Identification of alternative 5′/3′ splice sites based on the mechanism of splice site competition. *Nucleic Acids Res.* 34, 6305–6313
40 Koren, E. *et al.* (2007) The emergence of alternative 3′ and 5′ splice site exons from constitutive exons. *PLoS Comput. Biol.* 3, e95
41 Roca, X. *et al.* (2005) Determinants of the inherent strength of human 5′ splice sites. *RNA* 11, 683–698
42 Zheng, C.L. *et al.* (2005) Characteristics and regulatory elements defining constitutive splicing and different modes of alternative splicing in human and mouse. *RNA* 11, 1777–1787
43 Hu, C.A. *et al.* (1999) Molecular enzymology of mammalian Delta1-pyrroline-5-carboxylate synthase. Alternative splice donor utilization generates isoforms with different sensitivity to ornithine inhibition. *J. Biol. Chem.* 274, 6754–6762
44 Koenig Merediz, S.A. *et al.* (2000) Cloning of an interferon regulatory factor 2 isoform with different regulatory ability. *Nucleic Acids Res.* 28, 4219–4224
45 Takeda, H. *et al.* (1994) A novel POU domain gene, zebrafish pou2: expression and roles of two alternatively spliced twin products in early development. *Genes Dev.* 8, 45–59
46 Yan, M. *et al.* (2000) Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* 290, 523–527
47 Xu, Q. *et al.* (2004) A second protein L-isoaspartyl methyltransferase gene in *Arabidopsis* produces two transcripts whose products are sequestered in the nucleus. *Plant Physiol.* 136, 2652–2664
48 Wang, Z. *et al.* (2006) General and specific functions of exonic splicing silencers in splicing control. *Mol. Cell* 23, 61–70
49 Pollard, A.J. *et al.* (2002) Alternative splicing of the adenylyl cyclase stimulatory G-protein G alpha(s) is regulated by SF2/ASF and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and involves the use of an unusual TG 3′-splice Site. *J. Biol. Chem.* 277, 15241–15251
50 Chua, K. and Reed, R. (1999) The RNA splicing factor hSlu7 is required for correct 3′ splice-site choice. *Nature* 402, 207–210
51 Lev-Maor, G. *et al.* (2003) The birth of an alternatively spliced exon: 3′ splice-site selection in Alu exons. *Science* 300, 1288–1291
52 Burgar, H.R. *et al.* (2002) Association of the signaling adaptor FRS2 with fibroblast growth factor receptor 1 (Fgfr1) is mediated by alternative splicing of the juxtamembrane domain. *J. Biol. Chem.* 277, 4018–4023
53 Li, L. and Howe, G.A. (2001) Alternative splicing of prosystemin pre-mRNA produces two isoforms that are active as signals in the wound response pathway. *Plant Mol. Biol.* 46, 409–419
54 Rivers, C. *et al.* (1999) Insertion of an amino acid in the DNA-binding domain of the glucocorticoid receptor as a result of alternative splicing. *J. Clin. Endocrinol. Metab.* 84, 4283–4286
55 Makielski, J.C. *et al.* (2003) A ubiquitous splice variant and a common polymorphism affect heterologous expression of recombinant human SCN5A heart sodium channels. *Circ. Res.* 93, 821–828
56 Treacy, M.N. *et al.* (1992) Twin of I-POU: a two amino acid difference in the I-POU homeodomain distinguishes an activator from an inhibitor of transcription. *Cell* 68, 491–505
57 Vogan, K.J. *et al.* (1996) An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity. *Mol. Cell. Biol.* 16, 6677–6686
58 Wagner, K.D. *et al.* (2003) The complex life of WT1. *J. Cell Sci.* 116, 1653–1658
59 Joyce-Brady, M. *et al.* (2001) gamma -glutamyltransferase and its isoform mediate an endoplasmic reticulum stress response. *J. Biol. Chem.* 276, 9468–9477
60 Hosoda, H. *et al.* (2000) Purification and characterization of rat des-Gln14-Ghrelin, a second endogenous ligand for the growth hormone secretagogue receptor. *J. Biol. Chem.* 275, 21995–22000
61 Barbaux, S. *et al.* (1997) Donor splice-site mutations in WT1 are responsible for Frasier syndrome. *Nat. Genet.* 17, 467–470
62 Maugeri, A. *et al.* (1999) The 2588G→C mutation in the ABCR gene is a mild frequent founder mutation in the Western European population and allows the classification of ABCR mutations in patients with Stargardt disease. *Am. J. Hum. Genet.* 64, 1024–1035
63 Rao, C.V. *et al.* (2002) Control, exploitation and tolerance of intracellular noise. *Nature* 420, 231–237
64 Graveley, B.R. (2005) Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell* 123, 65–73
65 Cohen, T. *et al.* (2005) The exon 8-containing prosaposin gene splice variant is dispensable for mouse development, lysosomal function, and secretion. *Mol. Cell. Biol.* 25, 2431–2440
66 Cajiao, I. *et al.* (2004) Bystander gene activation by a locus control region. *EMBO J.* 23, 3854–3863
67 Hiller, M. *et al.* (2006) Alternative splicing at NAGNAG acceptors: simply noise or noise and more? *PLoS Genet.* 2, e207
68 Hiller, M. *et al.* (2008) Assessing the fraction of short-distance tandem splice sites under purifying selection. *RNA* 14, 616–629
69 Hiller, M. *et al.* (2008) Selection against tandem splice sites affecting structured protein regions. *BMC Evol. Biol.* 8, 89
70 Hiller, M. *et al.* (2005) Non-EST based prediction of exon skipping and intron retention events using Pfam information. *Nucleic Acids Res.* 33, 5611–5621
71 Wen, F. *et al.* (2004) The impact of very short alternative splicing on protein structures and functions in the human genome. *Trends Genet.* 20, 232–236
72 Modrek, B. and Lee, C.J. (2003) Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat. Genet.* 34, 177–180
73 Roca, X. *et al.* (2003) Intrinsic differences between authentic and cryptic 5′ splice sites. *Nucleic Acids Res.* 31, 6321–6333
74 Parmley, J.L. *et al.* (2007) Splicing and the evolution of proteins in mammals. *PLoS Biol.* 5, e14
75 Sheth, N. *et al.* (2006) Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res.* 34, 3955–3967
76 Szafranski, K. *et al.* (2007) Violating the splicing rules: TG dinucleotides function as alternative 3′ splice sites in U2-dependent introns. *Genome Biol.* 8, R154

77 Crispino, J.D. and Sharp, P.A. (1995) A U6 snRNA:pre-mRNA interaction can be rate-limiting for U1-independent splicing. *Genes Dev.* 9, 2314–2323
78 Caceres, J.F. *et al.* (1994) Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. *Science* 265, 1706–1709
79 Brackenridge, S. *et al.* (2003) Efficient use of a 'dead-end' GA 5′ splice site in the human fibroblast growth factor receptor genes. *EMBO J.* 22, 1620–1631
80 Chen, S. *et al.* (2000) Evidence for a linear search in bimolecular 3′ splice site AG selection. *Proc. Natl. Acad. Sci. U. S. A.* 97, 593–598
81 Chua, K. and Reed, R. (2001) An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing. *Mol. Cell. Biol.* 21, 1509–1514
82 Lallena, M.J. *et al.* (2002) Splicing regulation at the second catalytic step by Sex-lethal involves 3′ splice site recognition by SPF45. *Cell* 109, 285–296
83 Bai, Y. *et al.* (1999) Control of 3′ splice site choice *in vivo* by ASF/SF2 and hnRNP A1. *Nucleic Acids Res.* 27, 1126–1134
84 Condorelli, G. *et al.* (1994) Two alternatively spliced forms of the human insulin-like growth factor I receptor have distinct biological activities and internalization kinetics. *J. Biol. Chem.* 269, 8510–8516
85 Lorkovic, Z.J. *et al.* (2005) Evolutionary conservation of minor U12-type spliceosome between plants and humans. *RNA* 11, 1095–1107
86 Caffrey, J.J. *et al.* (2000) Discovery of molecular and catalytic diversity among human diphosphoinositol-polyphosphate phosphohydrolases. An expanding Nudt family. *J. Biol. Chem.* 275, 12730–12736
87 Bourinet, E. *et al.* (1999) Splicing of alpha 1A subunit gene generates phenotypic variants of P- and Q-type calcium channels. *Nat. Neurosci.* 2, 407–415

Elsevier celebrates two anniversaries with a gift to university libraries in the developing world

In 1580, the Elzevir family began their printing and bookselling business in the Netherlands, publishing works by scholars such as John Locke, Galileo Galilei and Hugo Grotius. On 4 March 1880, Jacobus George Robbers founded the modern Elsevier company intending, just like the original Elzevir family, to reproduce fine editions of literary classics for the edification of others who shared his passion, other 'Elzevirians'. Robbers co-opted the Elzevir family printer's mark, stamping the new Elsevier products with a classic symbol of the symbiotic relationship between publisher and scholar. Elsevier has since become a leader in the dissemination of scientific, technical and medical (STM) information, building a reputation for excellence in publishing, new product innovation and commitment to its STM communities.

In celebration of the House of Elzevir's 425th anniversary and the 125th anniversary of the modern Elsevier company, Elsevier donated books to ten university libraries in the developing world. Entitled 'A Book in Your Name', each of the 6700 Elsevier employees worldwide was invited to select one of the chosen libraries to receive a book donated by Elsevier. The core gift collection contains the company's most important and widely used STM publications, including *Gray's Anatomy, Dorland's Illustrated Medical Dictionary, Essential Medical Physiology, Cecil Essentials of Medicine, Mosby's Medical, Nursing and Allied Health Dictionary, The Vaccine Book, Fundamentals of Neuroscience*, and *Myles Textbook for Midwives.*

The ten beneficiary libraries are located in Africa, South America and Asia. They include the Library of the Sciences of the University of Sierra Leone; the library of the Muhimbili University College of Health Sciences of the University of Dar es Salaam, Tanzania; the library of the College of Medicine of the University of Malawi; and the University of Zambia; Universite du Mali; Universidade Eduardo Mondlane, Mozambique; Makerere University, Uganda; Universidad San Francisco de Quito, Ecuador; Universidad Francisco Marroquin, Guatemala; and the National Centre for Scientific and Technological Information (NACESTI), Vietnam.

Through 'A Book in Your Name', these libraries received books with a total retail value of approximately one million US dollars.

For more information, visit www.elsevier.com