PERSPECTIVE

A simple principle to explain the evolution of pre-mRNA splicing

José-María Izquierdo¹ and Juan Valcárcel^{2,3,4,5}

¹Centro de Biología Molecular Severo Ochoa, Universidad Autónoma de Madrid, Cantoblanco 28049 Madrid, Spain;

²Institució Catalana de Recerca i Estudis Avançats, Barcelona 08003, Spain; ³Centre de Regulació Genòmica, Barcelona 08003, Spain; ⁴Universitat Pompeu-Fabra, Barcelona 08003, Spain

One of the most surprising discoveries of molecular biology was the realization that eukaryotic messenger RNAs (mRNA) are usually transcribed as precursors containing internal noncoding sequences (introns) that need to be excised to generate translatable mRNAs. The process of intron removal (or pre-mRNA splicing) requires precise definition of the intron boundaries. This is achieved in part through base-pairing interactions involving specific U-rich small nuclear RNAs (U snRNAs) and particular intronic sequences near the splice sites (for review, see Valadkhan 2005). While the U snRNA sequences involved in these interactions have been conserved during evolution, pre-mRNA splicing signals are conserved in introns of the yeast *Saccharomyces cerevisiae*, but they are significantly more variable in higher eukaryotes (Fig. 1A; for review, see Ast 2004). To compensate for this sequence divergence and consequent loss of base-pairing to U snRNAs, recognition of higher eukaryotic introns often relies on additional sequences located in exons and introns that act as splicing enhancers. One class of factors recognizing enhancer sequences are serine-arginine-rich (SR) proteins, which are a diverse family of splicing factors and regulators containing arginine-serine-rich (RS) domains. In contrast, SR-like proteins and RS domains are rare in budding yeast and have not been implicated in the splicing process. Results from Shen and Green (2006) in this issue of *Genes & Development* indicate that SR proteins allow for higher sequence variation at splicing signals because their RS domains interact with double-stranded RNA (dsRNA) and help to stabilize base-pairing interactions. Shen and Green (2006) show that yeast introns can afford sequence variation at splice sites provided that an heterologous RS domain is targeted to their vicinity. Conversely, the requirement for an SR protein to remove a higher eukaryotic intron can be waived by increasing the complementarity of a splicing signal with a particular U snRNA. These observations offer a surprisingly simple rationale for the concerted evolution of splicing signals and trans-acting factors and have important implications for un-

derstanding alternative splicing regulation. Sustaining this picture is the intriguing property of phosphorylated RS domains to recognize short (often imperfectly base-paired) stretches of dsRNA, which poses an interesting additional question for structural biologists.

Reconciling accuracy and flexibility in splice-site recognition

There are only ~250 annotated introns in the genome of *S. cerevisiae*, affecting 3% of the genes, of which only six contain more than one intron (Ares et al. 1999; López and Séraphin 1999). Most yeast introns are relatively short, and only a handful of instances of splicing regulation—in the form of intron retention—have been documented. In contrast, hundreds of thousands of introns have been identified in the human genome, and most genes contain multiple introns, which are on average 10–20 times longer than the corresponding exons (Ast 2004). A minimum of 40%–75% of human genes generate alternatively spliced transcripts using a variety of patterns, often providing distinct gene functions (Black 2003; Mendes Soares and Valcárcel 2006). Alternative splicing is now recognized as an important component of gene expression regulation in multicellular organisms, contributing to the control of many aspects of cellular function, development, and homeostasis.

How did regulation of splice site selection evolve? The emergence of regulatory mechanisms must be compatible with efficient and accurate recognition of constitutive splice sites, which is required to sustain gene expression. Even for yeast cells, splicing represents an important energy investment because some of the most transcribed genes (e.g., those encoding ribosomal proteins) contain introns (Ares et al. 1999; López and Séraphin 1999). An example of splicing control in *S. cerevisiae* can serve to illustrate one important regulatory principle. At least three introns are spliced exclusively during meiosis in this organism (Spingola and Ares 2000). A common feature of these introns is that their 5' splice sites depart from the consensus sequence. 5' splice sites are recognized by U1 snRNP. Binding of U1 involves base-pairing between the 5' end of its RNA component—U1 snRNA—and the first six nucleotides of the

⁵Corresponding author.

E-MAIL juan.valcarcel@crg.es; FAX 34-93-2240899.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1449106>.

Figure 1. Recognition of splicing signals. (A) Splicing signals in yeast and mammals. The scheme represents a pre-mRNA containing two exons and an intervening intron and the location of sequences relevant for the splicing process. Only the first (GT) and last (AG) two nucleotides of the intron are highly conserved across evolution. Conservation of other sequences is variable in different organisms. This is illustrated for 5' splice site and branchpoint sequences by representing the most frequent nucleotide at each position, the size of each nucleotide being proportional to its relative use. The underlined A indicates the nucleotide undergoing 2'–5' phosphodiester bond formation after the first catalytic step. (B) 5' splice sites are recognized by U1 snRNP, the 5' end of U1 snRNA (in yellow) establishing base-pairing interactions with six nucleotides at the 5' end of the intron (indicated by thin lines). (C) Cooperative binding of BBP/SF1 and the two subunits of U2AF facilitates recognition of the three sequence elements relevant for definition of the 3' end of introns in higher eukaryotes, described in A.

intron (Fig. 1B; Ast 2004). Departure from the 5' splice site consensus implies a reduction in the number of possible base pairs with U1 snRNA and therefore a lower free energy of the interaction. Meiosis-specific splicing of these introns requires expression of the protein Mer1p, which is expressed only during this phase of the life cycle (Engbrecht et al. 1991). Binding of Mer1p to an intronic enhancer sequence allows recruitment of U1 snRNP to suboptimal 5' splice sites, as expression of a U1 snRNA variant that extends base-pairing with a regulated splice site renders splicing of the corresponding intron Mer1p-independent (Nandabalan et al. 1993; Spingola and Ares 2000). The principle illustrated by this example is that regulation can be achieved by (1) weakening a standard molecular recognition event and (2) making recognition dependent on a dedicated regulatory factor. It is unclear how Mer1p facilitates U1 snRNP recruitment, but regulators that promote 5' splice site recognition in *Drosophila* and mammals do so by binding to sequence motifs adjacent to the splice site sequence and establishing protein–protein interactions with U1 snRNP-specific protein components (Labourier et al. 2001; Förch et al. 2002). Because loss of a single base pair contact suffices to make a 5' splice site regulatable in yeast, it seems likely that the significant deviation from the consensus observed in higher eukary-

otic sites contributes to the versatility of splice site selection mechanisms operating in these organisms.

Recognition of the 3' end of introns offers another example of the same principle and further illustrates the differences in splice site recognition between yeast and higher eukaryote organisms. Intron removal involves two successive transesterification reactions. The first reaction entails cleavage of the phosphodiester bond between the first nucleotide of the intron and the last nucleotide of the preceding exon. Concomitant with this cleavage is formation of a 2'–5' phosphodiester bond between the 5' end of the intron and a specific adenosine residue—known as the branchpoint—located toward the end of the intron. The second transesterification reaction renders the flanking exons covalently bound and the intron released in a lariat configuration. In *S. cerevisiae*, the branchpoint is specified by binding of the branchpoint-binding protein (BBP, also known as SF1) to the branch adenosine and flanking sequences, followed by recruitment of U2 snRNP (Berglund et al. 1997; Liu et al. 2001). The sequences flanking the branchpoint are conserved in yeast introns, whereas they are much more degenerate in higher eukaryotes (Fig. 1A). Identification of the branchpoint by BBP/SF1 in higher eukaryotes relies heavily on additional downstream sequences: the polypyrimidine (Py)-tract and the 3' splice site AG. None of these additional sequences are essential for U2 snRNP binding or for undergoing the first catalytic step in yeast pre-mRNAs, whereas they are important for the majority of pre-mRNAs in higher eukaryotes. The Py-tract and the 3' splice site AG are recognized by the 65- and 35-kDa subunits, respectively, of the U2 Auxiliary Factor (U2AF) (Fig. 1C; Moore 2000). Interactions between BBP/SF1 and U2AF facilitate cooperative binding to assemblies of branchpoint/Py-tract/AG sequences despite the limited sequence conservation and/or limited affinity of each individual component for the corresponding sequence elements (Berglund et al. 1998; Selenko et al. 2003). As with regulated 5' splice sites, this partition of the energy of recognition opens the potential for modulation of 3' splice site usage (e.g., by enhancing or inhibiting the binding of U2AF).

The examples above document how regulation of splice site choice can be achieved at specific sites or genes depending on particular sets of auxiliary sequences, factors, and interactions. Can more general principles and basic molecular operations be envisaged as the forces that drive the evolution of the pre-mRNA splicing process and its regulation? The results by Shen and Green (2006) unveil a general mechanism that exploits the principle of sequence diversification using RS domains, which are characteristic of a family of splicing factors and regulators prominent in higher eukaryotes but not in budding yeast.

The debated functions of SR proteins and RS domains

In addition to helping define the 3' splice site region, U2AF facilitates U2 snRNP assembly on the branchpoint. This involves base-pairing between an internal se-

quence of U2 snRNA and nucleotides flanking the branch adenosine. As with 5' splice sites and U1 snRNA, the yeast branchpoint sequence is conserved and has maximum base-pairing potential with U2 snRNA, while variation of the U2 target sequence in higher eukaryotes decreases the potential to form base pairs (Fig. 1A). An RS domain located at the N-terminal region of U2AF⁶⁵ is necessary for U2 snRNP recruitment and U2 snRNA/branchpoint base-pairing (Zamore et al. 1992; Valcárcel et al. 1996). The RS domain contacts the branchpoint region upon binding of U2AF⁶⁵ RNA recognition motifs (RRM) 1 and 2 to the Py-tract (Valcárcel et al. 1996; Banerjee et al. 2003; Kent et al. 2003; Shen and Green 2004). Mutational studies indicated that net positive charges are the only requirement for U2AF⁶⁵ RS domain function in *in vitro* assays, and it was therefore proposed that the motif functions as targeted annealing activity to facilitate otherwise unstable base-pairing interactions (Lee et al. 1993; Valcárcel et al. 1996). Positive charges may function by the same principle that polyamines use to enhance formation of double-stranded nucleic acids by neutralizing the negative charges of the approaching strands. Indeed, U2AF⁶⁵ alone can promote interactions between U2 snRNA and the pre-mRNA branchpoint in a reconstitution assay using purified components (Valcárcel et al. 1996).

A different function was proposed for RS domains of SR proteins, a prominent family of splicing factors and regulators that also contain RRM motifs (Hertel and Graveley 2005; Singh and Valcárcel 2005). Yeast two-hybrid and coimmunoprecipitation assays documented that RS domains of SR proteins engage in protein-protein interactions (Wu and Maniatis 1993; Kohtz et al. 1994). Serine phosphorylation was found to be required for the interaction, suggesting that matching of alternate positive and negative charges between two domains could drive formation of a "polar zipper" mediating the contact (Perutz 1994). These observations provided an appealing model to explain splice site communication and the function of exonic splicing enhancers, which recruit complexes of SR proteins and promote the use of upstream 3' splice sites (Fig. 2). RS domain-mediated interactions between enhancer-bound SR proteins and U2AF³⁵ can stabilize U2AF binding to otherwise weakly recognized 3' splice site regions (Zuo and Maniatis 1996; Graveley et al. 2001). Although the recruitment model can be recapitulated with purified components, these interactions have not yet been documented to occur in splicing complexes (Hertel and Graveley 2005). In addition, increasing U2AF binding is only one of the functions of U2AF³⁵ in enhancer-dependent splicing (Graveley 2000; Graveley et al. 2001; Guth et al. 2001), and U2AF³⁵ is dispensable for the function of some exonic enhancers (Rudner et al. 1996).

An experimental approach that has been useful to study the function of exonic splicing enhancers takes advantage of the tight binding of the bacteriophage MS2 coat protein to its cognate stem-loop-binding site to tether a protein domain as a MS2 fusion protein to a particular location within an RNA molecule. Using this

Figure 2. Sequential interactions between RS domains and RNA during formation of splicing complexes. The figure highlights the known interactions between RS domains and splicing signals in different complexes along the spliceosome assembly pathway. (A) In early (E) complexes, the RS domain of U2AF⁶⁵ contacts the branchpoint region, which can assist initial interactions between U2 snRNA and this region of the pre-mRNA. (B) At the time of stable binding of U2 snRNP (prespliceosome or complex A), the RS domain of an SR protein bound to an exonic enhancer (ESE) interacts with the branchpoint region base-paired to U2 snRNA. (C) After recruitment of the U4,5,6 tri-snRNP to form mature spliceosomes (complexes B/C), the RS domain of another SR protein (possibly interacting with intronic sequences through its RRM domain) contacts the 5' splice site region base-paired to U6 snRNA.

experimental setup, it was possible to show that simple repeats of RS dipeptides can activate splicing from an exonic tethering site (Graveley and Maniatis 1998; Phillips et al. 2003). Surprisingly, Shen et al. (2004) found, using a UV cross-linking assay, that the RS domain contacts the branchpoint region in prespliceosomal complexes, implying a physical bridge between the enhancer and the branch site (Fig. 2). Arguing for the functional relevance of this contact, mutational analyses showed a correlation between the amino acid requirements for branch site cross-linking and for splicing activation. Two features distinguished the U2AF⁶⁵ RS domain/branchpoint interaction from that of the enhancer-bound RS motif with the branchpoint. First, whereas only net positive charges were required for the U2AF⁶⁵ RS domain contact, serines—presumably phosphorylated—were additionally required for the MS2-RS/branchpoint contact. Second, whereas the interaction of the U2AF⁶⁵ RS domain with the branchpoint appears to be an intrinsic property of the protein when bound to the downstream Py-tract, the MS2-RS contact (or that of the RS domain of an SR protein bound to the enhancer) requires functional U2 snRNP and a functional branch site sequence.

Although these observations were intriguing, they did not reveal the functional role of RS domains in splicing, nor did they entirely rule out the possibility that the observed cross-linking was just a byproduct of a nearby protein–protein interaction. Shen and Green (2006) now document that targeting of the enhancer-bound RS domain to the branch site is driven by the potential of the sequence to form a dsRNA region with U2 snRNA (or, in general, with complementary RNAs within a distance window from the tethering site). They also show that the proximity of the RS domain, in turn, facilitates base-pairing between the complementary strands. An appealing feature of these results is that they simultaneously provide a mechanism and a purpose for the function of an RS domain in an essential step for the assembly of splicing complexes.

A simple mechanism may allow (and drive) evolution

How general is this mechanism for the function of SR proteins and for other events in the splicing process promoted by these factors? Shen and Green (2004) had previously observed that, in addition to the MS2–RS domain fusion bound to the enhancer, splicing in cytoplasmic S100 extracts—which are devoid of SR proteins—required another SR protein. The additional SR protein was not involved in U2 snRNP binding, but rather in subsequent steps in spliceosome assembly involving the recruitment of the U4/U5/U6 tri-snRNP. One of the interactions established by these snRNPs with the pre-mRNA substrate is a set of base-pairing contacts between U6 snRNA and nucleotides in the 5' splice site region that partially overlap with those of U1 snRNA (Fig. 2; Wassarman and Steitz 1992; Kandels-Lewis and Séraphin 1993). The U6 snRNA/5' splice site interaction replaces the earlier base-pairing with U1 snRNA prior to catalytic activation of the spliceosome (Konforti et al. 1993). Analogous to the results obtained at the branch-point region, the RS domain of the SR protein was found to contact the 5'-splice site region and promote base-pairing between the splice site and U6 snRNA (Fig. 2; Shen and Green 2004). This model is now further substantiated by the observation that the additional SR protein is no longer required for splicing if the 5' splice site is mutated to increase the number of base pairs that can form with U6 snRNA (Shen and Green 2006).

This result has two important implications. First, it shows that RS domains, functioning as identifiers/promoters of base-pairing interactions, act in sequential order at various steps in spliceosome assembly: Interaction of the branchpoint with the RS domain of U2AF⁶⁵ in early complexes can help initial base-pairing between the branchpoint and U2 snRNA; this is followed by interaction of this base-paired region with the RS domain of exonic enhancer-bound SR proteins in prespliceosomal complexes containing stably bound U2 snRNP; finally, interaction between the 5' splice site region and U6 snRNA occurs in fully assembled spliceosomal complexes (Shen and Green 2004). Given the multiple effects of SR proteins in splicing and the multiple base-pairing

interactions that build the spliceosome or mediate its conformational changes leading to (probably RNA-based) catalysis (Staley and Guthrie 1998), it seems quite likely that other RNA–RNA transactions in the spliceosome are modulated by RS domains.

The property of RS domains to contact RNA has also suggestive evolutionary implications. If the function of an SR protein can be made superfluous by increasing the base-pairing between a U snRNA and pre-mRNA, is this the way in which organisms without SR proteins manage to remove introns? Shen and Green (2006) turned the question around and tested whether tethering an RS domain (again, as an MS2 fusion) close to a splicing signal could help yeast cells to splice pre-mRNAs containing mutations that reduce the base-pairing potential with their cognate U snRNAs. Although only some mutants were rescued and splicing efficiency was only partially restored, the striking result was that RS domains could provide function (in a splicing signal-specific manner) in an organism that has diverged during hundreds of millions of years from those that exploit extensively the activities of these protein motifs. This is consistent with RS domains providing a relatively simple molecular operation that does not require an array of cofactors and coevolved interactions. The immediate corollary is that the emergence of SR proteins may have allowed evolutionary divergence of splicing signals as well as diversification of these sequences in the transcripts of a given organism. These, in turn, provided versatility in splice site recognition, the possibility to regulate the process, and the generation of alternatively spliced transcripts. It is therefore not unexpected that SR proteins and other RS domain-containing proteins play important roles as regulators of alternative pre-mRNA splicing (Cáceres and Kornblihtt 2002; Singh and Valcárcel 2005).

What remains to be understood

The results from Shen and Green (2006) leave many interesting open questions for the future. Why are some yeast mutants rescued by SR proteins and others are not? The observation that the RS domain is not cross-linked to mutants that are not rescued argues that the interaction is functionally relevant, but there is no obvious rationale to distinguish between the two groups depending on, for example, remaining base-pairing potential. A step in splice site recognition previous to the establishment of base-pairing may be deficient in these mutants, thus leaving no room for the base-pairing-promoting activity of SR proteins to have an effect. There is precedent for this, as it is known that the U1 snRNP protein U1C recognizes the 5' splice site sequence before U1 snRNA establishes base-pairing interactions (Du and Rosbash 2002; Lund and Kjems 2002). An alternative would be that the RS domain distinguishes between different partially complementary dsRNA sequences. In this context, the requirement of serine phosphorylation for the RS domain to contact dsRNA is also intriguing. While at least some serines are phosphorylated when the RS domain interacts with the RNA (Shen and Green 2006), the ex-

tent to which this affects only a few or the majority of the serine residues present in the motif is currently unclear. This may be relevant to models invoking simple electrostatic interactions shielding the negative charges of the approaching complementary RNA strands. As argued by Shen and Green, the higher charge density of duplexes versus single-stranded regions may act as a signal for targeting the RS domain. RS domains could thus serve both as identifiers of base-pairing interactions and as RNA chaperones that stabilize them in the proximity of the tethering site.

Given the general nature of the structures recognized and of the activities involved, it seems likely that RS domains can contribute to other RNA–RNA interactions, including formation or stabilization of particular secondary structures or assembly of other RNAs or RNP complexes, from 3' end formation (Dettwiler et al. 2004) to target recognition by microRNAs. SR proteins have been implicated in other steps of mRNA metabolism, including RNA decay and translation (Lemaire et al. 2002; Sanford et al. 2004). It is conceivable that some of these effects are based on stabilization of RNA–RNA interactions, alternative RNA conformations, etc. It is noteworthy, in this regard, that *S. cerevisiae* SR-like proteins have been implicated in mRNA packaging and translational control (Hurt et al. 2004; Windgassen et al. 2004), suggesting that these could have been the ancestral functions of these factors before their expansion and recruitment for the splicing regulation process. Conversely, it will be interesting to consider the possibility that other families of regulatory factors harboring other protein domains exploit similar general strategies to achieve their functions in RNA biology.

While the collective results from Shen and Green document an important function of RS domains on RNA recognition, they do not rule out that other activities of these domains—including the establishment of protein–protein interactions—are also relevant for the splicing activation properties of SR proteins. An interesting possibility would be that RS domains engage in both protein–protein and RNA–protein interactions at different stages of the splicing process or as a way to recycle the complexes and allow the dynamic changes in RNA–RNA contacts that drive spliceosome assembly and catalysis (Staley and Guthrie 1998). The extent of RS domain phosphorylation, which is known to modulate SR protein function (Cáceres and Kornblihtt 2002), could serve as a switch between these modes of action. Phosphorylation could also serve to facilitate transitions between alternative RNA configurations. It may be interesting, for example, to study the effect of RS domains on the transition between alternative base-pairing schemes involving U6 snRNA and the 5' splice site occurring between the first and second catalytic steps of the splicing reaction (Konarska et al. 2006).

The evolutionary picture of sequence divergence associated with the emergence and expansion of RS domain-containing factors has been drawn from two, arguably extreme, examples. Careful analyses of sequence divergence and of SR protein-like factors in other yeasts (Bon

et al. 2003; Webb et al. 2005) and other classes of unicellular organisms will be needed to evaluate this hypothesis, identify intermediate steps in the evolution of splicing signals and in the acquisition of RS domains, and gauge their potential value as evolutionary clocks.

Acknowledgments

We thank Michael Green, Sara Evans, Gil Ast, and Josep Vilardell for discussions and useful comments on the manuscript. Work in the authors' laboratories is supported by grants from Fondo de Investigaciones Sanitarias, Ministerio de Educación y Ciencia, EU FP6 (EURASNET), and Generalitat de Catalunya.

References

- Ares Jr., M., Grate, L., and Pauling, M.H. 1999. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* **5**: 1138–1139.
- Ast, G. 2004. How did alternative splicing evolve? *Nat. Rev. Genet.* **5**: 773–782.
- Banerjee, H., Rahn, A., Davis, W., and Singh, R. 2003. Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF⁶⁵) recognize polypyrimidine tracts using multiple modes of binding. *RNA* **9**: 88–99.
- Berglund, J.A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89**: 781–787.
- Berglund, J.A., Abovich, N., and Rosbash, M. 1998. A cooperative interaction between U2AF⁶⁵ and mBBP/SF1 facilitates branchpoint region recognition. *Genes & Dev.* **12**: 858–867.
- Black, D.L. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* **72**: 291–336.
- Bon, E., Casaregola, S., Blandin, G., Llorente, B., Neugeglise, C., Munsterkotter, M., Guldener, U., Mewes, H.W., Van Helden, J., Dujon, B., et al. 2003. Molecular evolution of eukaryotic genomes: Hemiascomycetous yeast spliceosomal introns. *Nucleic Acids Res.* **31**: 1121–1135.
- Cáceres, J.F. and Kornblihtt, A.R. 2002. Alternative splicing: Multiple control mechanisms and involvement in human disease. *Trends Genet.* **18**: 186–193.
- Dettwiler, S., Aringhieri, C., Cardinale, S., Keller, W., and Barabino, S.M. 2004. Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein–protein interactions, and subcellular localization. *J. Biol. Chem.* **279**: 35788–35797.
- Du, H. and Rosbash, M. 2002. The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature* **419**: 86–90.
- Engbrecht, J.A., Voelkel-Meiman, K., and Roeder, G.S. 1991. Meiosis-specific RNA splicing in yeast. *Cell* **20**: 1257–1268.
- Förch, P., Puig, O., Martinez, C., Seraphin, B., and Valcárcel, J. 2002. The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites. *EMBO J.* **21**: 6882–6892.
- Graveley, B.R. 2000. Sorting out the complexity of SR protein functions. *RNA* **6**: 1197–1211.
- Graveley, B.R. and Maniatis, T. 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol. Cell* **1**: 765–771.
- Graveley, B.R., Hertel, K.J., and Maniatis, T. 2001. The role of U2AF³⁵ and U2AF⁶⁵ in enhancer-dependent splicing. *RNA* **7**: 806–818.
- Guth, S., Tange, T.O., Kellenberger, E., and Valcárcel, J. 2001.

- Dual function for U2AF³⁵ in AG-dependent pre-mRNA splicing. *Mol. Cell. Biol.* **21**: 7673–7681.
- Hertel, K.J. and Graveley, B.R. 2005. RS domains contact the pre-mRNA throughout spliceosome assembly. *Trends Biochem. Sci.* **30**: 115–118.
- Hurt, E., Luo, M.J., Rother, S., Reed, R., and Strasser, K. 2004. Cotranscriptional recruitment of the serine–arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc. Natl. Acad. Sci.* **101**: 1858–1862.
- Kandels-Lewis, S. and Seraphin, B. 1993. Involvement of U6 snRNA in 5' splice site selection. *Science* **262**: 2035–2039.
- Kent, O.A., Reayi, A., Foong, L., Chilibeck, K.A., and MacMillan, A.M. 2003. Structuring of the 3' splice site by U2AF⁶⁵. *J. Biol. Chem.* **278**: 50572–50577.
- Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R., García-Blanco, M.A., and Manley, J.L. 1994. Protein–protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* **368**: 119–124.
- Konarska, M.M., Vilardell, J., and Query, C.C. 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol. Cell* **21**: 543–553.
- Konforti, B.B., Koziolkiewicz, M.J., and Konarska, M.M. 1993. Disruption of base pairing between the 5' splice site and the 5' end of U1 snRNA is required for spliceosome assembly. *Cell* **75**: 863–873.
- Labourier, E., Adams, M.D., and Rio, D.C. 2001. Modulation of P-element pre-mRNA splicing by a direct interaction between PSI and U1 snRNP 70K protein. *Mol. Cell* **8**: 363–373.
- Lee, C.G., Zamore, P.D., Green, M.R., and Hurwitz, J. 1993. RNA annealing activity is intrinsically associated with U2AF. *J. Biol. Chem.* **268**: 13472–13478.
- Lemaire, R., Prasad, J., Kashima, T., Gustafson, J., Manley, J.L., and Lafyatis, R. 2002. Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: A novel function for SR proteins. *Genes & Dev.* **16**: 594–607.
- Liu, Z., Luyten, I., Bottomley, M.J., Messias, A.C., Houngrinou-Molango, S., Sprangers, R., Zanier, K., Kramer, A., and Sattler, M. 2001. Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science* **294**: 1098–1102.
- López, P.J. and Séraphin, B. 1999. Genomic-scale quantitative analysis of yeast pre-mRNA splicing: Implications for splice-site recognition. *RNA* **5**: 1135–1137.
- Lund, M. and Kjems, J. 2002. Defining a 5' splice site by functional selection in the presence and absence of U1 snRNA 5' end. *RNA* **8**: 166–179.
- Mendes Soares, L.M. and Valcárcel, J. 2006. The expanding transcriptome: The genome as the 'Book of Sand.' *EMBO J.* **25**: 923–931.
- Moore, M.J. 2000. Intron recognition comes of AGE. *Nat. Struct. Biol.* **7**: 14–16.
- Nandabalan, K., Price, L., and Roeder, G.S. 1993. Mutations in U1 snRNA bypass the requirement for a cell type-specific RNA splicing factor. *Cell* **73**: 407–415.
- Perutz, M. 1994. Polar zippers: Their role in human disease. *Protein Sci.* **3**: 1629–1637.
- Philipps, D., Celotto, A.M., Wang, Q.Q., Tarng, R.S., and Graveley, B.R. 2003. Arginine/serine repeats are sufficient to constitute a splicing activation domain. *Nucleic Acids Res.* **31**: 6502–6508.
- Rudner, D.Z., Kanaar, R., Breger, K.S., and Rio, D.C. 1996. Mutations in the small subunit of the *Drosophila* U2AF splicing factor cause lethality and developmental defects. *Proc. Natl. Acad. Sci.* **93**: 10333–10337.
- Sanford, J.R., Gray, N.K., Beckmann, K., and Cáceres, J.F. 2004. A novel role for shuttling SR proteins in mRNA translation. *Genes & Dev.* **18**: 755–768.
- Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Kramer, A., and Sattler, M. 2003. Structural basis for the molecular recognition between human splicing factors U2AF⁶⁵ and SF1/mBBP. *Mol. Cell* **11**: 965–976.
- Shen, H. and Green, M.R. 2004. A pathway of sequential arginine–serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Mol. Cell* **16**: 363–373.
- . 2006. RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes & Dev.* (this issue).
- Shen, H., Kan, J.L., and Green, M.R. 2004. Arginine–serine-rich domains bound at splicing enhancers contact the branch-point to promote prespliceosome assembly. *Mol. Cell* **13**: 367–376.
- Singh, R. and Valcárcel, J. 2005. Building specificity with non-specific RNA-binding proteins. *Nat. Struct. Mol. Biol.* **12**: 645–653.
- Spingola, M. and Ares Jr., M. 2000. A yeast intronic splicing enhancer and Nam8p are required for Mer1p-activated splicing. *Mol. Cell* **6**: 329–338.
- Staley, J.P. and Guthrie, C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* **92**: 315–326.
- Valadkhan, S. 2005. snRNAs as the catalysts of pre-mRNA splicing. *Curr. Opin. Chem. Biol.* **9**: 603–608.
- Valcárcel, J., Gaur, R.K., Singh, R., and Green, M.R. 1996. Interaction of U2AF⁶⁵ RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA. *Science* **273**: 1706–1709.
- Wassarman, D.A. and Steitz, J.A. 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science* **257**: 1918–1925.
- Webb, C.J., Romfo, C.M., van Heeckeren, W.J., and Wise, J.A. 2005. Exonic splicing enhancers in fission yeast: Functional conservation demonstrates an early evolutionary origin. *Genes & Dev.* **19**: 242–254.
- Windgassen, M., Sturm, D., Cajigas, I.J., Gonzalez, C.I., Seedorf, M., Bastians, H., and Krebber, H. 2004. Yeast shuttling SR proteins Npl3p, Gbp2p, and Hrb1p are part of the translating mRNPs, and Npl3p can function as a translational repressor. *Mol. Cell. Biol.* **24**: 10479–10491.
- Wu, J.Y. and Maniatis, T. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**: 1061–1070.
- Zamore, P.D., Patton, J.G., and Green, M.R. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355**: 609–614.
- Zuo, P. and Maniatis, T. 1996. The splicing factor U2AF³⁵ mediates critical protein–protein interactions in constitutive and enhancer-dependent splicing. *Genes & Dev.* **10**: 1356–1368.

A simple principle to explain the evolution of pre-mRNA splicing

José-María Izquierdo and Juan Valcárcel

Genes Dev. 2006, **20**:

Access the most recent version at doi:[10.1101/gad.1449106](https://doi.org/10.1101/gad.1449106)

Related Content **RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans**
Haihong Shen and Michael R. Green
[Genes Dev. UNKNOWN , 2006 20: 1755-1765](#)

References This article cites 50 articles, 20 of which can be accessed free at:
<http://genesdev.cshlp.org/content/20/13/1679.full.html#ref-list-1>

Articles cited in:
<http://genesdev.cshlp.org/content/20/13/1679.full.html#related-urls>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).