Elsevier Editorial System(tm) for Gene
Manuscript Draft

Manuscript Number: GENE-D-12-00204R1

Title: Functional coupling of transcription and splicing

Article Type: Review Article

Keywords: processing of RNA; co-transcriptional splicing; RNA polymerase II

Corresponding Author: Dr. Carlos Suñé, Ph.D.

Corresponding Author's Institution: IPBLN-CSIC

First Author: Marta Montes

Order of Authors: Marta Montes; Soraya Becerra; Miguel Sánchez-Álvarez; Carlos Suñé, Ph.D.

Manuscript Region of Origin: SPAIN

Abstract: The tightly regulated process of precursor messenger RNA (pre-mRNA) alternative splicing is a key mechanism to increase the number and complexity of proteins encoded by the genome. Evidence gathered in recent years has established that transcription and splicing are physically and functionally coupled and that this coupling may be an essential aspect of the regulation of splicing and alternative splicing. Recent advances in our understanding of transcription and of splicing regulation have uncovered the multiple interactions between components from both types of machinery. These interactions help to explain the functional coupling of RNAPII transcription and pre-mRNA alternative splicing for efficient and regulated gene expression at the molecular level. Recent technological advances, in addition to novel cell and molecular biology approaches, have led to the development of new tools for addressing mechanistic questions to achieve an integrated and global understanding of the functional coupling of RNAPII transcription and pre-mRNA alternative splicing. Here, we review major milestones and insights into RNA polymerase II transcription and pre-mRNA alternative splicing as well as new concepts and challenges that have arisen from multiple genome-wide approaches and analyses at the single-cell resolution.

Opposed Reviewers:

Functional coupling of transcription and splicing

Marta Montes[a], Soraya Becerra[a,†], Miguel Sánchez-Álvarez[b,†], and Carlos Suñé[a,*]

[a]*Department of Molecular Biology, Instituto de Parasitología y Biomedicina "López-Neyra" (IPBLN-CSIC), 18100 Armilla, Spain*

[b]*Dynamical Cell Systems Team, Section of Cellular and Molecular Biology, The Institute of Cancer Research, London SW3 6JB, UK*

*Corresponding author. Department of Molecular Biology, Instituto de Parasitología y Biomedicina "López-Neyra" (IPBLN-CSIC), PT Ciencias de la Salud, Av. del Conocimiento s/n., 18100 Armilla, Spain.

Tel.: +34.958181645

Fax: +34.958181632

Email: csune@ipb.csic.es

[†]These authors contributed equally to this work

Abstract

The tightly regulated process of precursor messenger RNA (pre-mRNA) alternative splicing is a key mechanism to increase the number and complexity of proteins encoded by the genome. Evidence gathered in recent years has established that transcription and splicing are physically and functionally coupled and that this coupling may be an essential aspect of the regulation of splicing and alternative splicing. Recent advances in our understanding of transcription and of splicing regulation have uncovered the multiple interactions between components from both types of machinery. These interactions help to explain the functional coupling of RNAPII transcription and pre-mRNA alternative splicing for efficient and regulated gene expression at the molecular level. Recent technological advances, in addition to novel cell and molecular biology approaches, have led to the development of new tools for addressing mechanistic questions to achieve an integrated and global understanding of the functional coupling of RNAPII transcription and pre-mRNA alternative splicing. Here, we review major milestones and insights into RNA polymerase II transcription and pre-mRNA alternative splicing as well as new concepts and challenges that have arisen from multiple genome-wide approaches and analyses at the single-cell resolution.

Introduction

The expression of protein-coding genes in eukaryotes is a complex process that involves numerous tightly regulated steps. The genomic coding DNA sequence must be transcribed into a precursor messenger RNA (pre-mRNA) by the nuclear RNA polymerase II (RNAPII). The pre-mRNA must then undergo a number of processing steps to yield a mature, functional messenger RNA (mRNA), which is then ready to be exported from the nucleus and used by the translational machinery of the cell. Thus, RNA transcripts undergo extensive maturation in the nuclei of eukaryotic cells. The complexity of the processing that occurs during the formation of a mature mRNA provides a basis for the multilayered regulation of gene expression. These regulatory networks correlate with the organizational complexity of a given organism and with the specialization of tissues such as the components of the central nervous system and the cardiac muscle (Maniatis and Tasic, 2002; Mendes Soares and Valcarcel, 2006; Taliaferro et al., 2011).

In higher eukaryotes, most protein-coding genes contain long non-coding stretches or intervening sequences (introns) that are transcribed in the pre-mRNA and must therefore be removed in a process called splicing, which leaves the resulting protein coding sequences (exons) appropriately aligned and ligated in the mRNA. The last decade has revealed the massive complexity and fundamental importance of this phenomenon in the regulation of gene expression. Splicing is a complex two-step trans-esterification reaction that is driven by the spliceosome, which is a large ribonucleoprotein complex of five uridine-rich small nuclear RNAs (snRNAs) and approximately 200 additional proteins. Alternative splicing, the phenomenon by which a single pre-mRNA can be potentially spliced in multiple different patterns to yield two or more different mRNA products, contributes significantly to the diversity of cell and

tissue-specific protein expression profiles (Maniatis and Tasic, 2002). The results of deep sequencing-based expression analyses suggest that more than 90% of multi-exon human genes undergo alternative splicing (Pan et al., 2008; Wang et al., 2008). Changes in the *cis*- or *trans*- regulation of this process can cause multiple pathologies as a result of general or specific aberrant pre-mRNA processing underscoring the fundamental importance of this regulatory process. Several reviews covering dysfunctions related to mRNA processing have been published elsewhere (Caceres and Kornblihtt, 2002; Garcia-Blanco et al., 2004; Licatalosi and Darnell, 2006; Wang and Cooper, 2007). Although our understanding of how alternative splicing is controlled in the context of global gene expression regulation and its relevance in specific metabolic processes has greatly advanced during the last decade, the *functional* integration of pre-mRNA processing with other nuclear machineries is not yet well understood.

It is widely accepted that transcription and splicing are physically and functionally coupled and that this coupling might be an important aspect of the regulation of splicing and alternative splicing. This coupling implies that splicing occurs co-transcriptionally (i.e., during transcription). However, this process can also take place post-transcriptionally (i.e., after transcription). Ten years have passed since the co-transcriptional splicing of pre-mRNAs, with a particular emphasis on the connections between transcriptional elongation and splicing, was last thoroughly discussed in this forum (Goldstrohm et al., 2001b). Since then, excellent reviews covering specific aspects of the coupling of transcription and alternative splicing have been published (for recent reviews see (Moore and Proudfoot, 2009; Munoz et al., 2010; Oesterreich et al., 2011)). In this review we aim to revisit current concepts and experimental observations, discuss the recent advances in the field, and highlight the major unanswered questions

regarding the functional coupling of RNAPII transcription and pre-mRNA alternative splicing.

The regulation of alternative splicing: general arguments for functional coupling

The predominant form of the spliceosome (the U2-spliceosome or "major spliceosome") is composed of a core of five snRNPs, termed U1, U2, U4, U5, and U6, and an additional estimated number of ~ 200 protein factors (Kramer, 1996; Wahl et al., 2009). Numerous *in vitro* and *in vivo* studies support the model that spliceosome assembly occurs in an ordered step-wise manner (Will and Luhrmann, 2001; Gornemann et al., 2005; Matlin and Moore, 2007; Wahl et al., 2009; Huranova et al., 2010; Hoskins et al., 2011). This unidirectional multistep assembly initially reflects the orderly recognition of the primary *cis* sequence elements that define the exon/intron boundaries (the 5' and 3' splice sites) and the associated 3' sequences for intron excision (the polypyrimidine (Py) tract, and the branch point sequence (BPS)). The first stable complex formed during the assembly of the spliceosome is the early spliceosome complex (E), also known as the commitment complex. In this step, the 5' splice site of the intron is bound by the U1snRNP, and the splicing factors SF1 and U2AF cooperatively recognize the BPS, the Py, and the 3' splice site (Berglund et al., 1997; Berglund et al., 1998). These two factors promote the recruitment of the U2snRNP to the branch point in an ATP-dependent manner to form complex A (MacMillan et al., 1994). Complex B is formed by the subsequent binding of the U4/U6-U5 tri-snRNP. After RNA-RNA and RNA-protein rearrangements at the heart of the spliceosome, U1 and U4 are released. The release of U1 and U4 produces the C (or catalytic) complex that is ultimately responsible for catalyzing the intron excision and exon-exon ligation reactions (Fig. 1A).

Alternative exons, as opposed to constitutive exons, can be alternatively spliced to give rise to two or more mRNAs from a single pre-mRNA. Interestingly, an essential property of the genome sequence of eukaryotes (especially in higher eukaryotes) is the suboptimal arrangement and poor conservation of the core *cis*-elements that drive the recruitment of the spliceosome (Izquierdo and Valcarcel, 2006; Mendes Soares and Valcarcel, 2006). This "loose" or suboptimal recognition of splicing signals in alternative exonic/intronic sequences is critical because it renders the sequences sensitive to the combinatorial regulation exerted by different elements, including transcription-coupled alternative splicing regulation. Additional *cis* elements, known as exonic and intronic splicing silencers or enhancers (ISS, ISE, ESS, and ESE), participate in the regulation of alternative splicing. These elements are most often recognized and acted on by specific RNA-binding proteins that include heterogeneous nuclear ribonucleoproteins (hnRNPs) and serine/arginine-rich (SR) proteins (Wu and Maniatis, 1993) (Fig. 1B). These positive and negative regulators together with an ever-increasing number of additional auxiliary regulators, act on a given alternatively processed pre-mRNA sequence as defined by the "splicing code". This "splicing code" is still incompletely characterized, and it is highly contextual in nature, varying among different tissues and distinct homeostatic environments (Zhang et al.; Lin and Fu, 2007; Barash et al., 2010).

Initial work supporting the existence of co-transcriptional splicing comes from direct observation. Electron microscopy has shown that intron lariat formation and/or the removal of intronic sequences can occur in nascent transcripts that are still tethered to their DNA templates (Beyer and Osheim, 1988; Osheim et al., 1988). Several high-resolution light microscopy studies using RNA-FISH probes to specifically discriminate between unprocessed and processed mRNA species revealed that in a significant

proportion of genes, either endogenous or transiently expressed synthetic reporter constructs, the processing of intronic sequences was performed at or very close to the active template gene (Johnson et al., 2000; Custodio et al., 2004). Analogous "direct" observation of co-transcriptional pre-mRNA processing has been achieved by comparing the patterns observed in nascent chromatin-associated RNA with those found after the release of the RNA into the nucleoplasm (Pandya-Jones and Black, 2009). Other more indirect approaches have also provided support for a model of co-transcriptional processing for the vast majority of intronic sequences. For example: (1) multiple components of the spliceosome physically interact with the transcription apparatus and components of the chromatin environment, including modified histones (see below); (2) ChIP analysis suggests a highly conserved stepwise assembly of spliceosomal subunits on the chromatin template (Listerman et al., 2006); (3) measurement of the rates of RNAPII transcription and pre-mRNA splicing *in vivo* on endogenous human genes show that splicing occurs long before the ends of the gens are transcribed (Singh and Padgett, 2009), and (4) co-transcriptional pre-mRNA processing can be reproduced *in vitro* in an RNAPII-specific manner (Natalizio and Garcia-Blanco, 2005) (Das et al., 2006; Yu et al., 2010). Consistent with these findings, total RNA sequencing reveals widespread co-transcriptional splicing in the human brain (Ameur et al., 2011). Similarly, measurements of complete and incomplete splicing rates through the analysis of genome-wide RNAseq datasets in various sub-cellular fractions is consistent with widespread co-transcriptional splicing in humans (H. Tilgner, R. Guigo, and the ENCODE CONSORTIUM, personal communication).

Based on these observations, it was initially thought that the physical coupling of transcription and downstream maturation events was a constitutive and intrinsic property of the machinery involved in these phenomena and that this coupling was

therefore not necessarily subject to regulation. Early models depicted the spliceosome as being optimally assembled when recruited co-transcriptionally to a ternary complex composed of the RNAPII holoenzyme, the chromatin template, and the nascent transcript. This view prompted pioneering *in vitro* studies demonstrating that indeed, RNAPII-associated co-transcriptional pre-mRNA constitutive splicing is significantly more efficient than the uncoupled processing of a purified, pre-transcribed template or the processing of a transcript that is being synthesized by an enzyme that cannot couple with the spliceosome, such as the T7 phage RNA polymerase (Natalizio and Garcia-Blanco, 2005; Hicks et al., 2006; Das et al., 2007). It is worth noting that these *in vitro* co-transcriptional models required an intact RNAPII CTD, but the addition of a complete CTD to the T7 polymerase did not confer enhanced co-transcriptional processing (Natalizio et al., 2009). However, we know that some pre-mRNA sequences are spliced after the release of the transcript depending on the position of the intron (Bauren and Wieslander, 1994; Pandya-Jones and Black, 2009). In some cases, splicing occurs significantly far away from the chromatin template (Johnson et al., 2000). Recently published work reporting the real-time visualization of intron removal in live mammalian cells has revealed that although constitutive introns are indeed co-transcriptionally removed, there are some particular cases where processing is delayed until after transcription has been completed (Vargas et al., 2011). This result most likely reflects some regulatory effects on gene expression programs. Thus, splicing may occur after transcription is completed and co-transcriptional splicing is not obligatory but is rather a general rule with exceptions.

Although it is generally accepted that the physical co-occurrence of transcription and splicing favors both processes, it is also likely that physical coupling is not absolutely required for either process. Rather, co-transcriptional processing is predominantly

regarded as a widespread property of a significant share of intronic sequences. Co-transcription allows the development of multiple mechanisms that exquisitely co-regulate pre-mRNA synthesis and alternative processing. We use the phrase "functional coupling" to refer to the mutual influence of transcription and pre-mRNA splicing on one another. The functional coupling of transcription and pre-mRNA processing appears to be a major element in the regulation of gene expression programs, driving complex cell functions such as tissue-specific differentiation and development (Barboric et al., 2009) and the integration of DNA damage signaling and apoptosis (see below). The immediate question is: which molecular mechanisms drive these functional coupling events? As we will see in the following sections, the functional coupling of transcription and alternative splicing is largely based on the shared properties of alternative splicing-regulatory sequences and their derived regulatory mechanisms. The available data show that transcription influences alternative splicing by recruiting positive and negative regulators to the nascent pre-mRNA ("physical coupling") or by presenting a specific sequence context to the spliceosome by altering the synthesis rate of the nascent transcript ("kinetic coupling"). These two pathways, which are proposed to explain the link between transcription and splicing, are distinct but not mutually exclusive models.

Physical coupling: interactions between the transcription and splicing machineries

One mechanism that may explain the functional coupling between transcription and splicing is the establishment of physical interactions among components of both machineries. Many examples of these interactions have been described throughout the literature (see below), and components of these macromolecular complexes have been purified using a variety of approaches. These interactions are also believed to underlie

regulatory mechanisms shared by both processes, such as post-translational modifications that transduce signaling cues or the reading of chromatin marks (Sims et al., 2007; Lenasi and Barboric, 2010).

It is believed that the carboxyl terminal domain (CTD) of the large subunit of RNAPII plays a major role in the co-transcriptional coupling of RNA biogenesis and processing. The CTD is comprised of 52 tandem repeats in mammals (26 in yeast) of the heptapeptide YSPTSPS, with a heterogeneous degree of degeneration. The CTD constitutes an extremely versatile and flexible platform for the recruitment of factors to the nascent transcript. This recruitment promotes the co-transcriptional accomplishment of gene expression-related functions, ranging from pre-mRNA capping, 3'-end processing and excision, pre-mRNA splicing, and chromatin remodeling (Phatnani and Greenleaf, 2006).

Specific residues within this repeat are targeted for post-translational modifications in a differential manner that depends on the stage of transcription. The modification of the CTD is believed to constitute an extremely versatile mechanism for encoding and transducing information about the functional status of the transcription complex to the different gene expression machineries. Dynamic changes in post-transcriptional modifications generate a "CTD code" (Fig. 2) that contributes to the recruitment and release of processing factors at specific stages of transcription (Komarnitsky, 2000; Buratowski, 2003; Egloff and Murphy, 2008). Hypophosphorylated RNAPII assembles into pre-initiation complexes positioned at gene promoters, whereas hyperphosphorylated RNAPII associates with the elongation stage. Cyclin-dependent kinase 7 (CDK7), which is part of the general transcription factor IIH (TFIIH), is responsible for the phosphorylation of Ser5. This modification has been linked with transcription initiation and is preferentially associated with the 5'-end of the genes. Ser2

phosphorylation by CDK9, the catalytic subunit of the positive transcription elongation factor (P-TEFb), is associated with the 3'-end of the genes. Classically, P-TEFb was thought to be the sole CTD kinase responsible for the conversion of RNAPII into its processive elongating form; however, recent studies have shown that *Drosophila* CDK12 and human CDK12 and CDK13 are also Ser2-CTD kinases (Bartkowiak et al., 2010). These results impact our understanding of CDK9/CTD phosphorylation and open the door to revisions of the current model on CTD phosphorylation and transcriptional elongation. Other residues within the consensus heptapeptide are substrates for phosphorylation, including threonine, tyrosine, and the less conserved Ser7 (Zhang and Corden, 1991; Baskaran et al., 1993; Zhou et al., 2009; Hsin et al., 2011). Recent genome-wide analysis in yeast detected Ser7 phosphorylation at the 5'-ends of genes and in introns, suggesting that Ser7 marks could facilitate elongation and splicing (Kim et al., 2010; Tietjen et al., 2010). Glycosylation, methylation, ubiquitylation and proline isomerization of residues within the CTD have also been reported (Kelly et al., 1993; Morris et al., 1999; Li et al., 2007b; Sims et al., 2011).

Evidence that the CTD is required for splicing has been observed both *in vivo* and *in vitro*. Truncation of the CTD reduces splicing efficiency *in vivo* (McCracken et al., 1997; Fong et al., 2003; Rosonina and Blencowe, 2004). The transcriptional activation of RNAPII-dependent genes induces the spatial recruitment of splicing factors to the transcription site, and this reaction is dependent on the integrity of the CTD (Du, 1997; Misteli, 1999). *In vitro* assays have shown that purified hyperphosphorylated RNAPII is able to activate pre-mRNA splicing, but hypophosphorylated RNAPII can inhibit splicing (Hirose, 1999). Finally, recombinant CTD can also stimulate the splicing reaction *in vitro*. However, this finding is only valid in the context of exon definition, when both the 3' and 5' splice sites are accessible (Zeng and Berget, 2000).

Despite the significant efforts focused on the characterization of the CTD, the specific role of the CTD in the regulation of splicing is not completely understood. The role of the CTD is likely to be variable and context-specific, and probably depends on the subset of effectors and/or regulators recruited in each condition. In some cases, RNAPII may not be sufficient for splicing factor accumulation and instead a more complex and dynamic mechanism for recruitment that involve direct interactions with the nascent RNAs may occur (Kotovic et al., 2003; Gornemann et al., 2005; Lacadie and Rosbash, 2005; Listerman et al., 2006; Das et al., 2007). It has been surmised that "scaffolding", adaptor or bifunctional proteins may regulate the functional coupling of transcription and splicing in a manner similar to the specification of signaling pathways in other cell systems. Indeed, some nuclear factors, also known as "cross-talk" factors or "coupling mediators" interact with both the transcription and splicing machineries. These nuclear factors have therefore been proposed to regulate the physical coupling of the two processes. Although their differential features are significant, these nuclear factors all share the following properties: first, they are capable of physically interacting with components of both the RNAPII transcription apparatus and the spliceosome, mostly through the CTD of RNAPII. In fact, the specific recognition of hyperphosphorylated forms of the CTD is currently regarded as a significant feature supporting the putative role of a given factor on co-transcriptional pre-mRNA processing. Second, these forms contain multiple protein-protein interaction domains or motifs, which confer the potential ability to act as scaffolding elements. Third, these factors are capable, at least under specific experimental conditions, of modulating transcription, splicing, or both (Kornblihtt et al., 2004; Allemand et al., 2008).

Transcription-splicing coupling factors

The family of serine/arginine rich-proteins (SR proteins) includes several conserved RNA-binding factors that have been primarily characterized as splicing regulators, although they seem to play a role in several other mRNA metabolism processes such as nucleocytoplasmic export and translation. The activity of SR proteins is tightly regulated through several mechanisms, including tissue-specific differential expression, signal-dependent phosphorylation, subnuclear distribution and conditional competition with other factors (for an extensive discussion of different aspects of these proteins, please refer to (Lin and Fu, 2007)). Physical interactions between SR proteins and the CTD of RNAPII have been reported, but the functional relevance of this association is not yet known (Yuryev et al., 1996; Bourquin et al., 1997). Indeed, this association was proposed to be required for the appropriate regulation of the recruitment of SR proteins from specific subnuclear structures called nuclear speckles to transcriptionally active sites (Misteli, 1999). SR proteins have been found to co-purify with the RNAPII holoenzyme (Das et al., 2007). Moreover, SR proteins partially mediate the enhancement of *in vitro* co-transcriptional pre-mRNA processing presumably by facilitating the recruitment and assembly of the early spliceosome on nascent RNAPII transcripts (Das et al., 2006 and 2007). The elongation-dependent influence of some of these factors, including SRp20 and SRSF1 (formerly called SF2/ASF), on specific alternative splicing minigene models has been shown to be dependent on the integrity of the RNAPII CTD as well as its phosphorylation state (de la Mata and Kornblihtt, 2006; Barboric et al., 2009). In cells, the depletion of the essential splicing factor SC35 led to a diminished recruitment of P-TEFb to the RNAPII complex and, consequently, a lower CTD phosphorylation level and increased accumulation of the RNAPII large subunit on the 5' ends of genes, both of which are hallmarks of impaired elongation (Lin et al., 2008). The specific mechanisms by which SR proteins establish the functional coupling

between transcription and alternative splicing are not yet completely understood. However, interactions involving both the co-transcriptional assembly of pre-spliceosomal complexes and the regulation of the progression towards an elongation-competent state by the RNAPII complex are likely to be important in this regulation.

Other splicing factors participate in the functional coupling of transcriptional elongation and alternative splicing. The splicing factor SKIP interacts with PTEF-b and positively influences transcription elongation by the HIV-1 Tat activator (Bres et al., 2005). Tat-SF1, the mammalian homolog of the yeast U2 snRNP component *CUS1*, has also been proposed to mediate the splicing-dependent regulation of transcription elongation *in vitro* (Fong and Zhou, 2001; Maniatis and Reed, 2002).

An interesting group of factors has been shown to play dual roles both in transcription and splicing regulation and/or to physically interact with components of both the splicing and transcription machinery. These factors include, among others, PSF, p54nrb/NonO, Prp40, FBP11, and TCERG1. Some of these factors contain tandem repeats of WW (Chan, 1996) and FF (Bedford and Leder, 1999) domains; this feature might define a subset of potential coupling factors.

PSF and $p54^{nrb}$/NonO are two RRM domain-containing proteins that are ubiquitously expressed in human tissues. These proteins share 71% identity over a 320-amino-acid region encompassing their RNA recognition motif (RRM) (Dong et al., 1993). Intriguingly, these proteins have been found to interact with each other through their RRM domains, although the functional consequences of this interaction are not known (Peng et al., 2002). PSF and $p54^{nrb}$/NonO interact with either hypophosphorylated or hyperphosphorylated forms of the RNAPII CTD (Emili et al., 2002). It has been shown that $p54^{nrb}$/NonO associates with 5' splice sites in high molecular weight complexes containing the U1 and U2 snRNPs and mostly (but not exclusively) phosphorylated

forms of RNAPII, both *in vitro* and *in vivo* (Kameoka, 2004). Similarly, PSF can be found in higher-order complexes assembled *in vitro* that contain all of the spliceosomal subunits (Peng et al., 2006). The fact that PSF was found to interact with the hypophosphorylated forms of the RNAPII and spliceosomal components involved in the 5' splice site is particularly interesting given that hypophosphorylated RNAPII is detected in the paraspeckle nuclear compartment. Moreover, PSF is enriched at paraspeckles, and mass spectrometry analysis of hypophosphorylated RNAPII complexes isolated both U1 snRNP proteins and SR proteins (Fox et al., 2005; Das et al., 2007; Fox and Lamond, 2010). Based on these data, it is tempting to speculate that PSF may prime interactions between the transcription initiation complex and the early spliceosome. Furthermore, the paraspeckle region might be the location at which these events take place. Consistent with these observations, *in vivo* experiments suggest that the PSF is capable of mediating transcription activator-dependent stimulation of pre-mRNA processing through its binding to the CTD of RNAPII (Rosonina et al., 2005).

The yeast splicing factor Prp40 contains two WW domains followed by four FF domain repeats, and it was first characterized as a splicing regulator. Prp40 is an essential U1 snRNP-associated protein that participates as a scaffold in the earlier steps of spliceosome complex formation (Kao and Siliciano, 1996). The WW domains of Prp40 recognize the consensus sequence PPXY in BBP, which is the yeast ortholog of the splicing factor SF1. BBP interacts directly with Mud2p, which is the yeast homolog of human U2AF^{65}, which binds to the polypyrimidine region at the 3'-end of introns. Based on these interactions, it has been proposed that Prp40 helps to define a bridging interaction that links both ends of the intron (Abovich and Rosbash, 1997). Prp40 also binds to the yeast U5-associated protein Prp8. Therefore, Prp40 might contribute to the recruitment of the U5 snRNP to the early spliceosome during cross-intron bridging.

According to the currently prevailing sequential model, this interaction may destabilize U1 snRNP binding, allowing for the progression towards a mature spliceosome (Abovich and Rosbash, 1997). Intriguingly, a recent study demonstrated that, in the absence of the WW domains of Prp40, no defects in *in vivo* splicing, *in vivo* U1 or U2 snRNP recruitment, or *in vitro* early splicing complex assembly were detected (Gornemann et al., 2011). Moreover, the Prp40 WW domains were dispensable for yeast viability. Delays in U5snRNP recruitment and altered spliceosome formation *in vitro* did not prevent spliceosome maturation or changes in RNAPII distribution along the transcripts. One plausible explanation that may reconcile these apparently opposing observations is that the Prp40 WW domains are required to act as a scaffold for early spliceosome assembly in only a subset of genes or in specific functional contexts. Finally, Prp40 was the first splicing factor that was shown to bind the phosphorylated CTD of RNAPII via its WW domains (Morris and Greenleaf, 2000).

There are two putative mammalian orthologs of Prp40, FBP11 and HYPC. FBP11, also known as HYPA or Prpf40A, comprises two WW and six FF domains. Similar to yeast Prp40, FBP11 interacts with SF1 and $U2AF^{65}$, which suggests that it may play a role in the recognition of the 3'-end of introns (Lin et al., 2004), However, unlike Prp40, FBP11 has not been strictly characterized as a component of the U1snRNP. Structural characterization of the binding mechanism and involved surfaces in the FF domain-containing region of FBP11 revealed that FBP11 is able to recognize phosphorylated sequences such as the CTD of RNAPII (Allen et al., 2002). Intriguingly, this interaction is preferentially established by the FF1 repeat, whereas Prp40 interacts weakly with the phosphorylated CTD through its FF2 repeat (Gasch et al., 2006). This observation illustrates that this family of factors displays significant differences regarding the

binding specificity of their WW and FF domains despite their apparent structural conservation.

HYPC, also known as Prpf40B, is the least well characterized of the putative mammalian Prp40 orthologs. It has an analogous modular structure of two WW domains and four FF domains. HYPC, as well as FBP11, has been implicated in the pathogenesis of neurological disorders, including Huntington's disease and Rett syndrome (Faber et al., 1998; Passani et al., 2000; Buschdorf and Stratling, 2004). High-resolution confocal microscopy analysis revealed that HYPC shows a precise overlap with SC35 in the splicing factor-rich nuclear speckles (S. Becerra and C. Suñé, unpublished data), which suggests that HYPC is involved in pre-mRNA processing.

FBP21 is the only member of the family of potential coupling factors that lacks FF domain repeats. However, similarities between its WW domains and the interdomain spacer length and those of the above-mentioned Prp40 paralogs indicate that FBP21 might play a conserved role in splicing as well. This theory is supported by data showing that FBP21 interacts with the U1 snRNP protein U1C, the core snRNP proteins SmB and SmB', and the branchpoint binding protein SF1, suggesting that FBP21 may play a role in the cross-intron bridging of U1 and U2 snRNPs in the mammalian A complex (Bedford et al., 1998). Importantly, the WW domains of FBP21 are critical for its splicing activator function and for its interaction with SIPP1 (Huang et al., 2009), a splicing factor that associates with polyglutamine-tract-binding protein (PQBP1) and protein phosphatase 1 (PP1).

TCERG1, also known as CA150, is a nuclear protein that has been implicated in transcriptional elongation and pre-mRNA splicing. TCERG1 contains three WW domains in its amino-terminal half and six FF repeat domains in its carboxyl-terminal half (Suñé et al., 1997) and associates with hyperphosphorylated RNAPII and with

transcriptional elongation and splicing components through both its WW and FF domains (Goldstrohm et al., 2001a; Lin, 2004; Smith et al., 2004; Sanchez-Alvarez et al., 2006). TCERG1 has been identified in highly purified spliceosome fractions (Neubauer et al., 1998; Makarov et al., 2002; Rappsilber et al., 2002; Deckert et al., 2006). TCERG1 localizes at the interface of nuclear speckles and what are presumably nearby transcription sites (Sanchez-Alvarez et al., 2006). Several studies have demonstrated that TCERG1 can affect the splicing of several splicing reporters (Carty et al., 2000; Lin, 2004; Cheng et al., 2007; Pearson et al., 2008; Sánchez-Álvarez et al., 2010) and of putative cellular targets identified by microarray analysis following TCERG1 knockdown (Pearson et al., 2008). Given these data, TCERG1 has been observed as a possible factor in coordinating transcriptional elongation and splicing, a suggestion that was in fact proposed in the first Gene review on co-transcriptional splicing (Goldstrohm et al., 2001b) (see below). Like other members of this family of proteins, TCERG1 has been implicated in the pathogenesis of neurodegenerative diseases (Holbert, 2001; Arango et al., 2006; Andresen et al., 2007).

Kinetic coupling: how elongation rate affects splice site choice

RNAPII transcription initiation and pre-initiation complex assembly were initially described as the rate-limiting steps of transcriptional activation. However, transcriptional elongation is also a very highly regulated process. Numerous genes have been shown to be controlled at this stage, and a great variety of factors that affect this step of the transcription cycle have been identified (Sims et al., 2004). Increasing evidence shows that after promoter clearance, which depends on CTD phosphorylation, RNAPII stalls approximately 30-50 nucleotides downstream of the transcription start site. This promoter-proximal pause was first described for the *Drosophila* heat shock

genes and in the c-myc and HIV-1 genes (Bentley and Groudine, 1986; Kao et al., 1987; O'Brien and Lis, 1991; Krumm et al., 1992). More recent, genome-wide studies suggest that the promoter-proximal pause is a global tool for gene regulation (Core and Lis, 2008; Nechaev and Adelman, 2008). The promoter-proximal pause has been proposed as a checkpoint to ensure that only correctly assembled RNAPII complexes continue elongating through the gene (Levine, 2011). Given that different splicing sites of a transcript are synthesized and co-transcriptionally exposed to the splicing machinery, it is then plausible that the RNAPII elongation rate might modulate alternative splicing. This "kinetic" regulation of alternative splicing model proposes that the transcriptional elongation rate is critical for the outcome of competing splicing reactions that occur co-transcriptionally. Some authors have viewed this regulation as a 'window of opportunity' (Perales and Bentley, 2009), referring to the period during which the spliceosome is able to functionally assemble at one splice site before it is subject to competition from a downstream splice site.

The first evidence of this kinetic coupling came from a study by Eperon and colleagues, which showed that the rate of RNA synthesis could affect the secondary structure of the mRNA and, in turn, affect alternative splicing (Eperon et al., 1988). A similar mechanism of regulation was also suggested by experiments in which the insertion of an RNAPII pause site downstream of weak alternative splice sites caused a reduction in the amount of regulated exon skipping by delaying the synthesis of an inhibitory element (Roberts, 1998). These studies hinted that the rate of transcript synthesis itself could define the specific sequence context presented to the co-transcriptional processing machinery.

One of the most studied experimental systems that largely (but not exclusively) fits this kinetic model and can be used to illustrate this model is that of the fibronectin extra

domain I (EDI). This exon is preceded by a rather weak or degenerate 3' splice site, which is best utilized when the processivity of RNAPII is low and the splicing machinery has sufficient time to recognize it as a splice site. When RNAPII processivity is high due to elements such as transcriptional coactivators, open chromatin states or strong upstream promoters, it is more likely that the weak splice site will be presented simultaneously with the downstream, strong 3' splice site. The strong 3' splice site can then outcompete the weak splice site, favoring the skipping of the regulated exon (Cramer et al., 1999; Kadener et al., 2001; Kadener et al., 2002; Nogues et al., 2002). Direct evidence supporting the kinetic model was obtained by the exogenous expression of RNAPII mutant proteins that exhibit intrinsically low processivity. In these elegant, seminal experiments, the RNAPII-dependent transcription of the cell was driven exclusively by different α-amanitin-resistant polymerases; invariably, the polymerases with intrinsically low processivity elicited greater inclusion of the EDI exon (de la Mata et al., 2003). Interestingly, one mutation, termed *C4* (Coulter and Greenleaf, 1982), which provokes a decrease in the rate of the RNAPII processivity, was found to be associated with changes in the alternative splicing patterns of transcripts encoded by the *Drosophila ultrabithorax* gene, underscoring the potential physiological relevance of this phenomenon (de la Mata et al., 2003).

These observations may reflect a more general mechanism. Bentley and coworkers reported several years ago that RNAPII is paused at both 5' and 3' ends of genes and that this pausing correlates with the recruitment of processing factors (Glover-Cutter et al., 2008). More recent studies using high-resolution ChIP methods, demonstrate the preferential accumulation of RNAPII at specific regions within the body of genes (Alexander et al., 2010; Carrillo Oesterreich et al., 2010). Similarly, global analysis of nascent transcripts in yeast have shown a peak of RNAPII accumulation within terminal

exons, located ~ 250 nt upstream of the poly(A) site (Carrillo Oesterreich et al., 2010). Pausing near the 3' end of introns that correlates with the recruitment of splicing factors has also been described in intron-containing genes (Alexander et al., 2010). Thus, it is proposed that transcriptional pausing is imposed by a regulatory checkpoint that is associated with co-transcriptional splicing (Alexander et al., 2010). Therefore, proteins acting at the interface of these processes (the "coupling factors") would act as checkpoint factors to regulate co-transcriptional splicing. Strong data that may support this hypothesis have been reported in recent work showing that the putative coupling factor TCERG1 regulates the alternative splicing of the *Bcl-x* gene by modulating the rate of RNAPII transcriptional elongation (Montes et al., 2012). The authors of the aforementioned study speculate that TCERG1 modulates the elongation rate of RNAPII to relieve pausing, thereby acting as a checkpoint regulator to promote co-transcriptional splicing.

The kinetic coupling model could explain the accurate control of the alternative processing of very specific sequences through the tight control of RNAPII processivity at the boundaries of the region of interest. This scenario has been proposed for the alternatively spliced variable region of the CD44 gene, in which a subset of regulatory RNA-binding factors could participate in the local control of RNAPII through the chromatin remodeling factor SWI/SNF, which in turn determines the processing profile of the resulting transcripts (Batsche et al., 2006). SWI/SNF can also associate with nascent transcripts to regulate alternative pre-mRNA splicing (Tyagi et al., 2009). Generally, RNAPII pausing could be a way to ensure its optimal activity, especially when transcribing large, highly regulated genes (Gilchrist et al., 2010; Levine, 2011). Insufficient time spent at the pause site could therefore result in more error-prone transcription and the production of inadequately processed transcripts.

The influence of chromatin in the regulation of alternative splicing

The fundamental unit of chromatin is the nucleosome, which is composed of DNA wrapped around an octamer of the four core histones (H3, H4, H2A, and H2B) and separated from its neighboring nucleosomes by linker DNA (Izban and Luse, 1991). The reversible modification of histones and the positioning of nucleosomes along the chromatin template regulate transcription (Orphanides and Reinberg, 2000; Hodges et al., 2009). This finding was, to some degree, expected. However, the existence of connections between chromatin state and splicing was rather unexpected. A link between chromatin and splicing was first suggested when different RNA products were obtained when studying transcription from a collection of adenovirus-based templates, with identical trans-acting factors (Adami and Babiss, 1991). These authors hypothesized that changes in elongation rate, transcriptional pauses, or the transcriptional machinery could affect splicing patterns. At the same time, it was reported that the distances between the 5' and 3' splice sites followed a periodic pattern resembling nucleosome phasing, suggesting that nucleosomes are positioned according to exon/intron boundaries and thus reflect connections between chromatin and splicing (Beckmann and Trifonov, 1991). Accumulating evidence supports these hypotheses: cell replication alters alternative splicing by favoring compact chromatin and slowing RNAPII processivity (Kadener et al., 2001); the inhibition of topoisomerase-I by camptothecin promotes RNAPII stalling, which favors the co-transcriptional recruitment of splicing factors (Listerman et al., 2006); SW1/SNF chromatin remodeling factors have been shown to promote the inclusion of a cluster of exons in the endogenous CD44 gene (Batsche et al., 2006) and to regulate the alternative pre-mRNA processing of a specific subset of transcripts in *Drosophila* (Waldholm et al.,

2011); histone modifying enzymes interact with various components of the splicing machinery (Martinez et al., 2001; Ohkura et al., 2005; Gunderson and Johnson, 2009); and high-throughput sequencing-based techniques have revealed that nucleosome positioning and specific histone marks are significantly correlated with the intron/exon structure of genes, underlining the potential impact of chromatin structure on splicing (see below).

Several recent studies have revealed a specific, genome-scale nucleosome occupancy pattern in exon-intron regions. High-throughput immunoprecipitation and sequencing of nucleosome-associated chromatin have provided evidence that nucleosomes are predominantly enriched in exons (Andersson et al., 2009; Schwartz et al., 2009; Spies et al., 2009; Tilgner et al., 2009). The preferential occupancy of nucleosomes on exons is independent of transcription because it is also observed in non-expressed genes (Tilgner et al., 2009). Exons have higher GC content compared to introns, and GC-enriched regions have been suggested to be preferred for the assembly of nucleosomes (Peckham et al., 2007). Therefore, some of these studies suggested a correlation between GC content and nucleosome occupancy. However, GC-richness is likely not the only factor that determines nucleosome patterning. Data obtained from bioinformatics studies strongly suggest a functional link between nucleosome occupancy and splicing. Notably, the average length of the DNA wrapped around a histone octamer is 147 nucleotides in length, which is approximately the length of an average mammalian internal exon, suggesting a conserved role for the nucleosome in the exon definition process (Schwartz et al., 2009; Tilgner et al., 2009). Moreover, nucleosome density in exons correlates with splice site strength and the properties of flanking introns. The correlation is stronger in weak exons as well as in exons that are surrounded by long introns (Spies et al., 2009; Tilgner et al., 2009). Intriguingly, alternatively spliced exons

are more enriched in nucleosomes than constitutively spliced exons (Schwartz et al., 2009), and longer exons are more enriched in nucleosomes than shorter exons (Schwartz et al., 2009; Spies et al., 2009; Tilgner et al., 2009). These observations might suggest that exonic nucleosomes do not act as a mark for spliceosome recognition. Despite the discrepancies, these data strongly support a role for nucleosome patterning in the regulation of RNA splicing. In the absence of further experimental evidence, two non-mutually exclusive hypotheses have been proposed. First, nucleosomes may act as "speed-bumps" for RNAPII, which would affect transcription rates and in turn modulate alternative splicing through the kinetic model discussed above (Schwartz et al., 2009; Spies et al., 2009). Secondly, nucleosomes could interact with splicing factors during transcription or facilitate interactions between splicing factors during exon definition and other splicing assembly steps (Tilgner et al., 2009). Additional experiments are required to elucidate the mechanisms by which nucleosomes affect splicing.

Histones are substrates for numerous post-translational modifications, and their N-terminal tails are subjected to acetylation, methylation, phosphorylation, ubiquitylation or sumoylation (Kouzarides, 2007). Most of these modifications are reversible, and the enzymes responsible of their removal have been characterized extensively (Bannister et al., 2002; Goto et al., 2002; Sugiyama et al., 2002; Yang and Seto, 2007). The combinatorial and dynamic nature of the post-translational modifications of histones enables them to be an extremely versatile and complex combinatorial system. The "histone code" is capable of encoding vast amounts of information, which regulates the expression of the underlying DNA sequences, by affecting the recruitment, assembly and function of multiple nuclear complexes.

Histone methylation, acetylation and phosphorylation have recently emerged as major splicing regulators. For example, treatment with deacetylase inhibitors has been shown

to affect the alternative splicing of several reporter and endogenous genes (Nogues et al., 2002; Hnilicova et al., 2011), and histone acetylation is important for co-transcriptional spliceosome assembly (Gunderson and Johnson, 2009); histone H3 phosphorylated on serine 10 releases SR proteins from chromatin during the cell cycle (Loomis et al., 2009); and histone H3 lysine 4 tri-methylation (H3K4me3) enhances the recruitment of spliceosomal components to gene units (Sims et al., 2007). These data support a model in which histone modifications actively participate in the coupling of transcription and splicing. However, further study is required to elucidate their functional impact in the cell and the extent to which these potential mechanisms function throughout the genome.

Genome-wide analyses revealed specific histone modification patterns within exons and introns. One early study revealed that histone H3 lysine 36 tri-methylation (H3K36me3) was globally enriched in exons (Kolasinska-Zwierz et al., 2009). Parallel studies have identified additional chromatin modifications in exons and introns (Andersson et al., 2009; Schwartz et al., 2009; Spies et al., 2009; Tilgner et al., 2009; Dhami et al., 2010; Huff et al., 2010; Luco et al., 2010). It remains controversial whether these marks reflect nucleosome occupancy within exons (Schwartz et al., 2009; Tilgner et al., 2009) or constitute independent, relevant patterns with an intrinsic functional significance (Kolasinska-Zwierz et al., 2009; Spies et al., 2009). In support of the latter hypothesis, the average enrichment landscape for the H3K36me3 mark is actually contradictory to that of nucleosome occupancy because H3K36me3 is more highly enriched in constitutive exons than in alternative exons. Importantly, the relative enrichment of a number of histone marks differs between intronic and exonic sequences. The enrichment of some of these modifications depends on the relative positioning of the exon within the gene. For example, histone H4 lysine 20 methylation (H4K20me1) and

histone H2B lysine 5 methylation (H2BK5me1) are found preferentially at exons located at the 5' end of genes. In contrast, H3K36me3 tends to accumulate at the 3' end of the gene (Hon et al., 2009). Other marks exhibit a relative enrichment throughout the body of a gene reflective of its activation state, such as the histone H3 lysine 27 tri-methylation (H3K27me3), which is highly enriched at the exons of weakly expressed genes and at the introns of highly expressed genes (Dhami et al., 2010). These observations suggest that histone modifications play an active role in the regulation of alternative splicing because they constitute potentially informative hallmarks of the identity of an exon or an intron within a specific activation context (Fig. 3).

How can chromatin labels modulate co-transcriptional alternative splicing? From our current knowledge, mainly gathered from the fields of transcription and genome maintenance, these post-translational modifications are recognized with a striking degree of specificity by highly specialized modules. Following recognition, these modifications either exert a direct influence on the gene expression machinery or, more often, play a role as adaptors or scaffolding molecules for different groups of effector complexes. This model seems to be valid for the regulation of alternative splicing through the control of histone marks. In a recent pioneering study based on the mutually exclusive alternative splicing of the exons IIIb and IIIc of the *FGFR2* gene (Carstens et al., 2000), it was observed that the H3K36me3 and H3K4me1 modifications are enriched when exon IIIc is included, but H3K27me3 and H3K4me3 accumulate preferentially when exon IIIb inclusion is favored (Luco et al., 2010). In this experimental system of alternative splicing, the RNA binding protein PTB functions as a major regulator by reducing the inclusion of IIIb and favoring the inclusion of IIIc. The PTB effect is mediated by MRG15, an adaptor protein that specifically binds H3K36me3, which is relatively enriched in cell types that exhibit greater IIIc inclusion

(Luco et al., 2010). Detailed integrative studies merging the available information about the alternative splicing profiles elicited when manipulating PTB levels, the preferential genome-wide binding sites for PTB and the H3K36me3 landscape might further explain the specificity of these mechanisms for PTB-mediated alternative splicing regulation.

DNA methylation has also been shown to be enriched in exonic regions, although it remains unknown how this modification affects the splicing outcome. Methylation could affect splicing in combination with other factors. For example, in the *CD45* gene, DNA methylation inhibits the inclusion of exon 5, and this methylation pattern opposes that of CTCF, which promotes the inclusion of exon 5. Genome-wide data indicate that CTCF binding downstream of alternative exons correlates with local RNAPII accumulation, suggesting that chromatin structure affects alternative splicing by modulating the RNAPII elongation rate (Shukla et al., 2011). This finding is especially relevant because CTCF acts as a general chromatin organizer and may be involved in numerous chromatin-dependent molecular processes. It is an exciting possibility that methylation patterns, which fluctuate during development, might modulate tissue-specific alternative splicing events through the differential recruitment of CTCF. Further studies are required to reveal whether other DNA binding proteins are involved in the regulation of alternative splicing through a similar mechanism.

Overall, histone modifications may elicit the modulation of pre-mRNA alternative processing both through the control of RNAPII kinetics and the co-transcriptional recruitment of splicing regulators. Histone modifications regulate the conversion between ‘closed’ (silent) and ‘open’ (active) chromatin conformations (Li et al., 2007a; Lee et al., 2010). Histone modifications can thus potentially modulate alternative splicing choices by regulating the RNAPII elongation rate. Supporting this model, intragenic hyperacetylated chromatin induced after neuronal depolarization increases

the RNAPII elongation rate and favours exon skipping in the neural cell adhesion molecule *(NCAM)* pre-mRNA (Schor et al., 2009). However, modifications related to transcriptional repression induce exon inclusion (Saint-Andre et al., 2011). Furthermore, targeting intronic sequences with siRNAs generates a closed chromatin context enriched in heterochromatin marks, blocking RNAPII elongation and affecting the splicing outcome (Allo et al., 2009).

Recent data demonstrate that splicing can also modulate histone modification. Splicing activation enhances the recruitment of methyltransferase HYPB/Setd2 and the subsequent methylation of histone H3 at lysine 36 (H3K36me3) (de Almeida et al., 2011), and splicing inhibition through splice site mutations or spliceostatin A (SSA) treatment produces a redistribution of H3K36me3 towards the 3' end of the gene (Kim et al., 2011). These data again suggest extensive bidirectional communication between epigenetic marks and splicing regulation. The characterization of genome-wide RNAPII elongation profiles compared with global alternative splicing programs and the distribution of the different histone marks will shed light on the functional relevance of these mechanisms in the regulation of gene expression.

Functional impact of the coupling between transcription and splicing

The functional coupling of transcription and alternative splicing provides a number of potential advantages when implemented in a given gene expression program. The first aspect promoted by co-transcriptional pre-mRNA processing is the coordination of these two simultaneous mechanisms of control, enabling the specification of a transcriptional regulation program that is tailored to each splicing isoform and increasing the available possibilities to correctly tune gene expression. The functional coordination of transcription and splicing is a mechanism that would help to restrict a specific alternative processing pattern to the subpopulation or bulk of transcripts whose

synthesis is promoted by a transcriptionally driven response. Coupling would decrease the leakage of a specific alternative splicing mechanism beyond the time window during which the coupled transcriptional response of the cell is activated. Thus coupling would minimize gene expression noise while maximizing the efficiency of the whole process, as recently predicted using mathematical modeling approaches (Aitken et al., 2011; Kalsotra and Cooper, 2011). This aspect may be especially relevant to adaptive mechanisms based on regulatory switches that can radically affect cell fate, such as cell cycle or apoptosis control.

The transitions between two distinct gene expression profiles that are associated with different contexts, such as consecutive developmental stages, are best implemented when the global coordinated regulation of the different involved mechanisms is ensured (Mendes Soares and Valcarcel, 2006; Keene, 2007; Kalsotra and Cooper, 2011). The coupling of transcription and post-transcriptional processing is likely essential for achieving this aim. Alternative splicing has been recently shown to influence the nonsense-mediated decay (NMD) and the usage of different 3' UTR stretches that in turn determine the miRNA sensitivity, subcellular localization, or stability of the resulting transcripts. Alternative splicing is extensively coupled to NMD because certain alternative exons introduce premature termination codons in their mature transcripts upon inclusion, and this can constitute a mechanism for diverting a subset of transcripts toward NMD-dependent degradation. A significant number of the targets of alternative splicing-determined NMD are splicing regulators themselves, potentially creating extensive arrays of regulatory feedback loops (Saltzman et al., 2008; Ip et al., 2010). These findings might support the evolution of an intricate NMD-mediated feedback regulation of splicing factors to achieve a high degree of robustness in post-transcriptional gene expression control, similar to what has been suggested for other

RNA-processing networks (Mittal et al., 2009). Importantly, NMD-eliciting alternative splicing events, which are overrepresented in RNA-binding protein-encoding genes, are affected by pharmacological treatments that alter RNAPII elongation (Ip et al., 2010) (see below).

Coupling mechanisms also enable the establishment of checkpoints or quality controls that would promote each step of pre-mRNA synthesis and processing while ensuring the accomplishment of the essential preceding stages. One example is the significant influence of co-transcriptional pre-mRNA 5' capping on the establishment of processive elongation itself and vice versa (Schroeder et al., 2000; Moteki and Price, 2002). In a similar scenario, a highly processive elongation rate would elicit the recruitment of components of the splicing machinery and other pre-mRNA processing modules, such as 3' end processing complexes (Glover-Cutter et al., 2008). For a number of experimental models, the splicing components and the proximity of the splicing sequences to the 5' terminus of the gene or other regulatory sequences are required for, or at least significantly influence, the activity of the RNAPII complex (Fong and Zhou, 2001; Lin et al., 2008). A potential function for this link is that splicing could be used as a quality checking mechanism in the control of transcription through a given gene. Although it does not fully overlap with the regulation of alternative splicing through co-transcriptional coupling, it is probable that similar molecular events govern this reverse coupling.

Another advantage of the functional coupling is the unification or simplification of the architecture of the signaling networks upstream of the coupled transcriptional and alternative splicing programs. The most intuitive example is represented by the proposed "kinetic model" of functional coupling: if the transcriptional elongation rate of RNAPII determines the alternative processing pattern of the transcripts, a single

signaling input primarily aimed at controlling the transcription of a given gene can also be used to specify the processing profile of the resulting transcripts. This option could be especially relevant in the case of regulatory programs that require rapid and efficient activation. As mentioned above, recent studies have demonstrated the widespread pausing or stalling of RNAPII complexes at gene promoters, relying on transcription elongation activation and release from pausing for induced expression (Muse et al., 2007; Zeitlinger et al., 2007; Gilchrist et al., 2008; Adelman and Rogatsky, 2010; Gilchrist et al., 2010; Nechaev et al., 2010). Importantly, a number of these elongation-driven genes encode regulators of stress-related responses, the innate immune response, and regulators of cell proliferation and apoptosis (Nechaev and Adelman, 2008; Adelman et al., 2009). Transcription elongation control during development may also be more important than previously thought (Levine, 2011). The production of a coherent alternative processing pattern that is regulated in part by transcription elongation control not only ensures the appropriate temporal coordination of transcription and splicing but also simplifies and equally prioritizes the triggering of the alternative splicing when necessary. The regulation of transcription-coupled alternative splicing seems to be deeply ingrained in the gene expression programs associated with such responses, including at least two different branches of the DNA damage response (DDR) and the regulation of apoptosis and cell cycle progression.

The DDR encompasses a heterogeneous group of signaling pathways that have evolved to detect errors introduced in the genomic DNA by different exogenous or endogenous agents. The DDR triggers functional programs to avoid (1) the transmission of damage to the offspring cell population and (2) the disruption of essential processes, such as DNA replication. Alternative splicing is one target of the DDR (Katzenberger et al., 2006; Matsuoka et al., 2007). To date, two different mechanisms that regulate different

control points of the DDR through transcription-coupled alternative splicing have been identified. One recent study identified a specific molecular and functional interaction between the Ewing Sarcoma EWS transcription factor and the stress-associated transcription factor YB-1. This interaction is essential for the appropriate co-transcriptional recruitment of the spliceosome and thus for the timely inclusion of a number of alternative exons of the MDM2 gene (Dutertre et al., 2010). The full-length mRNA sequence yields a functional MDM2 protein, which is in turn capable of down-regulating p53, one of the master regulators of DDR-triggered programs. Triggering a canonical DDR with different genotoxic agents, such as cisplatin and camptothecin, abrogated the interaction between EWS and YB-1. As a result, genome-wide alterations in the alternative splicing pattern of the transcriptome of the cell, including the splicing of the MDM2 gene, were observed. The mechanism of action of EWS/YB-1 in alternative splicing is currently poorly understood. Recently, EWS was found to interact co-transcriptionally with its target RNAs and to be redistributed to nucleoli upon UV light exposure (Paronetto et al., 2011). However, the molecular details of these regulatory mechanisms are currently unknown. Importantly, both inclusion and exclusion events were detected as specific consequences of camptothecin treatment, which suggests that an EWS/YB-1-based mechanism is unlikely to be the sole transcription-coupled effector of the DDR.

A second potential mechanism that links DDR with transcription-coupled alternative splicing was proposed when analyzing RNAPII activity and alternative splicing in UV-treated cells (Munoz et al., 2009). UV exposure elicits a significant global increase in the phosphorylation level of the RNAPII CTD, through an as yet not understood mechanism. Interestingly, UV exposure also elicits a decrease in RNAPII processivity on different gene templates. This change in processivity, in turn, influences pre-mRNA

processing, including that of two essential regulators of apoptosis, *Bcl-x* and *caspase-9*. This mechanism was found to be independent of other DNA damage signaling events, such as BRAD DNA damage-labeling proteins, ATM-dependent signaling, and p53-dependent regulation. This mechanism could thus constitute a novel "pro-apoptotic checkpoint" in UV-mediated DNA damage and cell fate determination.

The data discussed above link the functional coupling of transcription and splicing to the control of apoptosis-regulating networks. A number of genes that control apoptosis exhibit alternative splicing-driven functional switches (Akgul et al., 2004). A typical example is the *Bcl-x* gene, which encodes a master regulatory switch of apoptosis through the modulation of mitochondrial permeability. Alternative splicing yields the large, antiapoptotic isoform Bcl-x_L or the short, proapoptotic regulator Bcl-x_S upon the usage of two alternative 5' splice sites. An analogous example is *Mcl-1*, another Bcl-2 family-related gene. A pioneering genome-wide siRNA survey for regulators of alternative splicing of both genes uncovered the extensive integration of cell cycle progression with these alternative splicing regulation events through aurora kinase A (AURKA) and SRSF1 (Moore et al., 2010). Suggesting a coherent, global apoptosis control pathway based on alternative splicing regulation, Bcl-x and Mcl-1 shared a large number of regulators with a disproportionate enrichment of alternative splicing and cell cycle regulators. Intriguingly, large subsets of RNAPII-dependent transcription regulators scored as positive modulators of the alternative splicing of *Bcl-x*, but not of *Mcl-1*. As mentioned before, the putative coupling factor TCERG1 has been shown to regulate alternative splicing of the *Bcl-x* gene by modulating the rate of RNAPII processivity (Montes et al., 2012). Regardless of the precise mechanisms underlying these differences, these data support an additional tuning mechanism specific for Bcl-x and other apoptosis-control genes, which would appropriately regulate apoptosis

sensitivity in different contexts. These observations lead to several interesting questions. (1) Why is the alternative splicing of Bcl-x, but not Mcl-1, functionally coupled to transcription control? (2) Is the transcription-coupled alternative splicing observed for Bcl-x related to the previously identified AURKA-SRSF1-driven, cell cycle-dependent axis? (3) What is the transcriptional coupling status for other apoptotic regulators that exhibit alternative splicing-dependent switches, such as *Fas* (Izquierdo et al., 2005; Bonnal et al., 2008)?

To date, a small but rapidly growing number of studies have attempted to characterize global transcription-coupled alternative splicing profiles. A recently published study describes the transcriptome landscape in cells treated with either camptothecin or 5,6-dichloro-1-β-D-furanosylbenzimidazole (DRB; a potent and specific inhibitor of the RNAPII CTD kinase CDK9). Both treatments inhibit RNAPII elongation and promote partially overlapping alternative splicing profiles. In both cases, enrichment for alternative splicing events associated with downstream NMD was noted, with a clear preference for splicing factor encoding-genes. The set of genes with alternative splicing patterns induced by camptothecin treatment were enriched in apoptosis-related and cell cycle-related genes, independently from RNA-binding proteins. These observations indicate that transcription-coupled alternative splicing likely depends on specific, conserved molecular mechanisms with functional significance. These observations also suggest that transcription-coupled alternative splicing (and most likely general alternative splicing) has been highly developed and evolutionary conserved in genes encoding RNA binding proteins and RNA splicing regulators. Alternative splicing in these genes most likely serves as a mechanism to regulate and ensure the robustness, stability and versatility of RNA-processing regulatory networks.

Importantly, these reports provide substantial data suggesting that uncoupled alternative splicing may be a significant component of the global regulation of downstream functions. Indeed, subsets of genes are found to either display changes only in their total transcript levels (regulation of transcription without significant change in splicing isoform) or changes in their alternative splicing (alternative splicing without apparent changes in transcriptional output). Similar observations have been reported by different surveys of global alternative splicing profiles associated with other biological mechanisms. These mechanisms include heart development (Kalsotra et al., 2008), neuron differentiation (Jelen et al., 2007; Ule and Darnell, 2007), or the triggering of specific major signaling routes, such as the insulin and wingless pathways (Hartmann et al., 2009). It must be noted that the transcriptome profiling in UV light-induced coupled alternative splicing regulation study discussed above (Munoz et al., 2009) also identified genes that are alternatively spliced but that are not affected at the level of expression. These genes represent approximately 15% of the total number of affected genes. However, it is conceivable that local changes in transcription elongation are capable of profoundly affecting the processing of a specific alternative exon but do not significantly (or detectably) impact the total output of that gene. Therefore, even in these documented models, transcription-dependent alternative splicing regulation may be more common than currently estimated. As discussed above, certain molecular mechanisms may have been implemented to attain precisely this localized effect on alternative splicing through transcriptional modulation without significantly affecting the total output of transcripts. Novel approaches exploiting unbiased, high-resolution technologies based on deep sequencing and integrative analysis with detailed information about RNAPII positioning, chromatin context and bioinformatics analysis of *cis*-features may help to further elucidate this question.

Concluding remarks

The functional coupling of transcription and alternative splicing is emerging as an essential component of gene expression regulation. Despite considerable efforts, numerous questions remain regarding the functional significance and global impact of this coupling on cellular and organismal homeostasis as well as its underlying molecular mechanisms. High-throughput transcriptome profiling techniques are beginning to fill the gaps in our knowledge of the genes that undergo alternative processing. We are learning when and where this coupling is happening, and the co-transcriptional transcriptome is being characterized. A further step of considerable importance consists of the integrated analysis of these datasets together with information about the genome-wide distribution of different forms of RNAPII, specific chromatin marks, and potential mediators of coupling. Due to its intrinsic potential as a mechanism for tissue-specific gene expression regulation, comparative analysis of co-transcriptional alternative splicing across distinct experimental models, such as panels of different cell lines, will also shed light on the global understanding of this exciting phenomenon.

Last, but not least, the spatial organization of the eukaryotic genome within the nucleus is an additional layer of regulation to be taken into account. Chromosomes are not randomly organized but occupy defined "territories" within the nucleus, and this distribution often results in the intermingling of specific subsets and combinations of genes that need to be co-regulated in a given situation (Branco and Pombo, 2006; Cavalli, 2007; Zhao et al., 2009; Sánchez-Álvarez et al., 2011). The functional coupling of transcription and pre-mRNA processing in the context of the highly compartmentalized eukaryotic nucleus still needs to be characterized. The recent purification and exhaustive analysis of the "transcription factory" (Melnik et al., 2011), the proposed structural unit for RNAPII transcription in the nucleus (Cook, 2010), will

substantially contribute to this goal. Notably, spliceosome components and associated factors were identified together with essential components of the RNAPII machinery as integral components of these higher-order structures (Melnik et al., 2011). It is also imperative to address the exact role of nuclear speckles in co- and post-transcriptional splicing. The recent observation showing the accumulation of intron-containing transcripts near these structures (Dias et al., 2010) has been strongly supported by single-molecule imaging studies, which revealed released unspliced pre-mRNA congregating around nuclear speckles (Vargas et al., 2011). This finding has important implications for our understanding of the mechanisms that regulate alternative splicing. This and other studies (Lionnet et al., 2010; Brody et al., 2011; Larson et al., 2011; Martins et al., 2011; Schmidt et al., 2011; Vargas et al., 2011; Gurskaya et al., 2012) that have used innovative microscopy and spectroscopy tools coupled with powerful statistical analyses and modeling allow us to monitor and quantify transcript biogenesis as well as co- and post-transcriptional processing at single-cell resolution *in vivo*. It is likely that, in the near future, numerous new insights and questions regarding the functional coordination of RNAPII transcriptional activity and pre-mRNA processing will continue to capture our attention.

Acknowledgments

This work was supported by grants from the Spanish Ministry of Science and Innovation (BFU2008-01599 and BFU2011-24577), by the Foundation for Research and Prevention of AIDS in Spain (FIPSE-36768/08), and by the Andalusian Government (Excellence Project CVI-4626/2009) to C.S. Support from the European

Region Development Fund, ERDF (FEDER) is also acknowledged. M.M. was supported by a fellowship from the Spanish Ministry of Education (FPU program).

References

Abovich, N. and Rosbash, M. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* **89** (1997), pp. 403-12.

Adami, G. and Babiss, L.E. DNA template effect on RNA splicing: two copies of the same gene in the same nucleus are processed differently. *EMBO J* **10** (1991), pp. 3457-65.

Adelman, K., Kennedy, M.A., Nechaev, S., Gilchrist, D.A., Muse, G.W., Chinenov, Y. and Rogatsky, I.: Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling, *Proc Natl Acad Sci U S A* (2009), pp. 18207-12.

Adelman, K. and Rogatsky, I. RNA polymerase II stalling mediates cytokine gene expression. *Cell Cycle* **9** (2010), pp. 630-1.

Aitken, S., Alexander, R.D. and Beggs, J.D. Modelling reveals kinetic advantages of co-transcriptional splicing. *PLoS Comput Biol* **7** (2011), p. e1002215.

Akgul, C., Moulding, D.A. and Edwards, S.W. Alternative splicing of Bcl-2-related genes: functional consequences and potential therapeutic applications. *Cell Mol Life Sci* **61** (2004), pp. 2189-99.

Alexander, R.D., Innocente, S.A., Barrass, J.D. and Beggs, J.D. Splicing-dependent RNA polymerase pausing in yeast. *Mol Cell* **40** (2010), pp. 582-93.

Allemand, E., Batsche, E. and Muchardt, C. Splicing, transcription, and chromatin: a menage a trois. *Curr Opin Genet Dev* **18** (2008), pp. 145-51.

Allen, M., Friedler, A., Schon, O. and Bycroft, M. The structure of an FF domain from human HYPA/FBP11. *J Mol Biol* **323** (2002), pp. 411-6.

Allo, M., Buggiano, V., Fededa, J.P., Petrillo, E., Schor, I., de la Mata, M., Agirre, E., Plass, M., Eyras, E., Elela, S.A., Klinck, R., Chabot, B. and Kornblihtt, A.R. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* **16** (2009), pp. 717-24.

Ameur, A., Zaghlool, A., Halvardson, J., Wetterbom, A., Gyllensten, U., Cavelier, L. and Feuk, L. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat Struct Mol Biol* **18** (2011), pp. 1435-40.

Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C. and Komorowski, J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res* **19** (2009), pp. 1732-41.

Andresen, J.M., Gayan, J., Cherny, S.S., Brocklebank, D., Alkorta-Aranburu, G., Addis, E.A., Cardon, L.R., Housman, D.E. and Wexler, N.S. Replication of twelve association studies for Huntington's disease residual age of onset in large Venezuelan kindreds. *J Med Genet* **44** (2007), pp. 44-50.

Arango, M., Holbert, S., Zala, D., Brouillet, E., Pearson, J., Regulier, E., Thakur, A.K., Aebischer, P., Wetzel, R., Deglon, N. and Neri, C. CA150 expression delays

striatal cell death in overexpression and knock-in conditions for mutant huntingtin neurotoxicity. *J Neurosci* **26** (2006), pp. 4649-59.

Bannister, A.J., Schneider, R. and Kouzarides, T. Histone methylation: dynamic or static? *Cell* **109** (2002), pp. 801-6.

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J. and Frey, B.J. Deciphering the splicing code. *Nature* **465** (2010), pp. 53-9.

Barboric, M., Lenasi, T., Chen, H., Johansen, E.B., Guo, S. and Peterlin, B.M. 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc Natl Acad Sci U S A* **106** (2009), pp. 7798-803.

Bartkowiak, B., Liu, P., Phatnani, H.P., Fuda, N.J., Cooper, J.J., Price, D.H., Adelman, K., Lis, J.T. and Greenleaf, A.L. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev* **24** (2010), pp. 2303-16.

Baskaran, R., Dahmus, M.E. and Wang, J.Y. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc Natl Acad Sci U S A* **90** (1993), pp. 11167-71.

Batsche, E., Yaniv, M. and Muchardt, C. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13** (2006), pp. 22-9.

Bauren, G. and Wieslander, L. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell* **76** (1994), pp. 183-92.

Beckmann, J.S. and Trifonov, E.N. Splice junctions follow a 205-base ladder. *Proc Natl Acad Sci U S A* **88** (1991), pp. 2380-3.

Bedford, M.T. and Leder, P. The FF domain: a novel motif that often accompanies WW domains. *Trends Biochem Sci* **24** (1999), pp. 264-5.

Bedford, M.T., Reed, R. and Leder, P. WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. *Proc Natl Acad Sci U S A* **95** (1998), pp. 10602-7.

Bentley, D.L. and Groudine, M. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. *Nature* **321** (1986), pp. 702-6.

Berglund, J.A., Abovich, N. and Rosbash, M. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev* **12** (1998), pp. 858-67.

Berglund, J.A., Chua, K., Abovich, N., Reed, R. and Rosbash, M. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89** (1997), pp. 781-7.

Beyer, A.L. and Osheim, Y.N. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2** (1988), pp. 754-65.

Bonnal, S., Martinez, C., Forch, P., Bachi, A., Wilm, M. and Valcarcel, J. RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. *Mol Cell* **32** (2008), pp. 81-95.

Bourquin, J.P., Stagljar, I., Meier, P., Moosmann, P., Silke, J., Baechi, T., Georgiev, O. and Schaffner, W. A serine/arginine-rich nuclear matrix cyclophilin interacts with the C-terminal domain of RNA polymerase II. *Nucleic Acids Res* **25** (1997), pp. 2055-61.

Branco, M.R. and Pombo, A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol* **4** (2006), p. e138.

Bres, V., Gomes, N., Pickle, L. and Jones, K.A. A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev* **19** (2005), pp. 1211-26.

Brody, Y., Neufeld, N., Bieberstein, N., Causse, S.Z., Bohnlein, E.M., Neugebauer, K.M., Darzacq, X. and Shav-Tal, Y. The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol* **9** (2011), p. e1000573.

Buratowski, S. The CTD code. *Nat Struct Biol* **10** (2003), pp. 679-80.

Buschdorf, J.P. and Stratling, W.H. A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *J Mol Med* **82** (2004), pp. 135-43.

Caceres, J.F. and Kornblihtt, A.R. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet* **18** (2002), pp. 186-93.

Carrillo Oesterreich, F., Preibisch, S. and Neugebauer, K.M. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol Cell* **40** (2010), pp. 571-81.

Carstens, R.P., Wagner, E.J. and Garcia-Blanco, M.A. An intronic splicing silencer causes skipping of the IIIb exon of fibroblast growth factor receptor 2 through involvement of polypyrimidine tract binding protein. *Mol Cell Biol* **20** (2000), pp. 7388-400.

Carty, S.M., Goldstrohm, A.C., Suñé, C., Garcia-Blanco, M.A. and Greenleaf, A.L. Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II. *Proc Natl Acad Sci U S A* **97** (2000), pp. 9015-20.

Cavalli, G. Chromosome kissing. *Curr Opin Genet Dev* **17** (2007), pp. 443-50.

Chan, D.C., M.T. Bedford, and P. Leder Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J.* **15** (1996), pp. 1045-1054.

Cheng, D., Cote, J., Shaaban, S. and Bedford, M.T. The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Mol Cell* **25** (2007), pp. 71-83.

Cook, P.R. A model for all genomes: the role of transcription factories. *J Mol Biol* **395** (2010), pp. 1-10.

Core, L.J. and Lis, J.T. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319** (2008), pp. 1791-2.

Coulter, D.E. and Greenleaf, A.L. Properties of mutationally altered RNA polymerases II of Drosophila. *J Biol Chem* **257** (1982), pp. 1945-52.

Cramer, P., Caceres, J.F., Cazalla, D., Kadener, S., Muro, A.F., Baralle, F.E. and Kornblihtt, A.R. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol Cell* **4** (1999), pp. 251-8.

Custodio, N., Carvalho, C., Condado, I., Antoniou, M., Blencowe, B.J. and Carmo-Fonseca, M. In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei. *RNA* **10** (2004), pp. 622-33.

Das, R., Dufu, K., Romney, B., Feldt, M., Elenko, M. and Reed, R. Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev* **20** (2006), pp. 1100-9.

Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P. and Reed, R. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell* **26** (2007), pp. 867-81.

de Almeida, S.F., Grosso, A.R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., Levezinho, H., Gut, M., Eick, D., Gut, I., Andrau, J.C., Ferrier, P. and Carmo-Fonseca, M. Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat Struct Mol Biol* **18** (2011), pp. 977-83.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A.R. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* **12** (2003), pp. 525-32.

de la Mata, M. and Kornblihtt, A.R. RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat Struct Mol Biol* **13** (2006), pp. 973-80.

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H. and Luhrmann, R. Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol Cell Biol* **26** (2006), pp. 5528-43.

Dhami, P., Saffrey, P., Bruce, A.W., Dillon, S.C., Chiang, K., Bonhoure, N., Koch, C.M., Bye, J., James, K., Foad, N.S., Ellis, P., Watkins, N.A., Ouwehand, W.H., Langford, C., Andrews, R.M., Dunham, I. and Vetrie, D. Complex exon-intron marking by histone modifications is not determined solely by nucleosome distribution. *PLoS One* **5** (2010), p. e12339.

Dias, A.P., Dufu, K., Lei, H. and Reed, R. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun* **1** (2010), p. 97.

Dong, B., Horowitz, D.S., Kobayashi, R. and Krainer, A.R. Purification and cDNA cloning of HeLa cell p54nrb, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and Drosophila NONA/BJ6. *Nucleic Acids Res* **21** (1993), pp. 4085-92.

Du, L., and S.L. Warren. A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. *J. Cell. Biol.* **136** (1997), pp. 5-18.

Dutertre, M., Sanchez, G., De Cian, M.C., Barbier, J., Dardenne, E., Gratadou, L., Dujardin, G., Le Jossic-Corcos, C., Corcos, L. and Auboeuf, D. Cotranscriptional exon skipping in the genotoxic stress response. *Nat Struct Mol Biol* **17** (2010), pp. 1358-66.

Egloff, S. and Murphy, S. Cracking the RNA polymerase II CTD code. *Trends Genet* **24** (2008), pp. 280-8.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe, B.J. and Ingles, C.J. Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD. *RNA* **8** (2002), pp. 1102-11.

Eperon, L.P., Graham, I.R., Griffiths, A.D. and Eperon, I.C. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* **54** (1988), pp. 393-401.

Faber, P.W., Barnes, G.T., Srinidhi, J., Chen, J., Gusella, J.F. and MacDonald, M.E. Huntingtin interacts with a family of WW domain proteins. *Hum Mol Genet* **7** (1998), pp. 1463-74.

Fong, N., Bird, G., Vigneron, M. and Bentley, D.L. A 10 residue motif at the C-terminus of the RNA pol II CTD is required for transcription, splicing and 3' end processing. *EMBO J* **22** (2003), pp. 4274-82.

Fong, Y.W. and Zhou, Q. Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414** (2001), pp. 929-33.

Fox, A.H., Bond, C.S. and Lamond, A.I. P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. *Mol Biol Cell* **16** (2005), pp. 5304-15.

Fox, A.H. and Lamond, A.I. Paraspeckles. *Cold Spring Harb Perspect Biol* **2** (2010), p. a000687.

Garcia-Blanco, M.A., Baraniak, A.P. and Lasda, E.L. Alternative splicing in disease and therapy. *Nat Biotechnol* **22** (2004), pp. 535-46.

Gasch, A., Wiesner, S., Martin-Malpartida, P., Ramirez-Espain, X., Ruiz, L. and Macias, M.J. The structure of Prp40 FF1 domain and its interaction with the crn-TPR1 motif of Clf1 gives a new insight into the binding mode of FF domains. *J Biol Chem* **281** (2006), pp. 356-64.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L. and Adelman, K. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143** (2010), pp. 540-51.

Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K., Collins, J.B., Li, L., Gilmour, D.S. and Adelman, K. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* **22** (2008), pp. 1921-33.

Glover-Cutter, K., Kim, S., Espinosa, J. and Bentley, D.L. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat Struct Mol Biol* **15** (2008), pp. 71-8.

Goldstrohm, A.C., Albrecht, T.R., Sune, C., Bedford, M.T. and Garcia-Blanco, M.A. The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. *Mol Cell Biol* **21** (2001a), pp. 7617-28.

Goldstrohm, A.C., Greenleaf, A.L. and Garcia-Blanco, M.A. Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing. *Gene* **277** (2001b), pp. 31-47.

Gornemann, J., Barrandon, C., Hujer, K., Rutz, B., Rigaut, G., Kotovic, K.M., Faux, C., Neugebauer, K.M. and Seraphin, B. Cotranscriptional spliceosome assembly and splicing are independent of the Prp40p WW domain. *RNA* **17** (2011), pp. 2119-29.

Gornemann, J., Kotovic, K.M., Hujer, K. and Neugebauer, K.M. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol Cell* **19** (2005), pp. 53-63.

Goto, H., Yasui, Y., Nigg, E.A. and Inagaki, M. Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells* **7** (2002), pp. 11-7.

Gunderson, F.Q. and Johnson, T.L. Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. *PLoS Genet* **5** (2009), p. e1000682.

Gurskaya, N.G., Staroverov, D.B., Zhang, L., Fradkov, A.F., Markina, N.M., Pereverzev, A.P. and Lukyanov, K.A. Analysis of alternative splicing of cassette exons at single-cell level using two fluorescent proteins. *Nucleic Acids Res* (2012).

Hartmann, B., Castelo, R., Blanchette, M., Boue, S., Rio, D.C. and Valcarcel, J. Global analysis of alternative splicing regulation by insulin and wingless signaling in Drosophila cells. *Genome Biol* **10** (2009), p. R11.

Hicks, M.J., Yang, C.R., Kotlajich, M.V. and Hertel, K.J. Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol* **4** (2006), p. e147.

Hirose, Y., R. Tacke, and J.L. Manley. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev.* **13** (1999), pp. 1234-1239.

Hnilicova, J., Hozeifi, S., Duskova, E., Icha, J., Tomankova, T. and Stanek, D. Histone deacetylase activity modulates alternative splicing. *PLoS One* **6** (2011), p. e16727.

Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M. and Bustamante, C. Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science* **325** (2009), pp. 626-8.

Holbert, S., I. Denghien, T. Kiechle, A. Rosenblatt, C. Wellington, M.R. Hayden, R.L. Margolis, C.A. Ross, J. Dausset, R.J. Ferrante, and C. Neri The Gln-Ala repeat transcriptional activator CA150 interacts with hungtingtin: Neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. *Proc. Natl. Acad. Sci. U S A* **98** (2001), pp. 1811-1816.

Hon, G., Wang, W. and Ren, B. Discovery and annotation of functional chromatin signatures in the human genome. *PLoS Comput Biol* **5** (2009), p. e1000566.

Hoskins, A.A., Friedman, L.J., Gallagher, S.S., Crawford, D.J., Anderson, E.G., Wombacher, R., Ramirez, N., Cornish, V.W., Gelles, J. and Moore, M.J. Ordered and dynamic assembly of single spliceosomes. *Science* **331** (2011), pp. 1289-95.

Hsin, J.P., Sheth, A. and Manley, J.L. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* **334** (2011), pp. 683-6.

Huang, X., Beullens, M., Zhang, J., Zhou, Y., Nicolaescu, E., Lesage, B., Hu, Q., Wu, J., Bollen, M. and Shi, Y. Structure and function of the two tandem WW domains of the pre-mRNA splicing factor FBP21 (formin-binding protein 21). *J Biol Chem* **284** (2009), pp. 25375-87.

Huff, J.T., Plocik, A.M., Guthrie, C. and Yamamoto, K.R. Reciprocal intronic and exonic histone modification regions in humans. *Nat Struct Mol Biol* **17** (2010), pp. 1495-9.

Huranova, M., Ivani, I., Benda, A., Poser, I., Brody, Y., Hof, M., Shav-Tal, Y., Neugebauer, K.M. and Stanek, D. The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *J Cell Biol* **191** (2010), pp. 75-86.

Ip, J.Y., Schmidt, D., Pan, Q., Ramani, A.K., Fraser, A.G., Odom, D.T. and Blencowe, B. Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res* (2010).

Izban, M.G. and Luse, D.S. Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. *Genes Dev* **5** (1991), pp. 683-96.

Izquierdo, J.M., Majos, N., Bonnal, S., Martinez, C., Castelo, R., Guigo, R., Bilbao, D. and Valcarcel, J. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol Cell* **19** (2005), pp. 475-84.

Izquierdo, J.M. and Valcarcel, J. A simple principle to explain the evolution of pre-mRNA splicing. *Genes Dev* **20** (2006), pp. 1679-84.

Jelen, N., Ule, J., Zivin, M. and Darnell, R.B. Evolution of Nova-dependent splicing regulation in the brain. *PLoS Genet* **3** (2007), pp. 1838-47.

Johnson, C., Primorac, D., McKinstry, M., McNeil, J., Rowe, D. and Lawrence, J.B. Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. *J Cell Biol* **150** (2000), pp. 417-32.

Kadener, S., Cramer, P., Nogues, G., Cazalla, D., de la Mata, M., Fededa, J.P., Werbajh, S.E., Srebrow, A. and Kornblihtt, A.R. Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing. *Embo J* **20** (2001), pp. 5759-68.

Kadener, S., Fededa, J.P., Rosbash, M. and Kornblihtt, A.R. Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. *Proc Natl Acad Sci U S A* **99** (2002), pp. 8185-90.

Kalsotra, A. and Cooper, T.A. Functional consequences of developmentally regulated alternative splicing. *Nat Rev Genet* **12** (2011), pp. 715-29.

Kalsotra, A., Xiao, X., Ward, A.J., Castle, J.C., Johnson, J.M., Burge, C.B. and Cooper, T.A. A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc Natl Acad Sci U S A* **105** (2008), pp. 20333-8.

Kameoka, S., P. Duque, and M.M. Konarska. p54(nrb) associates with the 5' splice site within large transcription/splicing complexes. *EMBO J.* **23** (2004), pp. 1782-1791.

Kao, H.Y. and Siliciano, P.G. Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **16** (1996), pp. 960-7.

Kao, S.Y., Calman, A.F., Luciw, P.A. and Peterlin, B.M. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* **330** (1987), pp. 489-93.

Katzenberger, R.J., Marengo, M.S. and Wassarman, D.A. ATM and ATR pathways signal alternative splicing of Drosophila TAF1 pre-mRNA in response to DNA damage. *Mol Cell Biol* **26** (2006), pp. 9256-67.

Keene, J.D. RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* **8** (2007), pp. 533-43.

Kelly, W.G., Dahmus, M.E. and Hart, G.W. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J Biol Chem* **268** (1993), pp. 10416-24.

Kim, H., Erickson, B., Luo, W., Seward, D., Graber, J.H., Pollock, D.D., Megee, P.C. and Bentley, D.L. Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat Struct Mol Biol* **17** (2010), pp. 1279-86.

Kim, S., Kim, H., Fong, N., Erickson, B. and Bentley, D.L. Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proc Natl Acad Sci U S A* **108** (2011), pp. 13564-9.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S. and Ahringer, J. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* **41** (2009), pp. 376-81.

Komarnitsky, P., E.J. Cho, and S. Buratowski. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14** (2000), pp. 2452-2460.

Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J. and Nogues, G. Multiple links between transcription and splicing. *Rna* **10** (2004), pp. 1489-98.

Kotovic, K.M., Lockshon, D., Boric, L. and Neugebauer, K.M. Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol Cell Biol* **23** (2003), pp. 5768-79.

Kouzarides, T. Chromatin modifications and their function. *Cell* **128** (2007), pp. 693-705.

Kramer, A. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem* **65** (1996), pp. 367-409.

Krumm, A., Meulia, T., Brunvand, M. and Groudine, M. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. *Genes Dev* **6** (1992), pp. 2201-13.

Lacadie, S.A. and Rosbash, M. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. *Mol Cell* **19** (2005), pp. 65-75.

Larson, D.R., Zenklusen, D., Wu, B., Chao, J.A. and Singer, R.H. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* **332** (2011), pp. 475-8.

Lee, J.S., Smith, E. and Shilatifard, A. The language of histone crosstalk. *Cell* **142** (2010), pp. 682-5.

Lenasi, T. and Barboric, M. P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. *RNA Biol* **7** (2010), pp. 145-50.

Levine, M. Paused RNA Polymerase II as a Developmental Checkpoint. *Cell* **145** (2011), pp. 502-11.

Li, B., Carey, M. and Workman, J.L. The role of chromatin during transcription. *Cell* **128** (2007a), pp. 707-19.

Li, H., Zhang, Z., Wang, B., Zhang, J., Zhao, Y. and Jin, Y. Wwp2-mediated ubiquitination of the RNA polymerase II large subunit in mouse embryonic pluripotent stem cells. *Mol Cell Biol* **27** (2007b), pp. 5296-305.

Licatalosi, D.D. and Darnell, R.B. Splicing regulation in neurologic disease. *Neuron* **52** (2006), pp. 93-101.

Lin, K.T., R.M. Lu, and W.Y. Tarn. The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo. *Mol. Cell. Biol.* **24** (2004), pp. 9176-9185.

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S. and Fu, X.D. The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* **15** (2008), pp. 819-26.

Lin, S. and Fu, X.D. SR proteins and related factors in alternative splicing. *Adv Exp Med Biol* **623** (2007), pp. 107-22.

Lionnet, T., Wu, B., Grunwald, D., Singer, R.H. and Larson, D.R. Nuclear physics: quantitative single-cell approaches to nuclear organization and gene expression. *Cold Spring Harb Symp Quant Biol* **75** (2010), pp. 113-26.

Listerman, I., Sapra, A.K. and Neugebauer, K.M. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol* **13** (2006), pp. 815-22.

Loomis, R.J., Naoe, Y., Parker, J.B., Savic, V., Bozovsky, M.R., Macfarlan, T., Manley, J.L. and Chakravarti, D. Chromatin binding of SRp20 and ASF/SF2 and dissociation from mitotic chromosomes is modulated by histone H3 serine 10 phosphorylation. *Mol Cell* **33** (2009), pp. 450-61.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli, T. Regulation of alternative splicing by histone modifications. *Science* **327** (2010), pp. 996-1000.

MacMillan, A.M., Query, C.C., Allerson, C.R., Chen, S., Verdine, G.L. and Sharp, P.A. Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev* **8** (1994), pp. 3008-20.

Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M. and Luhrmann, R. Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* **298** (2002), pp. 2205-8.

Maniatis, T. and Reed, R. An extensive network of coupling among gene expression machines. *Nature* **416** (2002), pp. 499-506.

Maniatis, T. and Tasic, B. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* **418** (2002), pp. 236-43.

Martinez, E., Palhan, V.B., Tjernberg, A., Lymar, E.S., Gamper, A.M., Kundu, T.K., Chait, B.T. and Roeder, R.G. Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo. *Mol Cell Biol* **21** (2001), pp. 6782-95.

Martins, S.B., Rino, J., Carvalho, T., Carvalho, C., Yoshida, M., Klose, J.M., de Almeida, S.F. and Carmo-Fonseca, M. Spliceosome assembly is coupled to RNA polymerase II dynamics at the 3' end of human genes. *Nat Struct Mol Biol* **18** (2011), pp. 1115-23.

Matlin, A.J. and Moore, M.J. Spliceosome assembly and composition. *Adv Exp Med Biol* **623** (2007), pp. 14-35.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S.P. and Elledge, S.J. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316** (2007), pp. 1160-6.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385** (1997), pp. 357-61.

Melnik, S., Deng, B., Papantonis, A., Baboo, S., Carr, I.M. and Cook, P.R. The proteomes of transcription factories containing RNA polymerases I, II or III. *Nat Methods* **8** (2011), pp. 963-8.

Mendes Soares, L.M. and Valcarcel, J. The expanding transcriptome: the genome as the 'Book of Sand'. *EMBO J* **25** (2006), pp. 923-31.

Misteli, T., and D.L. Spector. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell.* **3** (1999), pp. 697-705.

Mittal, N., Roy, N., Babu, M.M. and Janga, S.C. Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. *Proc Natl Acad Sci U S A* **106** (2009), pp. 20300-5.

Montes, M., Cloutier, A., Sanchez-Hernandez, N., Michelle, L., Lemieux, B., Blanchette, M., Hernandez-Munain, C., Chabot, B. and Sune, C. TCERG1 regulates alternative splicing of Bcl-x gene by modulating the rate of RNAPII transcription. *Mol Cell Biol* **32** (2012), pp. 751-762.

Moore, M.J. and Proudfoot, N.J. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136** (2009), pp. 688-700.

Moore, M.J., Wang, Q., Kennedy, C.J. and Silver, P.A. An alternative splicing network links cell-cycle control to apoptosis. *Cell* **142** (2010), pp. 625-36.

Morris, D.P. and Greenleaf, A.L. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **275** (2000), pp. 39935-43.

Morris, D.P., Phatnani, H.P. and Greenleaf, A.L. Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-End formation. *J Biol Chem* **274** (1999), pp. 31583-7.

Moteki, S. and Price, D. Functional coupling of capping and transcription of mRNA. *Mol Cell* **10** (2002), pp. 599-609.

Munoz, M.J., de la Mata, M. and Kornblihtt, A.R. The carboxy terminal domain of RNA polymerase II and alternative splicing. *Trends Biochem Sci* **35** (2010), pp. 497-504.

Munoz, M.J., Perez Santangelo, M.S., Paronetto, M.P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J.J., Bird, G., Bentley, D., Bertrand, E. and Kornblihtt, A.R. DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. *Cell* **137** (2009), pp. 708-20.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J. and Adelman, K. RNA polymerase is poised for activation across the genome. *Nat Genet* **39** (2007), pp. 1507-11.

Natalizio, B.J. and Garcia-Blanco, M.A. In vitro coupled transcription splicing. *Methods* **37** (2005), pp. 314-22.

Natalizio, B.J., Robson-Dixon, N.D. and Garcia-Blanco, M.A. The Carboxyl-terminal Domain of RNA Polymerase II Is Not Sufficient to Enhance the Efficiency of Pre-mRNA Capping or Splicing in the Context of a Different Polymerase. *J Biol Chem* **284** (2009), pp. 8692-702.

Nechaev, S. and Adelman, K. Promoter-proximal Pol II: when stalling speeds things up. *Cell Cycle* **7** (2008), pp. 1539-44.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y. and Adelman, K. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. *Science* **327** (2010), pp. 335-8.

Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A. and Mann, M. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat Genet* **20** (1998), pp. 46-50.

Nogues, G., Kadener, S., Cramer, P., Bentley, D. and Kornblihtt, A.R. Transcriptional activators differ in their abilities to control alternative splicing. *J Biol Chem* **277** (2002), pp. 43110-4.

O'Brien, T. and Lis, J.T. RNA polymerase II pauses at the 5' end of the transcriptionally induced Drosophila hsp70 gene. *Mol Cell Biol* **11** (1991), pp. 5285-90.

Oesterreich, F.C., Bieberstein, N. and Neugebauer, K.M. Pause locally, splice globally. *Trends Cell Biol* **21** (2011), pp. 328-35.

Ohkura, N., Takahashi, M., Yaguchi, H., Nagamura, Y. and Tsukada, T. Coactivator-associated arginine methyltransferase 1, CARM1, affects pre-mRNA splicing in an isoform-specific manner. *J Biol Chem* **280** (2005), pp. 28927-35.

Orphanides, G. and Reinberg, D. RNA polymerase II elongation through chromatin. *Nature* **407** (2000), pp. 471-5.

Osheim, Y.N., Miller, O.L., Jr. and Beyer, A.L. Visualization of Drosophila melanogaster chorion genes undergoing amplification. *Mol Cell Biol* **8** (1988), pp. 2811-21.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40** (2008), pp. 1413-5.

Pandya-Jones, A. and Black, D.L. Co-transcriptional splicing of constitutive and alternative exons. *RNA* **15** (2009), pp. 1896-908.

Paronetto, M.P., Minana, B. and Valcarcel, J. The Ewing sarcoma protein regulates DNA damage-induced alternative splicing. *Mol Cell* **43** (2011), pp. 353-68.

Passani, L.A., Bedford, M.T., Faber, P.W., McGinnis, K.M., Sharp, A.H., Gusella, J.F., Vonsattel, J.P. and MacDonald, M.E. Huntingtin's WW domain partners in

Huntington's disease post-mortem brain fulfill genetic criteria for direct involvement in Huntington's disease pathogenesis. *Hum Mol Genet* **9** (2000), pp. 2175-82.

Pearson, J.L., Robinson, T.J., Munoz, M.J., Kornblihtt, A.R. and Garcia-Blanco, M.A. Identification of the cellular targets of the transcription factor TCERG1 reveals a prevalent role in mRNA processing. *J Biol Chem* **283** (2008), pp. 7949-61.

Peckham, H.E., Thurman, R.E., Fu, Y., Stamatoyannopoulos, J.A., Noble, W.S., Struhl, K. and Weng, Z. Nucleosome positioning signals in genomic DNA. *Genome Res* **17** (2007), pp. 1170-7.

Peng, R., Dye, B.T., Perez, I., Barnard, D.C., Thompson, A.B. and Patton, J.G. PSF and p54nrb bind a conserved stem in U5 snRNA. *RNA* **8** (2002), pp. 1334-47.

Peng, R., Hawkins, I., Link, A.J. and Patton, J.G. The splicing factor PSF is part of a large complex that assembles in the absence of pre-mRNA and contains all five snRNPs. *RNA Biol* **3** (2006), pp. 69-76.

Perales, R. and Bentley, D. "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* **36** (2009), pp. 178-91.

Phatnani, H.P. and Greenleaf, A.L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20** (2006), pp. 2922-36.

Rappsilber, J., Ryder, U., Lamond, A.I. and Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res* **12** (2002), pp. 1231-45.

Roberts, G.C., C. Gooding, H.Y. Mak, N.J.. Proudfoot, and C.W. Smith. Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res.* **26** (1998), pp. 5568-5572.

Rosonina, E. and Blencowe, B.J. Analysis of the requirement for RNA polymerase II CTD heptapeptide repeats in pre-mRNA splicing and 3'-end cleavage. *RNA* **10** (2004), pp. 581-9.

Rosonina, E., Ip, J.Y., Calarco, J.A., Bakowski, M.A., Emili, A., McCracken, S., Tucker, P., Ingles, C.J. and Blencowe, B.J. Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Mol Cell Biol* **25** (2005), pp. 6734-46.

Saint-Andre, V., Batsche, E., Rachez, C. and Muchardt, C. Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat Struct Mol Biol* **18** (2011), pp. 337-44.

Saltzman, A.L., Kim, Y.K., Pan, Q., Fagnani, M.M., Maquat, L.E. and Blencowe, B.J. Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol Cell Biol* **28** (2008), pp. 4320-30.

Sanchez-Alvarez, M., Goldstrohm, A.C., Garcia-Blanco, M.A. and Sune, C. Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through its amino and carboxyl regions. *Mol Cell Biol* **26** (2006), pp. 4998-5014.

Sánchez-Álvarez, M., Montes, M., Sánchez-Hernández, N., Hernández-Munain, C. and Suñé, C. Differential effects of sumoylation on transcription and alternative splicing by transcription elongation regulator 1 (TCERG1). *J Biol Chem* **285** (2010), pp. 15220-33.

Sánchez-Álvarez, M., Sanchez-Hernández, N. and Suñé, C.: Spatial Organization and Dynamics of Transcription Elongation and Pre-mRNA Processing in Live Cells, *Genetics Research International* (2011).

Schmidt, U., Basyuk, E., Robert, M.C., Yoshida, M., Villemin, J.P., Auboeuf, D., Aitken, S. and Bertrand, E. Real-time imaging of cotranscriptional splicing

reveals a kinetic model that reduces noise: implications for alternative splicing regulation. *J Cell Biol* **193** (2011), pp. 819-29.

Schor, I.E., Rascovan, N., Pelisch, F., Allo, M. and Kornblihtt, A.R. Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci U S A* **106** (2009), pp. 4325-30.

Schroeder, S.C., Schwer, B., Shuman, S. and Bentley, D. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev* **14** (2000), pp. 2435-40.

Schwartz, S., Meshorer, E. and Ast, G. Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol* **16** (2009), pp. 990-5.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R. and Oberdoerffer, S. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* (2011).

Sims, R.J., 3rd, Belotserkovskaya, R. and Reinberg, D. Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18** (2004), pp. 2437-68.

Sims, R.J., 3rd, Millhouse, S., Chen, C.F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L. and Reinberg, D. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* **28** (2007), pp. 665-76.

Sims, R.J., 3rd, Rojas, L.A., Beck, D., Bonasio, R., Schuller, R., Drury, W.J., 3rd, Eick, D. and Reinberg, D. The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* **332** (2011), pp. 99-103.

Singh, J. and Padgett, R.A. Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol* **16** (2009), pp. 1128-33.

Smith, M.J., Kulkarni, S. and Pawson, T. FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. *Mol Cell Biol* **24** (2004), pp. 9274-85.

Spies, N., Nielsen, C.B., Padgett, R.A. and Burge, C.B. Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* **36** (2009), pp. 245-54.

Sugiyama, K., Sugiura, K., Hara, T., Sugimoto, K., Shima, H., Honda, K., Furukawa, K., Yamashita, S. and Urano, T. Aurora-B associated protein phosphatases as negative regulators of kinase activation. *Oncogene* **21** (2002), pp. 3103-11.

Suñé, C., Hayashi, T., Liu, Y., Lane, W.S., Young, R.A. and Garcia-Blanco, M.A. CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. *Mol Cell Biol* **17** (1997), pp. 6029-39.

Taliaferro, J.M., Alvarez, N., Green, R.E., Blanchette, M. and Rio, D.C. Evolution of a tissue-specific splicing network. *Genes Dev* **25** (2011), pp. 608-20.

Tietjen, J.R., Zhang, D.W., Rodriguez-Molina, J.B., White, B.E., Akhtar, M.S., Heidemann, M., Li, X., Chapman, R.D., Shokat, K., Keles, S., Eick, D. and Ansari, A.Z. Chemical-genomic dissection of the CTD code. *Nat Struct Mol Biol* **17** (2010), pp. 1154-61.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcarcel, J. and Guigo, R. Nucleosome positioning as a determinant of exon recognition. *Nat Struct Mol Biol* **16** (2009), pp. 996-1001.

Tyagi, A., Ryme, J., Brodin, D., Ostlund Farrants, A.K. and Visa, N. SWI/SNF associates with nascent pre-mRNPs and regulates alternative pre-mRNA processing. *PLoS Genet* **5** (2009), p. e1000470.

Ule, J. and Darnell, R.B. Functional and mechanistic insights from genome-wide studies of splicing regulation in the brain. *Adv Exp Med Biol* **623** (2007), pp. 148-60.

Vargas, D.Y., Shah, K., Batish, M., Levandoski, M., Sinha, S., Marras, S.A., Schedl, P. and Tyagi, S. Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell* **147** (2011), pp. 1054-65.

Wahl, M.C., Will, C.L. and Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136** (2009), pp. 701-18.

Waldholm, J., Wang, Z., Brodin, D., Tyagi, A., Yu, S., Theopold, U., Farrants, A.K. and Visa, N. SWI/SNF regulates the alternative processing of a specific subset of pre-mRNAs in Drosophila melanogaster. *BMC Mol Biol* **12** (2011), p. 46.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456** (2008), pp. 470-6.

Wang, G.S. and Cooper, T.A. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* **8** (2007), pp. 749-61.

Will, C.L. and Luhrmann, R. Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13** (2001), pp. 290-301.

Wu, J.Y. and Maniatis, T. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75** (1993), pp. 1061-70.

Yang, X.J. and Seto, E. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* **26** (2007), pp. 5310-8.

Yu, Y., Das, R., Folco, E.G. and Reed, R. A model in vitro system for co-transcriptional splicing. *Nucleic Acids Res* **38** (2010), pp. 7570-8.

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M. and Corden, J.L. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc Natl Acad Sci U S A* **93** (1996), pp. 6975-80.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M. and Young, R.A. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat Genet* **39** (2007), pp. 1512-6.

Zeng, C. and Berget, S.M. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol Cell Biol* **20** (2000), pp. 8290-301.

Zhang, C., Frias, M.A., Mele, A., Ruggiu, M., Eom, T., Marney, C.B., Wang, H., Licatalosi, D.D., Fak, J.J. and Darnell, R.B. Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* **329**, pp. 439-43.

Zhang, J. and Corden, J.L. Identification of phosphorylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J Biol Chem* **266** (1991), pp. 2290-6.

Zhao, R., Bodnar, M.S. and Spector, D.L. Nuclear neighborhoods and gene expression. *Curr Opin Genet Dev* **19** (2009), pp. 172-9.

Zhou, M., Huang, K., Jung, K.J., Cho, W.K., Klase, Z., Kashanchi, F., Pise-Masison, C.A. and Brady, J.N. Bromodomain protein Brd4 regulates human immunodeficiency virus transcription through phosphorylation of CDK9 at threonine 29. *J Virol* **83** (2009), pp. 1036-44.

Figure Legends

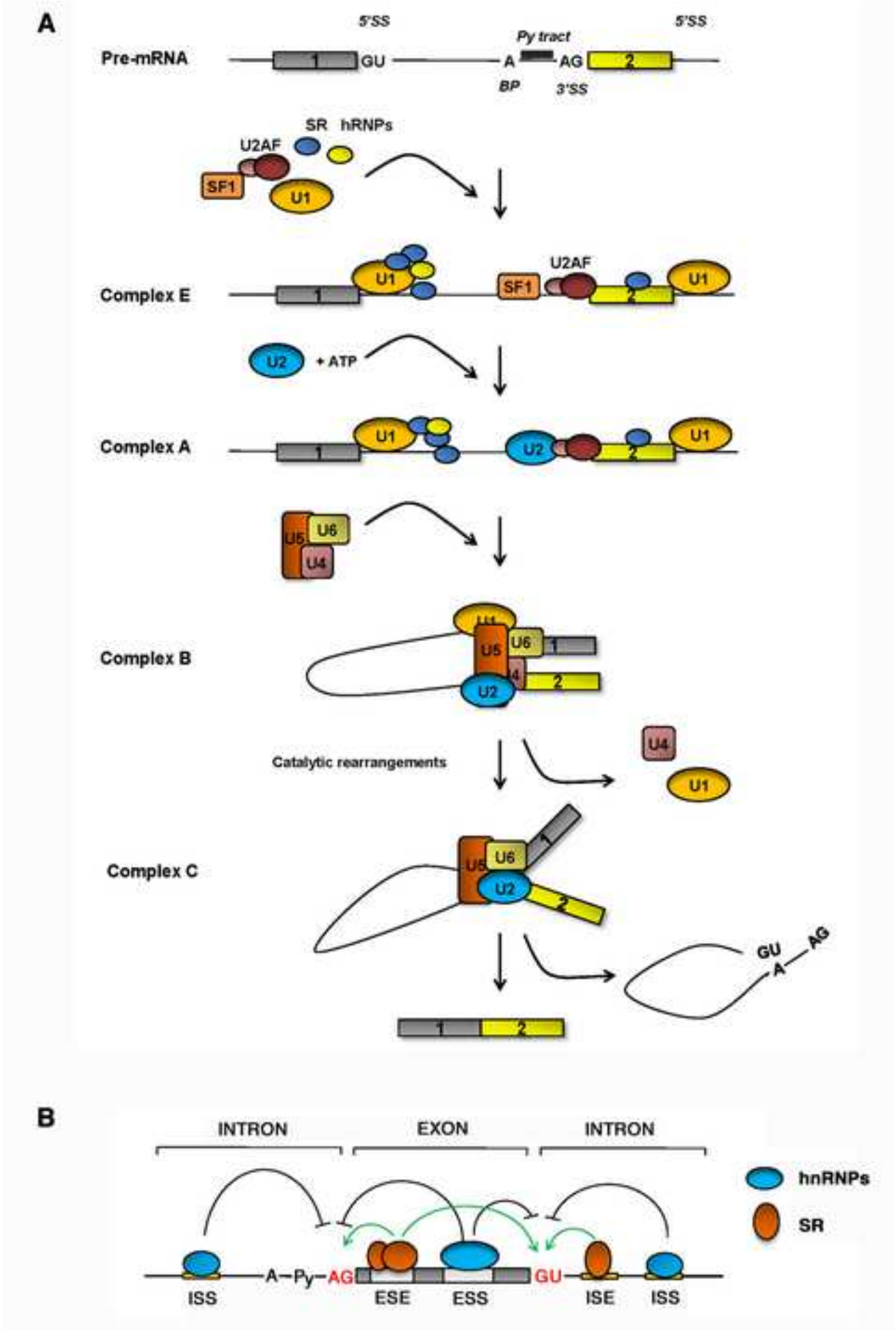
Fig. 1. A. Schematic representation of spliceosome assembly and the splicing of a pre-mRNA. The spliceosome is composed of a core of five small ribonucleoprotein particles (snRNPs), U1, U2, U4, U5, and U6, and ~ 200 additional proteins. The first step in spliceosome assembly is the formation of complex E (the commitment complex). The 5' splice site (GU, 5'ss) is bound by the U1 snRNP, and the splicing factors SF1 and U2AF cooperatively recognize the branch point sequence (BP), the Py, and the 3' splice site (AG, 3'ss). In an ATP-dependent manner, the pairing of the U2snRNP with the branchpoint results in the pre-spliceosomal complex A. Subsequent steps lead to the binding of the U4–U5–U6 tri-snRNP and the formation of the complex B. The catalytic complex C, which performs two trans-esterification reactions at the splice sites, is formed after rearrangements that detach the U1 and U4 snRNPs. These reactions result in the ligation of the exons and the excision of the intron, which is removed as the lariat RNA. B. Alternative splicing regulation by *cis*-elements and *trans*-acting factors. The core *cis* sequence elements that define the exon/intron boundaries (5' and 3' splice sites, GU-AG in red) and associated 3' sequences (polypyrimidine tract, Py; and branch point sequence, A) are poorly conserved. Additional enhancer and silencer elements in exons and in introns (ESE/ESS: exonic splicing enhancers/silencers, ISE/ISS: intronic splicing enhancers/silencers) add another layer of alternative splicing regulation. *Trans*-acting splicing factors, e.g., serine/arginine-rich (SR) family proteins and heterogeneous nuclear ribonucleoprotein particles (hnRNPs), bind to enhancers and silencers and interact with spliceosomal components. In general, SR proteins bound to enhancers facilitate exon recognition (green arrows) but hnRNPs inhibit this process (black arrows).

Fig. 2. The CTD-code and pre-mRNA processing. A. Hypophosphorylated RNAPII (IIa) is recruited to the preinitiation complex (PIC) at the promoter. B. Transcription is initiated upon the phosphorylation of the serines at position 5 (Ser5-P, green) and position 7 (Ser7-P, red) of the CTD by CDK7, the kinase component of the TFHII factor. Ser5-P helps to recruit the machinery necessary for the capping of the 5' end of the transcript (pink). C. During active elongation, the CTD is phosphorylated at Ser2 (Ser2-P, blue) by CDK9 and CDK12/13 (see text for more details) to produce the hyperphosphorylated form of RNAPII (IIo). Numerous splicing factors are recruited to the hyperphosphorylated RNAPII (orange). D. During elongation towards the 3' end of the gene, Ser5 phosphorylation decreases by the action of Ser5-specific phosphatases. Polyadenylation and cleavage factors (blue) bind to the CTD to allow transcription termination and the release of the mature transcript. E. After dissociation from the DNA template, Ser2-specific phosphatases remove Ser5-P before recycling or reinitiation.

Fig. 3. The role of chromatin in alternative splicing. A. Changes in chromatin organization due to the presence of chromatin remodeling factors (blue ovals), histone tail modifications (stars), or nucleosome positioning can affect the rate of RNAPII elongation and in turn affect the alternative splicing outcome. Slow elongation allows the binding of splicing factors to the nascent pre-mRNA and the recognition of weak splices sites, favoring exon inclusion (yellow box), while fast elongation favors exon exclusion (red box). B. Histone marks can also influence splicing decisions independently of the rate of RNAPII transcription by recruiting splicing factors via interaction with chromatin adaptors.

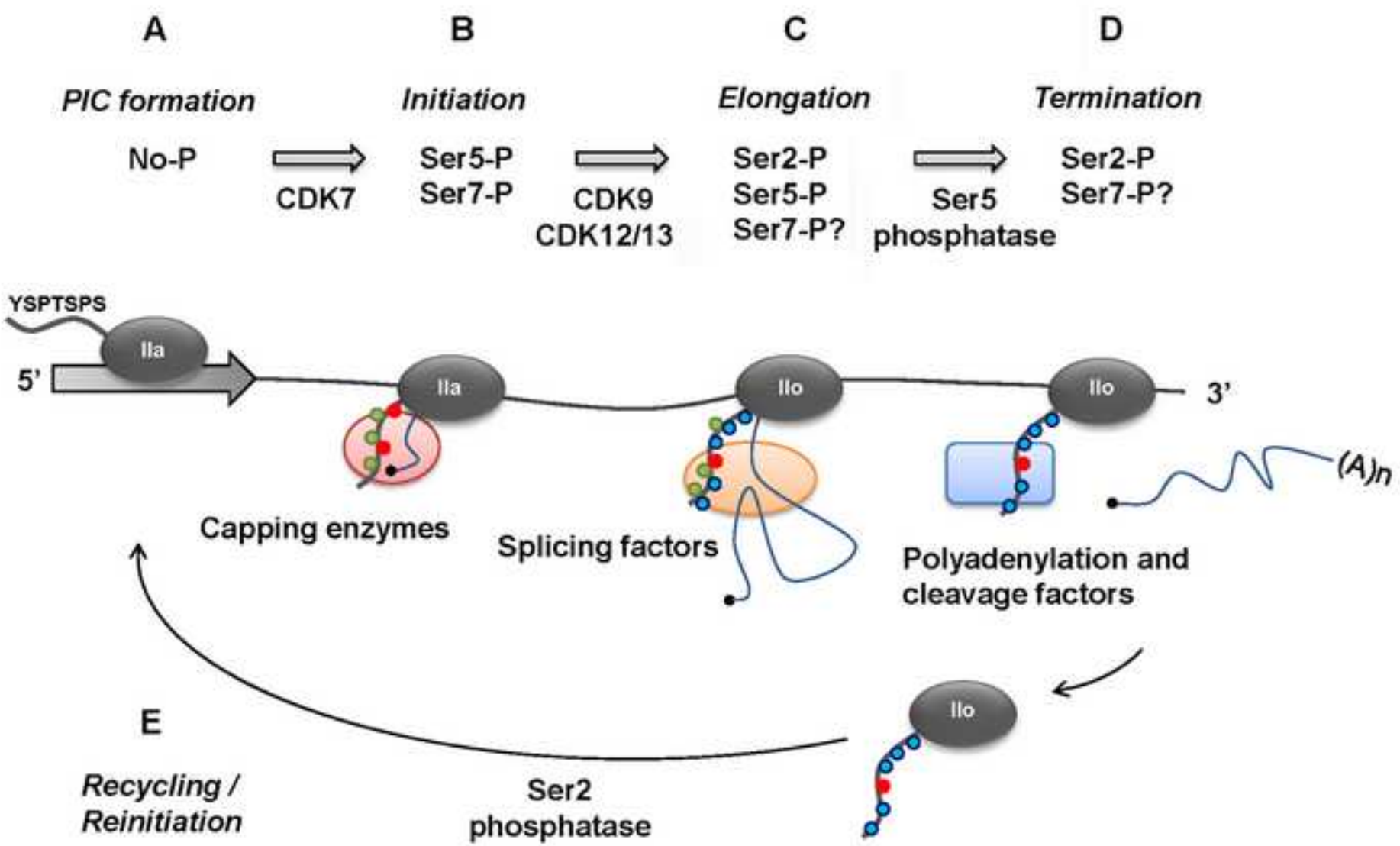
Figure(s)

Click here to download high resolution image



Figure(s)

Click here to download high resolution image



Figure(s)

Click here to download high resolution image