The EMBO Journal (2010), 1–8 | © 2010 European Molecular Biology Organization | All Rights Reserved 0261-4189/10
www.embojournal.org

THE EMBO JOURNAL

EMBO Member's Review

Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing

Schraga Schwartz and Gil Ast*

Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel-Aviv University, Ramat Aviv, Israel

How are short exonic sequences recognized within the vast intronic oceans in which they reside? Despite decades of research, this remains one of the most fundamental, yet enigmatic, questions in the field of pre-mRNA splicing research. For many years, studies aiming to shed light on this process were focused at the RNA level, characterizing the manner by which splicing factors and auxiliary proteins interact with splicing signals, thereby enabling, facilitating and regulating splicing. However, we increasingly understand that splicing is not an isolated process; rather it occurs co-transcriptionally and is presumably also regulated by transcription-related processes. In fact, studies by our group and others over the past year suggest that DNA structure in terms of nucleosome positioning and specific histone modifications, which have a well established role in transcription, may also have a role in splicing. In this review we discuss evidence for the coupling between transcription and splicing, focusing on recent findings suggesting a link between chromatin structure and splicing, and highlighting challenges this emerging field is facing.

The EMBO Journal advance online publication, 20 April 2010; doi:10.1038/emboj.2010.71
Subject Categories: chromatin & transcription; RNA
Keywords: bioinformatics; chromatin; exons; nucleosomes; splicing

Mammalian RNA is composed of short stretches of exonic sequences, typically around ~140 nt in length, interrupted by long, non-coding introns that are typically thousands of nucleotides long. In order for RNA polymerase-II (RNAPII) transcripts to mature into mRNA molecules, a sophisticated machinery known as the spliceosome excises the introns and ligates exons in a process termed splicing. One of the major and most fundamental challenges facing the splicing machinery is to correctly identify and ligate the short exons and precisely remove the long intronic sequences. This challenge is compounded by the fact that the splicing machinery may recognize a given sequence as an exon in one tissue or disease state, but as an intron in another, in a process termed alternative splicing (Smith and Valcarcel, 2000; Black, 2003; Nilsen and Graveley, 2010). Classically, attempts to understand the mechanism and regulation of splicing have focused on the pre-mRNA molecules and researchers have sought to characterize the *cis* and *trans* factors that facilitate and regulate exon recognition. Four signals residing at the exon–intron boundaries have a well-characterized role in helping direct the splicing machinery: The 5′ and 3′ splice sites (5′ss and 3′ss), at the respective ends of the introns, the polypyrimidine tract (PPT), which is upstream from the 3′ss, and the branch site (BS), located upstream from the PPT. These signals, along with a vast array of splicing-regulatory elements (SREs) in the exons and introns, direct the spliceosomal machinery to the exon–intron boundaries and allow precise recognition of exons (Berget, 1995; Smith and Valcarcel, 2000; Graveley, 2001; Black, 2003; Nilsen and Graveley, 2010). However, these signals are far from containing sufficient information to allow precise differentiation between exonic and intronic sequences (Lim and Burge, 2001; Schwartz *et al*, 2009a), implying that there are additional regulatory layers providing additional information, remaining to be discovered. Another open puzzle is the mechanism underlying tissue-specific patterns of alternative splicing, despite the identity in terms of the underlying sequence. In some cases this specificity can be explained through tissue-specific expression of splicing factors. However, in other cases the involved splicing factors are ubiquitously expressed (see references Kornblihtt *et al*, 2004 and Wang and Cooper, 2007), again indicating that there are additional regulatory layers leading to tissue-specific regulation of splicing patterns.

Where is this additional layer of information? One attractive possibility is that it lies therein that splicing *in vivo* does not occur as an isolated process; rather, it occurs co-transcriptionally. Although transcription-independent splicing has been documented both *in vitro* and *in vivo* (Green *et al*, 1983; Wetterberg *et al*, 1996), and it was found that in yeast the majority of splicing occurs post-transcriptionally (Tardiff *et al*, 2006), the emerging consensus is that splicing is initiated co-transcriptionally and that introns are removed while the nascent transcript is still tethered to the DNA by RNAPII (Neugebauer, 2002; Proudfoot *et al*, 2002; Allemand *et al*, 2008; Pandya-Jones and Black, 2009; Singh and Padgett, 2009; Wada *et al*, 2009). This is based on observations indicative of co-transcriptional splicing in human (Wuarin and Schibler, 1994; Tennyson *et al*, 1995; Roberts *et al*, 1998; Dye *et al*, 2006; Listerman *et al*, 2006; Pandya-Jones and Black, 2009), in *Drosophila* (Beyer and Osheim, 1988; LeMaire and Thummel, 1990), and in other systems and organisms (for review see references Neugebauer, 2002;

*Corresponding author. Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel-Aviv University, Medical School, Ramat Aviv 69978, Israel. Tel.: +972 3 640 6893; Fax: +972 3 640 5168; E-mail: gilast@post.tau.ac.il

Received: 4 February 2010; accepted: 26 March 2010

Allemand *et al*, 2008; Perales and Bentley, 2009). In itself, the temporal overlap between transcription and splicing is of no functional consequence in terms of guiding the splicing machinery towards the exon boundaries. However, it sets the foundation for coupling, or crosstalk, between these two processes, thereby providing possibilities for splicing to be controlled and regulated by transcription.

Different lines of evidence suggest that such cross-talk exists. First, there is evidence for coordinated regulation of transcription and splicing (Auboeuf *et al*, 2002). Furthermore, there are a large number of transcriptional regulators and elongation factors with documented roles in splicing (Allemand *et al*, 2008), and similarly there is an increasing number of splicing factors with documented roles in transcription (e.g., Bres *et al*, 2005; Lin *et al*, 2008). However, the most dominant factor implicated in this cross-talk between transcription and splicing is RNAPII as indicated by the finding that that in the absence of intact RNAPII, synthetic pre-mRNAs is spliced less efficiently (Bird *et al*, 2004; Das *et al*, 2006; Hicks *et al*, 2006).

How does RNAPII facilitate splicing? Evidence exists for at least two mechanisms. The first mechanism by which RNAPII affects splicing is by serving as a 'piggyback' or 'landing pad' for various splicing factors. In particular, the C-terminal domain (CTD) of RNAPII recruits a wide range of proteins, including splicing factors, to the nascent transcript (McCracken *et al*, 1997; Goldstrohm *et al*, 2001; Howe, 2002; Fong *et al*, 2003; de Almeida and Carmo-Fonseca, 2008). For example, human U1 snRNP components co-immunoprecipitate with RNAPII (Das *et al*, 2007) and presence of U1 snRNP at a 5′ss can promote recruitment of RNAPII and general transcription factors to the promoter (Damgaard *et al*, 2008). Also SR proteins, which help mediate spliceosomal assembly by binding to the U1 and U2 components of the spliceosome (Ram and Ast, 2007; Long and Caceres, 2009), co-immunoprecipitate with RNAPII (Yuryev *et al*, 1996; Das *et al*, 2007) and SRp20 regulation is dependent on the presence of the CTD in transcribing RNAPII (de la Mata and Kornblihtt, 2006). The second mechanism by which RNAPII is thought to regulate splicing is through kinetic coupling. According to this model, decreased RNAPII elongation rates facilitate the recognition of weakly defined exons that would otherwise not be recognized by the splicing machinery. There are different sources of support for this model: First, a mutant RNAPII with lower transcription rates than the wild-type polymerase increases the inclusion levels of certain alternatively spliced exons (de la Mata *et al*, 2003; Kornblihtt, 2006; Schor *et al*, 2009). Second, the topoisomerase-I inhibitor camptothecin, which stalls elongating RNAPII, increases co-transcriptional splicing factor accumulation and splicing (Listerman *et al*, 2006). Third, factors that increase the rate of RNAPII elongation, such as specific splicing enhancers or factors acting in *trans*, lead to increased exon skipping, whereas the elongation inhibitor dichlororibofuranosylbenzimidazole (DRB) enhances exon inclusion (Nogues *et al*, 2002; Kornblihtt *et al*, 2004). Finally, upon UV-induced damage, the CTD becomes hyper-phosphorylated leading to slowing down of transcription and subsequent alterations in splicing patterns (Munoz *et al*, 2009).

The coupling of splicing with transcription raised the possibility that factors classically known to regulate transcription could also have a role in the regulation of splicing. One such factor is chromatin structure. By chromatin structure we refer to two main elements: nucleosome positioning and histone modifications. Approximately 147 base pairs of DNA are wrapped around protein octamers, termed nucleosomes, that are separated by linker regions of variable length (Kouzarides, 2007). Positioning of nucleosomes along the genome is determined, to a large extent, by the genome sequence (Ioshikhes *et al*, 2006; Segal *et al*, 2006; Kaplan *et al*, 2009), but can also be subjected to modulation by chromatin remodellers (Vignali *et al*, 2000). In turn, nucleosome positioning modulates the properties of the sequence coiled around the nucleosome. For example, promoter DNA sequences covered by nucleosomes are less accessible to transcription factors (Kornberg and Lorch, 1999; Cairns, 2009). The chemical properties of nucleosomes can be subjected to fine-tuning, through post-translational histone modifications. The N-terminal tails of histones are subjected to at least eight types of modifications, including acetylation, methylation, phosphorylation, and ubiquitination; these modifications alter the chromatin structure and accessibility, and subsequently gene expression levels (Bernstein *et al*, 2007; Jones and Baylin, 2007; Kouzarides, 2007).

Considerable evidence has accumulated over the past years in support of a link between nucleosome organization and splicing. One of the first studies suggesting such a link was based on changes in splicing patterns after perturbation with drugs that inhibit histone acetylases (Nogues *et al*, 2002). Additional support for such a link came from the fact that different chromatin remodellers have a role in splicing. For example, the Brahma subunit of the chromatin-remodelling complex SWI/SNF interacts with splicing factors and regulates alternative splicing, presumably by pausing RNAPII elongation and thereby leading to inclusion of otherwise skipped exons (Batsche *et al*, 2006). Also, one isoform of the coactivator-associated arginine methyltransferase-1 associates with the U1 component U1C and affects 5′ ss selection (Ohkura *et al*, 2005). In addition, the CHD1 chromatin-remodelling ATPase has a role in splicing (Sims *et al*, 2007) and conversely the splicing factors SRp20 and ASF/SF2 are associated with chromatin (Loomis *et al*, 2009). Moreover, it was recently found that Gcn5, which encodes the histone acetyltransferase (HAT) activity of the SAGA complex, affects splicing (Gunderson and Johnson, 2009). Is there, then, a direct cross-talk between chromatin structure, exon–intron architecture, and splicing?

Link between nucleosome positioning and splicing

Studies of several groups, including ours, published over the last year have found large-scale evidence for a link between nucleosome positioning and exon–intron architecture (Andersson *et al*, 2009; Hon *et al*, 2009; Nahkuri *et al*, 2009; Spies *et al*, 2009; Tilgner *et al*, 2009; Schwartz *et al*, 2009b; Chen *et al*, 2010). These studies relied on analysis of MNase-digested chromatin followed by next-generation sequencing in cells from human, *Drosophila melanogaster*, *Caenorhabditis elegans*, and Japanese killifish, and consistently found an approximately 1.5-fold higher level of nucleosome occupancy in exons than in introns. By nucleosome occupancy we refer to the average levels of nucleosome positioning at a given genomic position. The factor underlying this differential nucleosome occupancy appears to be

differences in sequence composition and, in particular, the higher GC content in exons than in introns. Two main lines of evidence support the hypothesis that differences in nucleosome occupancy are sequence-based: First, higher nucleosome occupancy in exons versus introns can be precisely reproduced by computational models predicting nucleosome positioning based solely on DNA sequence (Tilgner *et al*, 2009; Schwartz *et al*, 2009b). Second, this pattern is also reproduced when measuring experimentally derived nucleosome occupancy levels in intergenic regions with exonic sequence composition flanked by sequences with intronic sequence composition (Spies *et al*, 2009). As nucleosomes are drawn by GC-rich sequences and repelled by GC-poor ones (Tillo and Hughes, 2009), and exons are characterized by higher GC content than introns (Tilgner *et al*, 2009; Schwartz *et al*, 2009b), it appears that the juxtaposition of intronic sequences to exonic ones serves as a force that drives nucleosomes to bind exons preferentially.

While positive correlation between exons and nucleosome positioning is indicative of a statistical association between the two, no experimental studies to date have established a causal link between nucleosome occupancy and splicing. Notably, some bioinformatics analyses do link nucleosome occupancy with splicing-related characteristics of the exons: More prominently defined nucleosomes were found in exons with weak splicing signals (Spies *et al*, 2009; Tilgner *et al*, 2009) and in isolated exons (Spies *et al*, 2009), suggesting that nucleosomes 'mark' these exons to insure their identification by the splicing machinery. In addition, the length of DNA wrapped around a mononucleosome, 147 nt, elegantly correlates with the evolutionarily conserved average length of metazoan exons (Tilgner *et al*, 2009; Schwartz *et al*, 2009b). However, somewhat contradicting these results, nucleosome occupancy levels are higher in longer exons than in shorter ones (Spies *et al*, 2009; Tilgner *et al*, 2009; Schwartz *et al*, 2009b), are higher in exons with stronger PPTs than in weaker ones, and were found to be lower in alternatively spliced exons as compared with constitutively spliced ones (Schwartz *et al*, 2009b). Thus, experimental studies that simultaneously manipulate nucleosome positioning while evaluating splicing patterns are required to fully evaluate the presence of a cross-talk between the nucleosome positioning and splicing. This is particularly important, because various alternative scenarios come to mind, which may underlie the association between nucleosome positioning and exon–intron architecture. For example, this association may merely reflect the gradient of GC content between exons and introns without having any regulatory role whatsoever. Alternatively, nucleosomes may confer protection to the exonic sequences coiled around them (Kogan and Trifonov, 2005), as sequences coiled around nucleosome are less sensitive to UV irradiation and chemical attack (Gale *et al*, 1987; Thrall *et al*, 1994) and less prone to indels longer than 1bp base pair (Sasaki *et al*, 2009), although nucleosome-bound sequences are more prone to 1-nt indels than are unbound sequences (Washietl *et al*, 2008; Sasaki *et al*, 2009).

Link between histone modifications and splicing

Considerable advance has been made over the past year in understanding the role of specific chromatin modifications in the context of splicing. An initial study by Kolasinska-Zwierz *et al* (2009) found the H3K36me3 modification, known to mark regions undergoing transcription (Talasz *et al*, 2005; Vakoc *et al*, 2006; Bell *et al*, 2007; Du *et al*, 2008), to be globally enriched in exons. Furthermore, this enrichment was greater in constitutive exons than in alternative ones, although this result was questioned by a subsequent study (Spies *et al*, 2009). As Kolasinska-Zwierz *et al* did not examine the underlying nucleosome positioning patterns, the question remained open whether the H3K36me3 exonic enrichment simply mirrors the underlying nucleosome enrichment, or whether it reflects an independent enrichment. Addressing this issue proved technically challenging, as nucleosome occupancy levels and histone modifications levels are typically derived on the basis of two completely different sets of experiments (MNase-digested chromatin followed by deep sequencing versus chromatin immunoprecipitation followed by deep sequencing or ChIP-seq, respectively) and comparison of data from these two types of experiments is not straightforward. Using different normalization schemes, some studies concluded that H3K36me3 appears to mirror nucleosome occupancy (Tilgner *et al*, 2009; Schwartz *et al*, 2009a), whereas others concluded that H3K36me3 is enriched even after taking into account nucleosome occupancy (Andersson *et al*, 2009; Spies *et al*, 2009). All studies are in agreement that the enrichment is dramatically reduced, once nucleosome occupancy is accounted for.

In parallel, these studies identified additional histone modifications enriched in exons. Surprisingly, although these studies all analysed the same data set of histone modifications in human T cells (Barski *et al*, 2007; Schones *et al*, 2008), they did not consistently identify the same set of histone modifications. In addition to H3K36me3, one study found H3K79me1, H4K20me1, and H2BK5me1 modifications to be enriched in exonic sequences (Schwartz *et al*, 2009b), another found H3K79me1, H2BK5me1, H3K27me1, H3K27me2, and H3K27me3 (Andersson *et al*, 2009), whereas the main candidates of another study were H3K27me2 and H3K4me1(Spies *et al*, 2009). Yet another analysis showed an anti-correlation between H4K20me1 and nucleosome levels (Tilgner *et al*, 2009). The differences among the studies appear to originate from different approaches to analysis of deep-sequencing data and in the different normalization schemes used.

As was the case for nucleosome occupancy, also for histone modifications it remained to establish a causal relationship between histone modifications and splicing; and until recently, only limited experimental data were available. The CHD1 chromatin-remodelling ATPase, which binds to H3K4me3, is known to bind to spliceosomal components, and knockdown of CHD1 and reduction of H3K4me3 levels were shown to alter the efficiency of splicing (Sims *et al*, 2007). In addition, two histone modifications, H3K9 acetylation and H3K36me3, were experimentally shown to be associated with exon skipping (Schor *et al*, 2009). An additional study found an association between splicing patterns and changes in H3K9me2 and H3K27me3 (Allo *et al*, 2009). However, as none of these studies monitored the underlying nucleosome occupancy patterns, it remained possible that changes in splicing attributed to specific histone modifications could stem from nucleosome positioning.

Recently, Luco *et al* (2010) made a considerable advance in establishing the causal role of histone modifications in splicing. The authors found that a range of histone modifications were differentially associated with the splicing pattern of a set of mutually exclusive exons regulated by the PPT-binding protein (PTB). These modifications included H3K36me3 and H3K4me1, acting in one direction, and H3K27me3, H3K4me3, and H3K9me1, acting in the opposite direction. Through modulations of the levels of SET2 and ASH2, the methyltransferases leading to H3K36me3 and H3K4me3 modifications, respectively (Du *et al*, 2008; Edmunds *et al*, 2008), Luco *et al* showed a causal link between the histone modifications and splicing. Furthermore, they showed that the effect was mediated through recruitment of PTB to H3K36me3-modified chromatin through the histone-tail-binding protein MRG15. Thus, this study not only established the causal link between histone modifications and splicing, but also provided experimental evidence of the mechanism by which these processes are linked. It is important to point out that this study did not attempt to examine the effect of nucleosome occupancy thus leaving open the question of to what extent chromatin modifications mirror nucleosome occupancy.

Another major question is whether there is a combinatorial cross-talk between different histone modifications in the context of splicing or whether the effect of each histone modification is independently exerted. Notably, this question is not specific to splicing, but also remains an open question in the field of histone modifications around promoters and within genes (Kouzarides, 2007). Studies using sequential ChIP-seq experiments may begin to shed light on these questions and refine our understanding of the precise chromatin structure around exons.

Additional frontiers

In addition to the questions discussed above, the link between chromatin structure and splicing opens up a plethora of questions pertaining to the characteristics of chromatin around exons, just as the nucleosome-free region upstream from the transcription start sites opened up questions pertaining to the chromatin organization at these sites. These questions include the following:

Mechanism

Assuming that nucleosome positioning and histone modifications can be unequivocally linked with splicing, the biological mechanism linking these two processes must be uncovered. RNAPII may mediate the link by either of the two mechanisms described above. Nucleosomes could serve as speed-bumps for RNAPII, slowing down the transcription rates in the vicinity of exons (Schwartz *et al*, 2009b), thereby improving their recognition by the splicing machinery through kinetic coupling (de la Mata *et al*, 2003; Kornblihtt, 2006; Schor *et al*, 2009) (Figure 1A). The potential of nucleosomes to slow down RNAPII was recently shown by Hodges *et al* (2009a). However, it is also possible that it is not, or not only, the nucleosomes, but also specific histone modifications, that slow down RNAPII. Conversely, they may also speed up RNAPII, as is thought to be the case with histone acetylations (Nogues *et al*, 2002; Figure 1B). Alternatively, by means of its 'landing pad' characteristics, RNAPII may recruit splicing factors to the pre-mRNA; these splicing factors may act in conjunction with splicing factors bound by chromatin (Figure 1C). Specific chromatin landscapes may recruit or interact directly with splicing factors, just as chromatin directs the interaction of transcription regulators with promoters (Cairns, 2009). SR proteins, for instance, may directly interact with chromatin as suggested by at least two studies (Kress *et al*, 2008; Loomis *et al*, 2009) and thereby contribute to splicing. Another possibility is that RNAPII may recruit chromatin modifiers altering the chromatin organization along exons, to allow recruitment of splicing factors to the exon either directly or indirectly, as was elegantly shown recently by Luco *et al* (2010) (Figure 1D). Yet another important question to be considered is the directionality of the association between chromatin structure and splicing. Although the currently favoured model is that nucleosome occupancy precedes splicing and thus helps guide the splicing machinery, the opposite model, in which chromatin structure is altered by the splicing machinery, cannot be ruled out.

Role of histone variants

Although one means of altering the chemical properties of nucleosomes are through histone modifications, a similar effect can be achieved by using different histone variants. Most core histone proteins have several variants (e.g., H2B.1, H2A.Z, H2A.Bbd, H3.3, CENP-A) differing in their amino-acid composition and manifesting different biological characteristics (Segal and Widom, 2009; Talbert and Henikoff, 2010). It will be interesting to find out whether there are differences in the binding profiles of histone variants in exons and introns, and to understand the biology underlying these differences.

Role of DNA methylation

Several studies point to a link between nucleosome positioning and DNA CpG methylation. The emerging consensus is that CpG methylation decreases DNA bending flexibility and hence influences nucleosome positioning (Tippin and Sundaralingam, 1997; Nathan and Crothers, 2002), although it cannot be excluded that the link between the two is mediated indirectly through a range of processes triggered by CpG methylation (Segal and Widom, 2009). In parallel, recent studies indicate that DNA methylation and histone modification pathways are also linked through a wide range of mechanisms (Cedar and Bergman, 2009; Choy *et al*, 2010). For instance, in early embryo development, DNA methylation patterns are established in regions lacking H3K4 methylation, as its presence inhibits DNA methyltransferases (Ooi *et al*, 2007; Cedar and Bergman, 2009). At pericentromeric satellite repeats, histone methyltransferases are responsible for heterochromatization through H3K9me3 histone modification and for recruitment of DNA methyltransferases (Lehnertz *et al*, 2003). Inactivation of Polycomb target genes and the X chromosomes are additional examples of mechanisms in which DNA methylation is believed to be induced by histone modifications (for review see references Cedar and Bergman, 2009). The opposite mechanism is also believed to exist: In this case, DNA methylation induces specific repressive histone modifications (Eden *et al*, 1998; Hashimshony *et al*, 2003). Thus, if exons differ from introns in terms of nucleosome positioning and in terms of histone modifications, they may also differ from each other in terms of DNA methylation status. In support of this conjunction, higher methylation levels were indeed

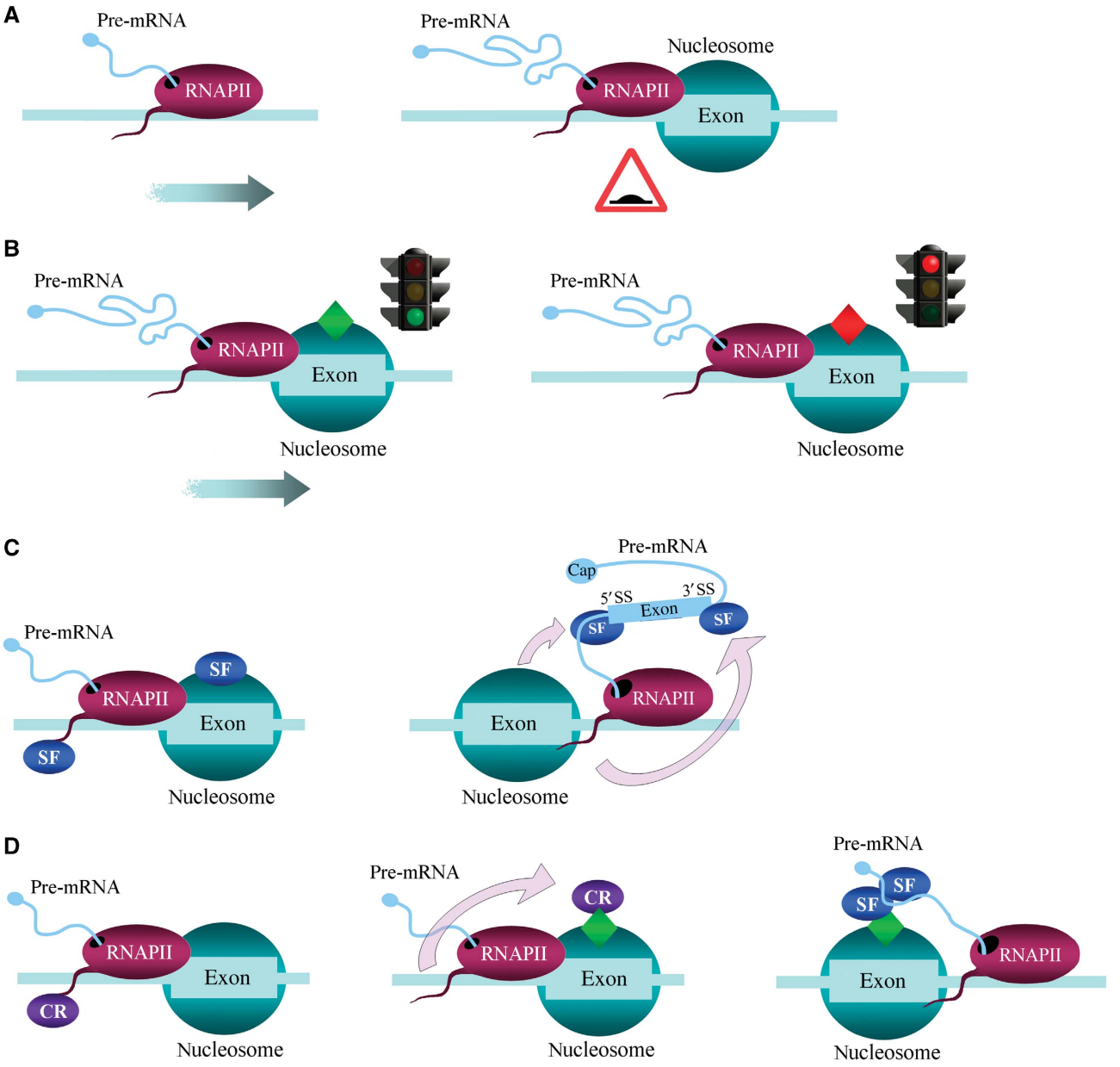

Figure 1 Mechanisms through which nucleosomes may be linked with exon–intron architecture and splicing. (**A**) Nucleosomes may serve as 'speed bumps' for RNAPII, slowing down transcription rates and affecting splicing through the kinetic coupling model. (**B**) Different histone modifications, as indicated by the different coloured diamonds over the nucleosomes, may serve as signals for RNAPII, with one histone modification (e.g., acetylation) increasing transcription rates and others slowing it down. (**C**) The CTD domain of RNAPII serves as a landing pad for splicing factors, which are recruited to the pre-mRNA upon transcription. In parallel, splicing factors directly binding to chromatin are recruited to the pre-mRNA. (**D**) The CTD domain of RNAPII serves as a landing pad for chromatin remodellers, which alter the chromatin conformation along the exons, and thereby recruit splicing factors to exonic regions. CR, chromatin remodeller; SF, splicing factor.

found in exons as compared with introns by two recent studies (Lister *et al*, 2009; Hodges *et al*, 2009b). Are these differential levels of functional consequence or do they only reflect indirect links between the processes? Do they impose an epigenetic layer of regulation on splicing? It will be fascinating to decipher the cross-talk between chromatin structure and DNA methylation in the context of exon–intron architecture and splicing.

Role of RNAPII kinetics

Recent studies indicate that the rate of transcription by RNAPII is 3100–3800bp per minute, whereas exon splicing is thought to be a considerably slower process, possibly in the timeframe of minutes (Singh and Padgett, 2009; Wada *et al*, 2009). Thus, by the time an exon is spliced, RNAPII may be several thousands of nucleotides downstream. In light of the different time scales between transcription and splicing, it will be important to assess to what extent the slowing down of RNAPII by nucleosomes (Hodges *et al*, 2009a) can be 'meaningful' in terms of splicing.

Evolutionary aspects

Prokaryotes lack both nucleosomes and splicing, whereas in eukaryotes both exons and nucleosomes, each of which involves sequences in the range of ~150 nt, have come to

exist. This raises speculations pertaining to an evolutionary link between the two. Did nucleosomes and splicing arise concurrently? Did one set the ground for the other? Furthermore, among mammals, and in particular primates, intron length has gradually expanded, whereas the length of exons has remained relatively constant. This and other findings has led to speculations that in mammals the spliceosome primarily recognizes exons in a process termed exon definition, as opposed to that in fungi in which introns are kept short and are thought to be the recognized unit in a process termed intron definition (Berget, 1995; Ast, 2004; Ram and Ast, 2007). Can this shift from intron definition to exon definition be linked with improved positioning of nucleosomes along mammalian exons?

Conclusion

We opened this paper by stating that it remains an enigma how short exons are recognized within their overwhelming intronic environment. We then summarized recent advances pertaining to our understanding of the potential role of chromatin organization in this process. Does, then, nucleosome positioning contain a substantial amount of information, which, in conjunction with previously found features, can account for the ability of the splicing machinery to recognize exons? Apparently not. First, the relatively low fold change ($\sim$1.5) in terms of nucleosome occupancy between exons and introns suggests that the information borne by nucleosome occupancy is limited. Second, in a study assessing this using a computational model, nucleosome occupancy was found to improve recognition of exons, but to a very limited degree (Spies *et al*, 2009). Thus, nucleosome occupancy brings us one small step forward in understanding the workings of the splicing machinery, but we remain far from 'explaining away' this remarkable process. While this is suggestive of additional features remaining to be discovered or quantified (including histone modification levels), we close with the speculation that perhaps it is precisely this inherent stochasticity in the genome sequence, this intrinsic absence of determinative information (see also Fox-Walsh and Hertel, 2009), that sets the ground for alternative splicing and for enriching our transcriptome with previously non-existing transcripts.

Acknowledgements

We are grateful to Dror Hollander for help with preparing the figure. SS is a fellow of the Edmond J Safra bioinformatics program at Tel Aviv University. Funding for open access charge: Israel Science Foundation (ISF 61/09); Joint Germany-Israeli Research Program (ca-139); Deutsche-Israel Project (DIP MI-1317); and European Alternative Splicing Network (EURASNET). This work was performed in partial fulfilment of the requirements for a PhD degree of SS (Sackler Faculty of Medicine, Tel Aviv University, Israel).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Allemand E, Batsche E, Muchardt C (2008) Splicing, transcription, and chromatin: a menage a trois. *Curr Opin Genetics Dev* **18:** 145–151

Allo M, Buggiano V, Fededa JP, Petrillo E, Schor I, de la Mata M, Agirre E, Plass M, Eyras E, Elela SA, Klinck R, Chabot B, Kornblihtt AR (2009) Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* **16:** 717–724

Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J (2009) Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res* **19:** 1732–1741

Ast G (2004) How did alternative splicing evolve? *Nat Rev* **5:** 773–782

Auboeuf D, Honig A, Berget SM, O'Malley BW (2002) Coordinate regulation of transcription and splicing by steroid receptor coregulators. *Science (New York, NY)* **298:** 416–419

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823–837

Batsche E, Yaniv M, Muchardt C (2006) The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13:** 22–29

Bell O, Wirbelauer C, Hild M, Scharf AN, Schwaiger M, MacAlpine DM, Zilbermann F, van Leeuwen F, Bell SP, Imhof A, Garza D, Peters AH, Schubeler D (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *EMBO J* **26:** 4974–4984

Berget SM (1995) Exon recognition in vertebrate splicing. *J Biol Chem* **270:** 2411–2414

Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. *Cell* **128:** 669–681

Beyer AL, Osheim YN (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2:** 754–765

Bird G, Zorio DA, Bentley DL (2004) RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3′-end formation. *Mol Cell Biol* **24:** 8963–8969

Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72:** 291–336

Bres V, Gomes N, Pickle L, Jones KA (2005) A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev* **19:** 1211–1226

Cairns BR (2009) The logic of chromatin architecture and remodelling at promoters. *Nature* **461:** 193–198

Cedar H, Bergman Y (2009) Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev* **10:** 295–304

Chen W, Luo L, Zhang L (2010) The organization of nucleosomes around splice sites. *Nucleic Acids Res* (advance online publication; doi:10.1093/nar/gkq007)

Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee TH. (2010) DNA methylation increases nucleosome compaction and rigidity. *J Am Chem Soc* **132:** 1782–1783

Damgaard CK, Kahns S, Lykke-Andersen S, Nielsen AL, Jensen TH, Kjems J (2008) A 5′ splice site enhances the recruitment of basal transcription initiation factors *in vivo*. *Mol Cell* **29:** 271–278

Das R, Dufu K, Romney B, Feldt M, Elenko M, Reed R (2006) Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev* **20:** 1100–1109

Das R, Yu J, Zhang Z, Gygi MP, Krainer AR, Gygi SP, Reed R (2007) SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell* **26:** 867–881

de Almeida SF, Carmo-Fonseca M (2008) The CTD role in cotranscriptional RNA processing and surveillance. *FEBS Lett* **582:** 1971–1976

de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR (2003) A slow RNA polymerase II affects alternative splicing *in vivo*. *Mol Cell* **12:** 525–532

de la Mata M, Kornblihtt AR (2006) RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat Struct Mol Biol* **13:** 973–980

Du HN, Fingerman IM, Briggs SD (2008) Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. *Genes Dev* **22:** 2786–2798

Dye MJ, Gromak N, Proudfoot NJ (2006) Exon tethering in transcription by RNA polymerase II. *Mol Cell* **21:** 849–859

Eden S, Hashimshony T, Keshet I, Cedar H, Thorne AW (1998) DNA methylation models histone acetylation. *Nature* **394:** 842

Edmunds JW, Mahadevan LC, Clayton AL (2008) Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J* **27:** 406–420

Fong N, Bird G, Vigneron M, Bentley DL (2003) 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:** 4274–4282

Fox-Walsh KL, Hertel KJ (2009) Splice-site pairing is an intrinsically high fidelity process. *Proc Natl Acad Sci USA* **106:** 1766–1771

Gale JM, Nissen KA, Smerdon MJ (1987) UV-induced formation of pyrimidine dimers in nucleosome core DNA is strongly modulated with a period of 10.3 bases. *Proc Natl Acad Sci USA* **84:** 6644–6648

Goldstrohm AC, Albrecht TR, Sune C, Bedford MT, Garcia-Blanco MA (2001) The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. *Mol Cell Biol* **21:** 7617–7628

Graveley BR (2001) Alternative splicing: increasing diversity in the proteomic world. *Trends Genet* **17:** 100–107

Green MR, Maniatis T, Melton DA (1983) Human beta-globin pre-mRNA synthesized *in vitro* is accurately spliced in *Xenopus* oocyte nuclei. *Cell* **32:** 681–694

Gunderson FQ, Johnson TL (2009) Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. *PLoS Genet* **5:** e1000682

Hashimshony T, Zhang J, Keshet I, Bustin M, Cedar H (2003) The role of DNA methylation in setting up chromatin structure during development. *Nat Genet* **34:** 187–192

Hicks MJ, Yang CR, Kotlajich MV, Hertel KJ (2006) Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol* **4:** e147

Hodges C, Bintu L, Lubkowska L, Kashlev M, Bustamante C (2009a) Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science (New York, NY)* **325:** 626–628

Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ, Hicks JB (2009b) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res* **19:** 1593–1605

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

Howe KJ (2002) RNA polymerase II conducts a symphony of pre-mRNA processing activities. *Biochim Biophys Acta* **1577:** 308–324

Ioshikhes IP, Albert I, Zanton SJ, Pugh BF (2006) Nucleosome positions predicted through comparative genomics. *Nat Genet* **38:** 1210–1215

Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell* **128:** 683–692

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, Segal E (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458:** 362–366

Kogan S, Trifonov EN (2005) Gene splice sites correlate with nucleosome positions. *Gene* **352:** 57–62

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* **41:** 376–381

Kornberg RD, Lorch Y (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98:** 285–294

Kornblihtt AR (2006) Chromatin, transcript elongation and alternative splicing. *Nat Struct Mol Biol* **13:** 5–7

Kornblihtt AR, de la Mata M, Fededa JP, Munoz MJ, Nogues G (2004) Multiple links between transcription and splicing. *RNA (New York, NY)* **10:** 1489–1498

Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128:** 693–705

Kress TL, Krogan NJ, Guthrie C (2008) A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. *Mol Cell* **32:** 727–734

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13:** 1192–1200

LeMaire MF, Thummel CS (1990) Splicing precedes polyadenylation during *Drosophila* E74A transcription. *Mol Cell Biol* **10:** 6059–6063

Lim LP, Burge CB (2001) A computational analysis of sequence features involved in recognition of short introns. *Proc Natl Acad Sci USA* **98:** 11193–11198

Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu XD (2008) The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* **15:** 819–826

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462:** 315–322

Listerman I, Sapra AK, Neugebauer KM (2006) Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol* **13:** 815–822

Long JC, Caceres JF (2009) The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* **417:** 15–27

Loomis RJ, Naoe Y, Parker JB, Savic V, Bozovsky MR, Macfarlan T, Manley JL, Chakravarti D (2009) Chromatin binding of SRp20 and ASF/SF2 and dissociation from mitotic chromosomes is modulated by histone H3 serine 10 phosphorylation. *Mol Cell* **33:** 450–461

Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T (2010) Regulation of alternative splicing by histone modifications. *Science (New York, NY)* **327:** 996–1000

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385:** 357–361

Munoz MJ, Perez Santangelo MS, Paronetto MP, de la Mata M, Pelisch F, Boireau S, Glover-Cutter K, Ben-Dov C, Blaustein M, Lozano JJ, Bird G, Bentley D, Bertrand E, Kornblihtt AR (2009) DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. *Cell* **137:** 708–720

Nahkuri S, Taft RJ, Mattick JS (2009) Nucleosomes are preferentially positioned at exons in somatic and sperm cells. *Cell Cycle (Georgetown, Tex)* **8:** 3420–3424

Nathan D, Crothers DM (2002) Bending and flexibility of methylated and unmethylated *Eco*RI DNA. *J Mol Biol* **316:** 7–17

Neugebauer KM (2002) On the importance of being co-transcriptional. *J Cell Sci* **115:** 3865–3871

Nilsen TW, Graveley BR (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463:** 457–463

Nogues G, Kadener S, Cramer P, Bentley D, Kornblihtt AR (2002) Transcriptional activators differ in their abilities to control alternative splicing. *J Biol Chem* **277:** 43110–43114

Ohkura N, Takahashi M, Yaguchi H, Nagamura Y, Tsukada T (2005) Coactivator-associated arginine methyltransferase 1, CARM1, affects pre-mRNA splicing in an isoform-specific manner. *J Biol Chem* **280:** 28927–28935

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* **448:** 714–717

Pandya-Jones A, Black DL (2009) Co-transcriptional splicing of constitutive and alternative exons. *RNA (New York, NY)* **15:** 1896–1908

Perales R, Bentley D (2009) 'Cotranscriptionality': the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* **36:** 178–191

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. *Cell* **108:** 501–512

Ram O, Ast G (2007) SR proteins: a foot on the exon before the transition from intron to exon definition. *Trends Genet* **23:** 5–7

Roberts GC, Gooding C, Mak HY, Proudfoot NJ, Smith CW (1998) Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res* **26:** 5568–5572

Sasaki S, Mello CC, Shimada A, Nakatani Y, Hashimoto S, Ogawa M, Matsushima K, Gu SG, Kasahara M, Ahsan B, Sasaki A, Saito T, Suzuki Y, Sugano S, Kohara Y, Takeda H, Fire A, Morishita S (2009) Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science (New York, NY)* **323:** 401–404

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132:** 887–898

Schor IE, Rascovan N, Pelisch F, Allo M, Kornblihtt AR (2009) Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci USA* **106:** 4325–4330

Schwartz S, Gal-Mark N, Kfir N, Oren R, Kim E, Ast G (2009a) Alu exonization events reveal features required for precise recognition of exons by the splicing machinery. *PLoS Comput Biol* **5:** e1000300

Schwartz S, Meshorer E, Ast G (2009b) Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol* **16:** 990–995

Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, Wang JP, Widom J (2006) A genomic code for nucleosome positioning. *Nature* **442:** 772–778

Segal E, Widom J (2009) What controls nucleosome positions? *Trends Genet* **25:** 335–343

Sims III RJ, Millhouse S, Chen CF, Lewis BA, Erdjument-Bromage H, Tempst P, Manley JL, Reinberg D (2007) Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* **28:** 665–676

Singh J, Padgett RA (2009) Rates of *in situ* transcription and splicing in large human genes. *Nat Struct Mol Biol* **16:** 1128–1133

Smith CW, Valcarcel J (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci* **25:** 381–388

Spies N, Nielsen CB, Padgett RA, Burge CB (2009) Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* **36:** 245–254

Talasz H, Lindner HH, Sarg B, Helliger W (2005) Histone H4-lysine 20 monomethylation is increased in promoter and coding regions of active genes and correlates with hyperacetylation. *J Biol Chem* **280:** 38814–38822

Talbert PB, Henikoff S (2010) Histone variants—ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11:** 264–275

Tardiff DF, Lacadie SA, Rosbash M (2006) A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly post-transcriptional. *Mol Cell* **24:** 917–929

Tennyson CN, Klamut HJ, Worton RG (1995) The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nat Genet* **9:** 184–190

Thrall BD, Mann DB, Smerdon MJ, Springer DL (1994) Nucleosome structure modulates benzo[a]pyrenediol epoxide adduct formation. *Biochemistry* **33:** 2210–2216

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

Tillo D, Hughes TR (2009) G + C content dominates intrinsic nucleosome occupancy. *BMC Bioinformatics* **10:** 442

Tippin DB, Sundaralingam M (1997) Nine polymorphic crystal structures of d(CCGGGCCCGG), d(CCGGGCCm5CGG), d(Cm5CGGGCCm5CGG) and d(CCGGGCC(Br)5CGG) in three different conformations: effects of spermine binding and methylation on the bending and condensation of A-DNA. *J Mol Biol* **267:** 1171–1185

Vakoc CR, Sachdeva MM, Wang H, Blobel GA (2006) Profile of histone lysine methylation across transcribed mammalian chromatin. *Mol Cell Biol* **26:** 9185–9195

Vignali M, Hassan AH, Neely KE, Workman JL (2000) ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* **20:** 1899–1910

Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H *et al* (2009) A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci USA* **106:** 18357–18361

Wang GS, Cooper TA (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev* **8:** 749–761

Washietl S, Machne R, Goldman N (2008) Evolutionary footprints of nucleosome positions in yeast. *Trends Genet* **24:** 583–587

Wetterberg I, Bauren G, Wieslander L (1996) The intranuclear site of excision of each intron in Balbiani ring 3 pre-mRNA is influenced by the time remaining to transcription termination and different excision efficiencies for the various introns. *RNA (New York, NY)* **2:** 641–651

Wuarin J, Schibler U (1994) Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing. *Mol Cell Biol* **14:** 7219–7225

Yuryev A, Patturajan M, Litingtung Y, Joshi RV, Gentile C, Gebara M, Corden JL (1996) 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 USA* **93:** 6975–6980

