COMMENTARY

Cell biology of transcription and pre-mRNA splicing: nuclear architecture meets nuclear function

Tom Misteli
National Cancer Institute, NIH, 41 Library Drive, Bethesda, MD 20892, USA
(e-mail: mistelit@mail.nih.gov)

Published on WWW 10 May 2000

SUMMARY

Gene expression is a fundamental cellular process. The basic mechanisms involved in expression of genes have been characterized at the molecular level. A major challenge is now to uncover how transcription, RNA processing and RNA export are organized within the cell nucleus, how these processes are coordinated with each other and how nuclear architecture influences gene expression and regulation. A significant contribution has come from cell biological approaches, which combine molecular techniques with microscopy methods. These studies have revealed that the mammalian cell nucleus is a complex but highly organized organelle, which contains numerous subcompartments. I discuss here how two essential nuclear processes – transcription and pre-mRNA splicing – are spatially organized and coordinated in vivo, and how this organization might contribute to the control of gene expression. The dynamic nature of nuclear proteins and compartments indicates a high degree of plasticity in the cellular organization of nuclear functions. The cellular organization of transcription and splicing suggest that the morphology of nuclear compartments is largely determined by the activities of the nucleus.

Key words: Pre-mRNA splicing, Transcription, Nuclear architecture, Phosphorylation

INTRODUCTION

Gene expression is a multistep process. In higher eukaryotes, it involves transcription, 5′ capping, pre-mRNA splicing, 3′ processing, RNA export and eventually translation in the cytoplasm. Powerful biochemical and genetic approaches have led to the identification of many factors that mediate these steps and have revealed the basic principles of each step at the molecular level. Most studies have analyzed the functions of single genes and of single steps in the gene expression process. Only recently has the focus shifted to a more ‘holistic’ approach in which large sets of genes are analyzed and the interplay of multiple steps is investigated. Cell biological approaches, such as high-resolution microscopy, are crucial tools in the study of gene expression in vivo, since they allow analysis of gene expression at the single cell level, in living cells and in a spatial and temporal context.

A large gap exists in our understanding of how gene expression processes are organized within the space of the cell nucleus and how they are related to nuclear architecture. It is unclear how molecules find their binding partners in the nucleus and how proteins are targeted to particular nuclear sites. Furthermore, one has to ask to what degree and by what mechanisms nuclear architecture might directly or indirectly affect expression of genes. Does it matter where in the nucleus a gene is positioned? What is the relationship between nuclear architecture and chromatin organization? A further key issue in understanding gene expression is how single gene expression steps are coordinated with each other. The molecular dissection of this aspect of gene expression has been hampered by the virtual absence of in vitro systems that faithfully reproduce multiple gene expression steps. Uncovering of these links and the mechanisms by which single steps are coupled will reveal important regulatory mechanisms of gene expression. Through the combined application of molecular methods and microscopy technology we are now for the first time in a position to address these fundamental cell biological aspects of gene expression.

THE HETEROGENEOUS CELL NUCLEUS

The mammalian cell nucleus is a complex, multi-functional organelle (Spector, 1993; Lamond and Earnshaw, 1998; Misteli and Spector, 1998; Matera, 1999). Not only is it the repository of the vast majority of hereditary information and the major site of gene expression, but it is also the site of DNA replication, repair, as well as DNA and RNA modifications. The high degree of functional complexity might suggest high morphological complexity as well. However, only a few years
ago, the commonly held view of the mammalian cell nucleus was one of a poorly organized organelle. The morphological methods that were so powerful in elucidating the various cytoplasmic organelles had been largely unsuccessful in finding significant morphologically defined structures within the nucleus. The absence of distinct morphological features led to the view that the nucleus is a ‘chaotic’ organelle that contains large amounts of disorganized DNA and homogeneously distributed proteins. In the early 1980s Cremer and colleagues revived ideas first proposed by Rabl and Boveri at the end of the 19th century suggesting a non-random organization of interphase chromosomes (Rabl, 1885; Boveri, 1888). In a series of elegant experiments Cremer et al. (1982a,b) demonstrated that particular chromosomes occupy distinct territories in the interphase nucleus. Those observations were confirmed by the development of methods for direct visualization of chromosomes by fluorescence microscopy (Pinkel et al., 1988). It is now accepted that chromosomes occupy distinct volumes of the nucleus, and in some cases, such as particular stages of embryonic development in Drosophila melanogaster or in plants, the position of chromosomes is well defined (Marshall et al., 1997).

Not only are chromosomes spatially organized within the nuclear space, but the nucleus also contains numerous distinct subcompartments. The presence of morphological compartments must be considered a fundamental feature of the mammalian cell nucleus. The best-characterized nuclear compartment and the only one that has a clearly identified function is the nucleolus, the site of rRNA synthesis (Fig. 1A; Scheer and Hock, 1999; Olson et al., 2000). The nucleolus is a self-assembling compartment formed by the aggregation of DNA sequences on several chromosomes, which contain repeated rRNA genes. Disruption of ribosomal genes by translocation results in the formation of multiple individual nucleoli (McClintock, 1934), and inhibition of RNA pol I transcription leads to reorganization of the nucleolus in interphase cells. This suggests that transcriptional activity is a driving force for the maintenance of nucleolar integrity (Melese and Xue, 1995). The nucleolus has recently attracted renewed interest because the compartment also appears to serve additional functions, such as sequestration of nuclear proteins (Pederson, 1998; Olson et al., 2000). For example, interaction of the Mdm2 protein with the tumor suppressor p53 is required for the export of p53 from the nucleus and its degradation in the cytoplasm. Mdm2 is sequestered to the nucleolus by Ink4/Arf, which prevents its interaction with p53 and therefore allows the tumor suppressor activity of p53 (Tao and Levine, 1999; Weber et al., 1999).

In addition to the nucleolus, a growing family of small nuclear bodies, often referred to as foci because of their appearance by fluorescence microscopy, is present in nuclei (Fig. 1B; Matera, 1999). The two best-characterized nuclear bodies, the PML body (promyelocytic leukemia body) and the Cajal Body*, are both thought to form in response to transcriptional activity of genes. However, the function of none of the small nuclear bodies has been asserted. A large volume of nuclear space is occupied by a compartment commonly referred to as ‘speckles’ and henceforth termed splicing factor compartments (SFCs; Spector, 1993; Misteli and Spector, 1998). SFCs occupy ~20% of total nuclear volume and were originally defined on the basis of the presence of high concentrations of pre-mRNA splicing factors (Fig. 1C; Beck, 1961; Spector, 1990; Spector et al., 1991). In addition to pre-mRNA splicing factors and snRNAs, SFCs also contain transcription factors (Larsson et al., 1995; Mortillaro et al., 1996; Zeng et al., 1997), 3'-processing factors (Krause et al., 1994; Schul et al., 1998b) and ribosomal proteins (Mintz et al., 1999). Biochemical purification of the nuclear structures, which correspond to SFCs, indicates that the compartment contains about 150 proteins (Mintz et al., 1999; D. Spector, personal communication), although it remains to be determined which are bona fide residents. By electron microscopy, SFCs correspond to two morphologically distinct structures: the central regions consist of clusters of 20-nm granules, the so-called interchromatin granule clusters; the peripheral regions are perichromatin fibrils, which are believed to represent nascent transcripts (Fakan, 1994). These two morphological structures cannot be distinguished by fluorescence microscopy.

Nuclear compartments are distinct suborganelles not unlike the cytoplasmic organelles. They contain a distinct set of resident proteins, they are morphologically identifiable by light and electron microscopy and, most importantly, some of them have been biochemically purified (Busch et al., 1967; Mintz et al., 1999; Y. W. Lam and A. Lamond, personal communication). The main difference from cytoplasmic organelles is that nuclear compartments are not defined by membranes. The absence of membranes raises the intriguing question of how nuclear compartments are formed and maintained in vivo. A possible answer comes from recent in vivo photobleaching experiments on fusion proteins between GFP and several nuclear proteins (Phair and Misteli, 2000). We have shown that, although nuclear compartments are stable

*Gall recently proposed the term ‘Cajal Body’ in place of ‘coiled body’ or ‘accessory body’ in honor of the structure’s discoverer, Ramon Y Cajal. The term ‘Cajal Body’ is used throughout this review.
overall, proteins rapidly associate and dissociate with them and that the compartments are in continuous flux. For example, we calculate that for splicing factor SF2/ASF the residence time in SFCs is not more than ~45 seconds and that at least 10,000 molecules of SF2/ASF are lost from SFCs per second per cell, which results in the complete turnover of SF2/ASF in SFCs within less than 2 minutes. These observations are consistent with a self-assembly process in which proteins interact with each other and in this way form ‘morphological’ structures (Phair and Misteli, 2000).

**TRANSCRIPTION AND PRE-mRNA SPlicING IN THE NUCLEUS**

In a growing mammalian cell, pre-mRNA splicing factors are enriched in SFCs. In addition, splicing factors are also diffusely distributed throughout the nucleoplasm (Fig. 1C). The majority of pre-mRNA splicing occurs co-transcriptionally and thus at the site of transcription (Beyer and Osheim, 1988; Bäuren and Wieslander, 1994; Neugebauer and Roth, 1997). Microscopy mapping data place the sites of most transcription away from SFCs or at their periphery (Fig. 2a; Wansink et al., 1993; Dirks et al., 1997). Similarly, in-situ hybridization studies using splice junction probes demonstrate that splicing of most genes occurs predominantly at the periphery of SFCs (Zhang et al., 1994; Xing et al., 1995; Huang and Spector, 1996). The association of transcription sites with SFCs does not reflect simply the amount of splicing factor bound to transcripts, but is influenced in a locus-specific manner (Jolly et al., 1999; Smith et al., 1999). Further evidence that transcription and splicing generally occur at the periphery of SFCs is the localization of acetylated histones, a general indicator of transcriptionally active or competent DNA, to the periphery of SFCs but very rarely to their interior (Hendzel et al., 1998). These studies point to the existence of mechanisms by which splicing factors, and presumably other processing proteins, are recruited from SFCs and are targeted to active sites of transcription.

Consistent with this notion are several observations that indicate that changes in the overall transcription level of a cell cause redistribution of splicing factors. Upon infection of cells with adenovirus, which produces extensively spliced transcripts, a large pool of splicing factors is recruited from SFCs to viral transcription and splicing sites, where the factors accumulate (Bridge et al., 1993; Jiménez-García and Spector, 1993). In contrast, treatment of cells with general inhibitors of transcription or heat-shock, resulting in transcriptional silencing of most endogenous genes, causes the enlargement of SFCs and the accumulation of splicing factors in SFCs (Fig. 2b; Sinclair and Brasch, 1978; Spector et al., 1991; Melcak et al., 2000). Upon infection with herpes virus 1, which is virtually devoid of intron-containing genes, no recruitment is observed (Phelan et al., 1993), which suggests that recruitment is intron dependent. The presence of recruitment to intron-less cDNAs confirms this; if introns are present they efficiently recruit splicing factors in vivo (Huang and Spector, 1996). We demonstrated directly that splicing factors can be recruited from SFCs to transcription sites by visualization of pre-mRNA splicing factors in living cells, using time-lapse microscopy (Misteli et al., 1997). Upon activation of stably expressed genes, splicing factor SF2/ASF dissociates from SFCs and moves towards newly induced transcription sites, where it accumulates. These observations demonstrate that splicing factors can be recruited from SFCs to transcription sites in vivo and suggest that one function of SFCs is the storage/assembly of spliceosomal components.

It is not clear at present to what extent, if at all, RNAs also move towards SFCs. Successful pre-mRNA splicing not only is an essential step in RNA maturation, but also serves as a quality-control step. Transcripts that are not spliced are generally not exported, and unspliced transcripts are not released from the spliceosome and the transcription site, which indicates that completion of splicing is a prerequisite for release from the transcription site (Custodio et al., 1999). However, RNAs that are very highly expressed or contain many introns might overwhelm the splicing machinery and be released from the transcription sites prior to completion of splicing (Fig. 2c). RNA movement towards and possibly into SFCs might be a feature of most RNAs, since newly synthesized transcripts accumulate in SFCs upon inhibition of transcription (Melcak et al., 2000). Furthermore, a population of fully processed polyadenylated RNA accumulates in SFCs, although the role of this fraction of RNA is unclear (Huang et al., 1994).

**SPLICING FACTOR TARGETING MECHANISMS**

If splicing factors are targeted to transcription sites in vivo,
what are the targeting mechanisms? Numerous lines of evidence point toward splicing factor phosphorylation as a major control mechanism for the localization of at least the members of the SR family of splicing factors (Fig. 3A; Misteli and Spector, 1997, 1998). SR proteins are characterized by the presence of either one or two RNA-binding domains and a highly phosphorylated C-terminal RS domain largely consisting of Arg-Ser dipeptides (Fu, 1995; Manley and Tacke, 1996). A first hint as to how the association of splicing factors with their compartments might be controlled came from the identification of a protein kinase termed SRPK-1 (SR protein kinase-1) that specifically phosphorylates SR proteins and is activated at the onset of M-phase (Gui et al., 1994). During this time, splicing factors are released from SFCs, which leads to the disassembly of the compartments. Overexpression of SRPK-1 mimicked this behavior by releasing factors from SFCs. Numerous splicing factor specific protein kinases, some of which are enriched in SFCs that have properties similar to those of SRPKs have recently been characterized (Table 1). The kinases belong to several distinct protein families and are components of various signal transduction pathways. What the specific role of each kinase is, and how much they functionally overlap, remains to be determined. We showed that phosphorylation of the RS domain is necessary for targeting to a transcription site by analysis of phosphorylation-deficient mutant SR proteins, which could not be recruited to transcription sites in vivo (Misteli et al., 1998). Furthermore, time-lapse microscopy experiments in living cells showed that release of splicing factors from SFCs was blocked in the presence of a kinase inhibitor (Misteli et al., 1997). Apart from a role in control of intranuclear distribution, phosphorylation of SR proteins has also been implicated in efficient nuclear import of SR proteins either after their synthesis or as part of a nuclear-cytoplasmic shuttling pathway (Cáceres et al., 1998; Yeakley et al., 1999).

If splicing factor phosphorylation promotes release of factors from SFCs, one might predict that dephosphorylation promotes association. That is indeed the case. In permeabilized cells, accumulation of splicing factors upon inhibition of transcription requires a Ser/Thr protein phosphatase I activity, which indicates that the reassociation of splicing factors requires the removal of a phosphate group(s) (Misteli and Spector, 1996). This observation is consistent with the demonstration that PP1 and PP2A activities are required for the progression of the splicing reaction in vitro (Mermoud et al., 1992, 1994). In addition, the Krainer laboratory recently identified a PP2C as a component of the spliceosome (Murray et al., 1999). These opposing effects of kinases and phosphatases on splicing factor localization strongly suggest that the distribution of splicing factors is controlled by a cycle of phosphorylation and dephosphorylation (Fig. 3A).

Phosphorylation also affects the function of splicing factors. In vitro binding studies and expression of human proteins in a yeast mutant strain that lacks SRPK-1 activity show that phosphorylation of a protein can differentially affect its interaction with binding partners (Xiao and Manley, 1997, 1998; Yeakley et al., 1999). For example, upon phosphorylation SF2/ASF interacts more readily with the U1-70K protein but interacts less well with SRp40, whereas its interaction with U2AF35 is not affected. These differential
effects on binding probably contribute to the dynamic reorganizations and the switching of binding partners during the splicing reaction (Misteli, 1999). The possibility that changes in localization also reflect the altered function of splicing factors cannot be ruled out at present. An argument against this scenario is the fact that a functionally inactive fusion protein in which the RS domain of SRp20 is linked to the normally cytoplasmic nucleoplasmic core domain is targeted in a normal pattern to SFCs and the nucleoplasm (Cáceres et al., 1997).

**FUNCTIONS FOR SPlicing FACTOR COMPARTMENTS**

Why are splicing factors organized in such a peculiar manner? One possibility is that SFCs are a means to control the concentration of splicing factors in the nucleoplasm and thus the sites of transcription and splicing (Fig. 3B, a). Balancing the concentration of splicing factors might contribute to optimizing in vivo splicing efficiency. A second possibility is a role for SFCs in recycling, or reactivation, of splicing factors (Fig. 3B, b). Splicing factors have a long half-life and only 10-20% of any given splicing factor is engaged in the splicing reaction at any given time in vivo (T. Misteli, unpublished). Therefore, each molecule participates in many splicing reactions during its lifetime and must be recycled. It is conceivable that recycling involves reversible phosphorylation, which might take place in the SFCs (Fig. 3B, b). SFCs are also ideally suited to play a role in alternative-splice-site selection (Fig. 3B, c). Splice-site usage is determined by the relative levels of these two proteins thus modulate splice-site selection (Mayeda and Krainer, 1992; Cáceres et al., 1994; Yang et al., 1994). In vivo, the intracellular localization of a splicing factor can be used to modulate alternative splice-site use. Upon stress signaling, hnRNP A1 accumulates in the cytoplasm, and the alternative splicing pattern of an adenovirus E1A pre-mRNA splicing reporter changes. These effects are mediated by the MKK-p38 signaling pathway (van der Houven van Oordt et al., 2000). The localization of splicing factors within the nucleus might also affect alternative splice-site choice. Overexpression of the Clk/Sty kinases results in the release of SR proteins from SFCs concomitant with a shift in splice sites in both exogenous and endogenous transcripts (Duncan et al., 1997, 1998), although this effect might also be a consequence of redistribution of SF2/ASF from the nucleus to the cytoplasm upon expression of Clk/STY (Cáceres et al., 1998).

The spatial organization of transcription and splicing might also contribute to the recently documented functional coupling of the two processes in that it promotes assembly of complexes containing components of the transcription and splicing machineries (Fig. 3B, d; Bentley, 1999). The C-terminal domain (CTD) of the largest subunit of RNA polymerase II (RNA pol II) has been implicated in linking the transcription machinery with the RNA-processing machineries. The CTD is a highly conserved sequence of YSPTSPS repeats (Corden and Patturajan, 1997). It is believed to be hypophosphorylated during initiation and hyperphosphorylated during the elongation phase of transcription (Corden and Patturajan, 1997). 3'-processing enzymes and 5' -capping enzymes associate with the hyperphosphorylated form, and both processes are inhibited when the CTD is deleted or severely truncated (Cho et al., 1997; McCracken et al., 1997a,b; Yue et al., 1997). The observation that truncation of the CTD prevents splicing in vivo and reduces splicing efficiency in vitro suggests that a direct link exists between the splicing and transcription machineries in vivo (McCracken et al., 1997b; Hirose et al., 1999). Pre-mRNA splicing factors can also be immunoprecipitated with RNA pol II subunits (Mortillaro et al., 1996; Vincent et al., 1996; Yuriev et al., 1996; Bourquain et al., 1997; Du and Warren, 1997; Kim et al., 1997; Tanner et

<table>
<thead>
<tr>
<th>Localization in SFCs</th>
<th>Substrates</th>
<th>Redistributes splicing factors</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRPK-1</td>
<td>+</td>
<td>SR proteins</td>
<td>+</td>
</tr>
<tr>
<td>SRPK-2</td>
<td>+</td>
<td>SR-like proteins</td>
<td>+</td>
</tr>
<tr>
<td>Clk/Sty 1</td>
<td>+</td>
<td>SR proteins</td>
<td>+</td>
</tr>
<tr>
<td>Clk/Sty-2</td>
<td>+</td>
<td>SR proteins</td>
<td>+</td>
</tr>
<tr>
<td>Clk/Sty-3</td>
<td>+</td>
<td>SR proteins</td>
<td>+</td>
</tr>
<tr>
<td>PFTSLRE</td>
<td>+</td>
<td>N.D.</td>
<td>-</td>
</tr>
<tr>
<td>PIP</td>
<td>+</td>
<td>N.D.</td>
<td>-</td>
</tr>
<tr>
<td>Topoisomerase I</td>
<td>+</td>
<td>SR proteins</td>
<td>N.D.</td>
</tr>
<tr>
<td>U1-70K kinase*</td>
<td>N.D.</td>
<td>U1-70K</td>
<td>N.D.</td>
</tr>
<tr>
<td>S. pombe Pp44p</td>
<td>-</td>
<td>SR-protiens</td>
<td>-</td>
</tr>
<tr>
<td>S. pombe Dsk1p</td>
<td>-</td>
<td>SR proteins</td>
<td>-</td>
</tr>
<tr>
<td>S. cerevisiae Sky1p</td>
<td>N.D.</td>
<td>SR protein</td>
<td>-</td>
</tr>
<tr>
<td>PP1-activity</td>
<td>-</td>
<td>SR proteins</td>
<td>+</td>
</tr>
<tr>
<td>PP2A-activity</td>
<td>N.D.</td>
<td>SR-protiens</td>
<td>N.D.</td>
</tr>
<tr>
<td>SCF1 (PP2Cγ)</td>
<td>-</td>
<td>N.D.</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

*Identical to SRPK-1 (S. Gunderson, personal communication).
NUCLEAR ARCHITECTURE IN GENE EXPRESSION

Nuclear architecture is important for the accurate spatial coordination of transcription and pre-mRNA splicing. Is this the only contribution nuclear architecture makes to gene expression? Likely not. One of the most intriguing unanswered questions in the field of gene function is still whether the position of a gene in the nucleus matters for its expression. Croft et al. (1999) have demonstrated that human chromosome 18, which contains very few active genes, is usually found at the periphery of nuclei, whereas the transcriptionally active chromosome 19 is typically found more in the center of the nucleus. Similarly, upon differentiation of PC12 cells, hypersensitive chromatin regions underwent a transposition from the nuclear interior to the periphery (Park and De Boni, 1996). What is clear is that not all regions in the nucleus support gene expression to similar levels. From a morphological point of view it is well established that the nuclear periphery contains large amounts of heterochromatin that contains largely inactive genes. In D. melanogaster, the brown locus is silenced by translocation from a euchromatic environment to a heterochromatic environment (Dernburg et al., 1996). A similar event occurs in maturing B cells, in which lymphoid-associated genes are silenced by association with centromeric regions (Brown et al., 1997, 1999). An impressive example of a correlation between transcriptional activity and gene position was recently reported. In stable cell lines expressing various forms of the 5′ HS2 enhancer from the β-globin locus control region, transcriptionally inactive reporter-gene mutants were closely associated with centromeric heterochromatin, whereas transcriptionally active mutants lay outside of heterochromatin (Francastel et al., 1999). Finally, in yeast, the mating-type locus is silenced by segregation to the nuclear periphery, where Sir proteins mediate its silencing (Cockell and Gasser, 1999). When a RNA pol II transcribed gene is inserted near the silenced locus, it too becomes transcriptionally inactive (Gottschling et al., 1990). These types of observation indicate that nuclear positioning is important for gene expression. They also provoke questions such as how is the position of chromosomes determined in the nucleus, what are the factors that determine chromosome structure and folding, and how is folding regulated?

How are nuclear bodies related to gene expression? Many nuclear bodies represent ‘hot spots’ of gene activity. The nucleolus is a site of high transcriptional activity of rRNA genes, and the structure is sensitive to inhibitors of transcription. Similarly, the Cajal Body is frequently associated with transcriptionally active histone loci, and it has been proposed to regulate histone gene expression (Frey and Matera, 1995; Gall et al., 1995; Schul et al., 1999). Similarly, a nuclear body termed the OPT domain (Oct1/PTF-transcription) that contains the transcription factors OCT, PTF, RNA pol II, TBP and SP1 forms small foci throughout the nucleus but also a large, prominent structure associated with the nucleolus (Pombo et al., 1998; Schul et al., 1998a). The latter form of the OPT domain is associated with the short arm of chromosome 6 and might form as a consequence of transcriptional activity of this region (Pombo et al., 1998).

NUCLEAR ARCHITECTURE AND DISEASE

Nuclear bodies have been implicated in various diseases. The increased size and number of nucleoli has been used as an indicator of transformed cells (Smetana et al., 1984). Similarly, the Cajal Body is found in increased numbers in tumor cells (Spector et al., 1992; Ochs et al., 1994). PML bodies gained prominence when it was found that they are absent from cells from patients with promyelocytic leukemia. Loss of the PML body is the consequence of a (t15;17) translocation giving rise to a fusion protein between the PML protein and the retinoic acid receptor α, which has altered localization properties (Ascoli and Maul, 1991; de The et al., 1991; Kakizuka et al., 1991). A strong correlation between the presence of a nuclear body and disease is the almost exclusive appearance of the
perinucleolar compartment (PNC) in transformed cells (Huang et al., 1997). The PNC forms as a ‘cap’ on nucleoli, and it probably associates with a particular chromosome, although the identity of the binding site is not clear. A further example of a connection between nuclear bodies and human disease is the so-called gems. Gems contain the SMN1 protein, which is the causative agent of spinal muscular atrophy (Liu and Dreyfuss, 1996). SMN1 mutations interfere with normal processing of snRNP particles and result in a block of pre-mRNA splicing (Fischer et al., 1997; Pellizzoni et al., 1998). Gems were initially described as distinct structures tightly associated with Cajal Bodies in HeLa cells, but subsequent investigation has revealed that gems in most cell lines are identical to Cajal Bodies (Carvalho et al., 1999). Although the function of nuclear bodies in these diseases is far from clear, their presence indicates that they might reflect aberrant nuclear processes, and identification of the function of these bodies might provide clues to disease mechanisms.

OUTLOOK

The existence of nuclear compartments is now firmly established. It is clear that they are distinct nuclear structures, although the function of most of them remains unknown. The study of nuclear architecture and nuclear compartments was initially purely descriptive, and this phase of study was crucial in establishing the identity of the compartments. Recently, investigations have moved into a molecular phase, and within the next few years the function of many compartments will hopefully be uncovered. Although there is considerable evidence to suggest a role for nuclear architecture in the coordination of transcription and pre-mRNA splicing, we do not know how other nuclear processes, such as RNA and protein degradation or DNA repair, are organized in vivo and whether particular nuclear structures are involved in those processes. One of the most important questions at present is to understand how nuclear compartments form. Do molecules that nucleate the formation of compartments exist or do compartments form as a consequence of mass action by association of proteins with each other, or both? A related and after many years still-unresolved issue is the existence of a ‘nuclear matrix’ or some sort of nuclear skeleton (see Pederson, 2000). Furthermore, we must define a link between expression of specific genes and nuclear bodies. Do nuclear bodies form at predetermined sites within the nucleus or the genome? If so, what are these sites? What is the function of nuclear bodies in regulation of those specific genes? Finally, we must determine whether positioning of genes within the nucleus is a general regulatory mechanism. The answers to these cell biological questions will contribute greatly to our understanding of gene expression in vivo.

I thank J. Cáceres, A. Lamond, I. Raska, S. Gunderson for sharing of unpublished data. I thank Javier Cáceres for critical comments on the manuscript.

REFERENCES


Nuclear architecture meets nuclear function 1847
translocation of the t(15;17) in human acute promyelocytic leukemia fuses RAR alpha with a novel putative transcription factor, PML. Cell 66, 663-674.


McCracken, S. et al. (1997a). 5′-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev. 11, 3306-3318.


