Nonrandom Gene Organization: Structural Arrangements of Specific Pre-mRNA Transcription and Splicing with SC-35 Domains

Yigong Xing, Carol V. Johnson, Phillip T. Moen, Jr., John A. McNeil, and Jeanne Bentley Lawrence

Department of Cell Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

Abstract. This work demonstrates a highly nonrandom distribution of specific genes relative to nuclear domains enriched in splicing factors and poly(A)$^+$ RNA, and provides evidence for the direct involvement of these in pre-mRNA metabolism. As investigated in hundreds of diploid fibroblasts, human collagen Iα1 and β-actin DNA/RNA showed a very high degree of spatial association with SC-35 domains, whereas three nontranscribed genes, myosin heavy chain, neurotensin, and albumin, showed no such preferential association. Collagen Iα1 RNA accumulates within the more central region of the domain, whereas β-actin RNA localizes at the periphery. A novel approach revealed that collagen RNA tracks are polarized, with the entire gene at one end, on the edge of the domain, and the RNA extending into the domain. Intron 26 is spliced within the RNA track at the domain periphery. Transcriptional inhibition studies show both the structure of the domain and the gene's relationship to it are not dependent upon the continued presence of accumulated collagen RNA, and that domains remaining after inhibition are not just storage sites. Results support a model reconciling light and electron microscopic observations which proposes that transcription of some specific genes occurs at the border of domains, which may also function in the assembly or distribution of RNA metabolic components. In contrast to the apparently random dispersal of total undefined hnRNA synthesis through interdomain space, transcription and splicing for some genes occurs preferentially at specific sites, and a high degree of individual pre-mRNA metabolism is compartmentalized with discrete SC-35 domains.

DESPITE the remarkable complexity of critical functions the nucleus performs, a simplified view of the extra-nucleolar nucleoplasm has persisted for many years. The idea that there exists some higher-level organization of the nucleoplasm that facilitates basic nuclear functions has been proposed for some time (see for example Comings, 1980; Blobel, 1985; Jackson, 1991; Lawrence et al., 1993), but the extent to which it exists is still largely unknown. It is known that the distributions of heterochromatin and satellite sequences are cell type specific (Manuelidis, 1984), that chromosomes occupy discrete nuclear "territories" (Cremer, 1982; Lichter, 1988; Pinkel et al., 1988), and that "chromosome position effects" can impact the expression of transgenes for largely unknown reasons (for example Al-Shawi et al., 1990). Relative to overall nuclear space, individual genes do not localize to precise coordinates, but distribute within preferred nuclear regions (Ward, W. S., J. A. McNeil, and J. B. Lawrence, manuscript submitted for publication; Lawrence et al., 1993). However, as explored in this work, greater order may become apparent when sequences are localized relative to defined internal reference points.

While some studies of nuclear structure investigate chromosomes or DNA within the nucleus, others focus on the localization of RNA or RNA metabolic components. These two broad aspects of nuclear structure are generally studied as separate entities and may in fact be independent of one another. Alternatively, these two compartments may be integrated, such that there is a nonrandom spatial arrangement of specific DNA sequences relative to nuclear regions devoted to the assembly, storage, transport, or activity of RNA metabolic components. This work investigates the possibility that there is nonrandom distribution of protein coding genes and/or their RNAs relative to discrete domains highly enriched in pre-mRNA splicing components and poly(A)$^+$ RNA.

It is now well established that many pre-mRNA splicing components (Spector 1990; Fu and Maniatis, 1990) as well as poly(A)$^+$ RNA (Carter et al., 1991; Visa et al., 1993) are highly concentrated at ~30–40 major nuclear sites or domains, in addition to a less concentrated signal dispersed throughout the nucleoplasm (reviewed in Lawrence et al., 1993; Spector, 1993; Carter, 1994; Fakan, 1994). Recent evidence indicates the presence of other nuclear matrix and SR proteins involved in pre-mRNA processing (Blencowe

Address all correspondence to Jeanne Bentley Lawrence, Dept. of Cell Biology, University of Massachusetts Medical Center, 55 Lake Ave. North, Worcester, MA 01655.

Drs. Xing and Johnson contributed equally to this work.

et al., 1994), poly(A)$^+$ RNA binding protein (Krause et al., 1994), and hyper-phosphorylated RNA polymerase II (Bregman et al., 1995) concentrated in these discrete domains. Earlier electron microscopic studies had described RNP-containing structures within the nucleoplasm (Bernhard, 1969; Fakan and Bernhard, 1971; Fakan and Puvion, 1980) termed interchromatin granule clusters (IGC)[1] and smaller, more amorphous perichromatin fibrils (PF). Although these may be heterogeneous categories, much evidence now indicates that the IGC are largely coincident with the prominent domains highly enriched in poly (A) RNA or spliceosome assembly factor, SC-35 (Spector et al., 1991; Visa et al., 1993). This work reflects a twofold interest in this nucleoplasmic compartment. First, irrespective of their precise function, discrete domains detected by anti-SC-35 immunofluorescence (Fu and Maniatis, 1990), provide internal reference points with respect to which the possibility of higher-level genomic organization may be explored. Second, the potential function of this nuclear compartment in the biogenesis of RNA processing components or in RNA processing itself is of obvious interest. The finding that poly(A)$^+$ RNA concentrated in these regions led us to raise the question of a role in pre-mRNA metabolism (Carter et al., 1991), an idea which would seem to be discounted by earlier observations that short pulses of ^{3}H-uridine label primarily the PF, rather than the IGC (reviewed in Fakan and Puvion, 1980; Fakan, 1994).

Biochemical studies indicate that the bulk of poly(A)$^+$ RNA gives rise to mRNA (for example, Scherrer et al., 1970; Salditt-Georgieff et al., 1981; reviewed in Brawerman, 1981; Lewin, 1990). However, recently XIST RNA provides a precedent for a poly(A)$^+$ RNA that functions in nuclear structure and appears to be long-lived in the nucleus (Brown et al., 1992; Clemson et al., in press). Hence, the finding of poly(A)$^+$ RNA in domains does not unequivocally show the presence of pre-mRNA (mRNA), and the question of whether mRNA metabolism is associated with this nuclear compartment remains a subject of interest and debate.

Studies which investigate broad classes of RNA, rely on drug inhibition, or introduce exogenous sequences in a foreign structural context yield readily detectable signals which can provide a general view, but may yield ambiguous conclusions. Furthermore, such studies cannot address the fundamental question of whether specific genes localize to and are transcribed in specific places. This requires an approach to identify directly the site of transcription and RNA processing for specific protein coding genes. Such methods showed the transcription and splicing site for fibronectin RNA to exhibit a high degree of association, mostly at the SC-35 domain periphery, leading us to suggest that there may be a nonrandom arrangement of some genetic loci relative to these regions (Lawrence et al., 1993; Xing et al., 1993). Although not interpreted in this way, the prior observation of c-fos RNA near snRNP "speckles" (Huang and Spector, 1991) is consistent with this hypothesis.

To investigate the relationship of other endogenous genes to this nuclear compartment, and to address key points concerning the function(s) of domains, we applied and further developed powerful molecular cytogenetic methods for precise cellular localization of specific genes, RNAs, and proteins (Langer-Safer et al., 1982; Lawrence et al., 1989; Carter et al., 1991; Xing et al., 1993). The resulting evidence demonstrates a nonrandom organization of specific sequences relative to these internal nuclear landmarks, and clearly demonstrates that the transcription and splicing of some pre-mRNAs is directly associated with them. Results further demonstrate that collagen RNA tracks are polar structures which extend beyond the gene and show a reproducible orientation with SC-35 domains.

Materials and Methods

Cell Culture and Fixation

Human diploid cultures either from lung (WI-38, CCL 75, American Type Culture Collection [ATCC]), Rockville, MD) or from foreskin (Detroit 551, CCL 110) were grown in Dulbecco's Modified Eagle's medium, high glucose (DME-high) supplemented with 10% FCS (GIBCO BRL, Gaithersburg, MD), and 10 μg/ml gentamicin (GIBCO BRL). Before fixation, cells were treated with cytoskeletal (CSK) buffer (100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 10 mM Pipes, pH 6.8; Fey et al., 1986) containing 0.5% Triton X-100 and 2 mM vanadyl ribonucleoside complex (VRC; GIBCO BRL), for 0.5–5 min on ice. Cells were then immediately fixed in 4% paraformaldehyde (PAF) in 1× PBS (pH 7.4) for 10 min and stored at 4°C in 70% EtOH until used.

Probes and Antibodies

A genomic probe for the human neurotensin gene, hNT14, which encompasses exons 1–3 was obtained from Dr. P. Dobner (University of Massachusetts Medical School, Worcester, MA). For detection of the α-cardiac myosin heavy chain gene, a genomic clone containing a 12.3-kb fragment encompassing 10.5 kb of the 5′ end of the gene was obtained from Dr. L. Leinwand (University of Colorado, Boulder, CO). A 20-kb genomic clone containing the human albumin gene, HSA-10, was obtained from ATCC. Specific localization of the β-actin locus was accomplished using a 14-kb genomic probe (p14T B-17) encompassing the 6-kb human β-actin gene (Leavitt et al., 1984; Ng et al., 1985). It was provided by Dr. Janet Stein (University of Massachusetts Medical School). Studies of the collagen Iα1 gene and/or RNA utilized several probes. A 40-kb cosmid clone containing the entire 18-kb human type Iα1 collagen gene (CG103) was obtained from ATCC. A 1.8-kb human collagen cDNA clone, pH F677, and a 24 kb genomic mouse collagen Iα1 gene probe encompassing and consisting almost entirely of the homologous mouse collagen gene were provided by Dr. William Strauss (Whitehead Institute for Biomedical Research, Cambridge, MA). Human collagen Iα1 intron 26-specific probes were generated by PCR amplification of intron 26 sequences only, and was provided by Dr. David Rowe (University of Connecticut, Farmington, CT). Hybridization of specific probes to metaphase chromosomes was used to confirm probe specificity for a defined genetic locus, as described elsewhere (Johnson et al., 1991, 1993).

A mouse monoclonal antibody against the spliceosome assembly factor SC-35, provided by Dr. T. Maniatis (Harvard University, Cambridge, MA) was used to delineate SC-35 domains by immunofluorescence. Anti-SC-35 antibody detects a non-snRNP spliceosome assembly factor (Fu and Maniatis, 1990) and may detect another structural protein which colocalizes to the same nuclear domains (X.-D. Fu, personal communication).

Fluorescent In Situ Hybridization and Immunofluorescence Staining

Probes were nick-translated using either biotin-16-dUTP or digoxigenin-11-dUTP as the substituted nucleotide. Briefly, for hybridization to DNA, cells were heat denatured in 70% formamide, 2× SSC at 70°C for 2 min, followed by hybridization with labeled probe in 50% formamide 2× SSC for 3 h to overnight at 37°C, as previously described (Johnson et al., 1991;

1. *Abbreviations used in this paper*: 3-D, three-dimensional; cMHC, cardiac myosin heavy chain; CSK, cytoskeletal; DAPI, 4,6-diamidine-2-phenylindole; IGC, interchromatin granule clusters; PAF, paraformaldehyde; PF, perichromatin fibrils; VRC, vanadyl ribonucleoside complex.

Xing et al., 1993). Hybridized probe was detected using either rhodamine or fluorescein conjugated anti-digoxigenin (Boehringer Mannheim Corp., Indianapolis, IN) or FITC-conjugated avidin (Boehringer Mannheim) (Johnson et al., 1991). In cases where DNA-specific hybridization was desired, removal of complementary RNA sequences was performed by either 0.07 N NaOH hydrolysis of RNA or incubation in RNAse A before hybridization. For RNA-specific hybridizations, no denaturation of cellular DNA was done, as previously described (Lawrence et al., 1989).

We previously reported the colocalization of the fibronectin gene and its RNA by targeting a nontranscribed 5′ flanking sequence as a probe for DNA, and an expressed sequence probe for detection of RNA (Xing et al., 1993). This method was supplemented by the development of a sequential hybridization protocol that makes it possible to detect DNA and RNA using the identical probe detected with two different fluorochromes. Collagen Iα1 RNA and the gene were detected using the 24-kb genomic collagen probe previously described. Nondenatured cells were hybridized with the digoxigenin-labeled probe to specifically detect collagen RNA, and the hybridized probe then detected using rhodamine anti-digoxigenin. The bound fluorescent antibody was then cross-linked by fixation in 4% PAF, rendering it impermeable to alkaline hydrolysis. The cells were then denatured in 0.07 N NaOH and then hybridized with the same sequence labeled with biotin to detect DNA specifically. Biotin was detected using FITC-avidin.

For all experiments, coverslips were rinsed in 4× SSC after the final detection incubation, when appropriate stained for 30 s in 0.1 μg/ml 4,6-diamidino-2-phenylindole (DAPI) in PBS and mounted on microscope slides in 90% glycerol/10% PBS containing 1 mg/ml phenylene diamine as an antioxidant. Slides were stored at −20°C until analysis.

Microscopy and Image Analysis

A Zeiss Axioplan microscope equipped with a triple-bandpass epifluorescence filter (Chroma Technology, Inc., Brattleboro, VT) was used for sample analysis by both standard and digital imaging microscopy. Digital images were captured with a Photometrics series 200 CCD camera (Photometrics, Ltd., Tucson, AZ) equipped with a custom-made color filter wheel, such that three colors can be captured and viewed simultaneously with no optical shift (J. McNeil, personal communication). Care was taken to score cells randomly, with the only selection criteria being clear quality of the hybridization signal and the SC-35 staining, which was evaluated first using separate filter sets. Placement of the signals relative to domains was then visualized through the standard microscope using a dual-band filter and scored by two investigators independently.

Scoring was done using a 100× 1.4 N.A. objective, which provides high-resolution and shallow depth of field, such that objects more than ~0.5 microns apart in the Z-axis will appear in a different focal plane (Carter et al., 1993). Hence, a significant degree of three-dimensional (3-D) information was discernible as cells were scored. As documented earlier by a thorough 3-D analysis using deconvolution and 3-D rendering, the SC-35 domains lie in a single plane in fibroblasts in the lower half of the nucleus (Carter et al., 1993), and hence are almost always all in focus in a single photograph. Optical sections of nuclei were obtained in 0.2 micron steps using a Z-axis stage controller, as detailed elsewhere (Carter et al., 1993).

As determined by computerized microfluorimetry of ten cells, the surface area occupied by SC-35 domains averaged 22%. However, correcting for the fact that domains have a planar arrangement and appear in the focal plane in only 1/3 to 1/2 of the Z-axis using standard microscopy (Carter et al., 1993; J. McNeil, unpublished data), the expected random distribution of spots which would associate with domains would be ~8–12%. This estimate is close to and consistent with the empirical observations presented here.

Results

The distribution of five single-copy genes and/or their nuclear RNAs was investigated using fluorescence in situ hybridization to nonsynchronized human diploid fibroblasts, using a combination of previously described (Lawrence et al., 1989; Johnson et al., 1991; Xing et al., 1993) and new techniques (Materials and Methods). Nuclear domains enriched in the spliceosome assembly protein SC-35 were detected using immunofluorescence with anti–SC-35 antibody (Fu and Maniatis, 1990). For clarity we will refer to these regions here as SC-35 domains to distinguish them from the more diffuse component seen for a subset of SC-35 signal. The SC-35 domains colocalize with discrete nuclear domains that have previously been termed transcript domains, which are regions enriched in poly(A)$^+$ RNA (Carter et al., 1991, 1993), and with the subset of Sm speckles that excludes coiled bodies (Lerner et al., 1981; Spector et al., 1983: Nyman et al., 1986; Fu and Maniatis, 1990). Generally 100–200 signals were examined by two investigators using 2-D analysis, and where necessary 3-D sectioning was done on a smaller cell sample. As previously described (Xing et al., 1993), "associated" denotes that the signal appears to contact the domain with no visible separation between the SC-35 domain and the gene/RNA, as opposed to "separate" in which a space between them is visible by fluorescence microscopy.

Localization of Five Genetic Loci Relative to SC-35 Domains

The first set of experiments addressed the question of whether there was a nonrandom organization of specific genetic loci relative to SC-35 domains. To investigate this we examined quantitatively the distribution of five different sequences, including two transcriptionally active and three inactive genes. Since DNA signals are not detected unless cellular DNA is denatured (Lawrence et al., 1989; Xing et al., 1993), the presence of signal in the absence of denaturation confirms the expression of nuclear RNA from the gene. In all cases below, the in situ analysis of RNA expression was consistent with the predicted expression pattern of that gene in human fibroblasts. As shown in other work from our lab and confirmed below, specific pre-mRNA and gene signals overlap, and were therefore only separately visualized where necessary.

Albumin. The albumin gene, expressed only in liver cells (see Hawkins and Dugaiczyk, 1982), is transcriptionally inactive in human diploid fibroblasts. Hybridization generally produced two small fluorescent spots in samples where cellular DNA was denatured, in keeping with single-copy gene detection. As illustrated in Fig. 1 *F*, and summarized in Table I, analysis of 160 signals for the spatial relationship of this gene to SC-35 domains showed 86% of genes clearly separate from domains. The remaining signals (14%) showed no visible separation by fluorescence microscopy, and hence were scored as associated. Interestingly, in most cells the gene localized near the nuclear periphery (Fig. 1 *F*), in a region devoid of poly(A)$^+$ RNA/splicing factor-rich domains in human fibroblasts (Carter et al., 1991, 1993), and an area in which heterochromatin is known to concentrate in many cell types (Lentz, 1971).

In human fibroblasts, poly(A)$^+$ RNA/SC-35 domains occupy ~5% of nuclear volume, as measured in deconvoluted and 3-D restored digital images (Carter et al., 1993). However, as considered in Materials and Methods, based on both average surface area and Z-axis distribution of domains using the optics applied here, a spot randomly distributed in the nucleus would appear associated with SC-35 domains on the order of 8–12% of the time (see Materials and Methods). We conclude, therefore, that the distribution of the albumin gene with respect to SC-35 domains is

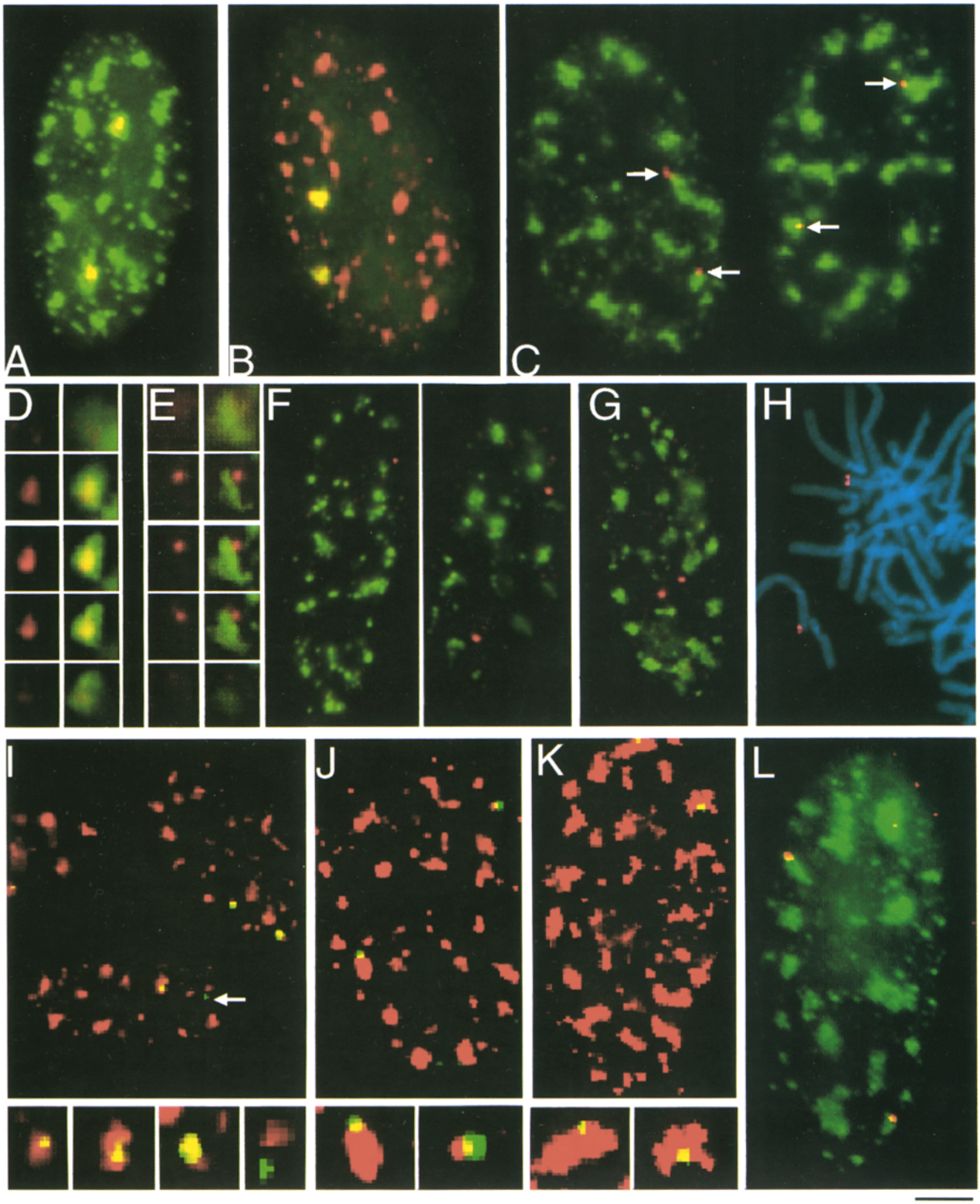

Figure 1. Localization of genes and/or their transcripts relative to SC-35 domains. Genes were probed in normal human diploid fibroblasts using fluorescence in situ hybridization in conjunction with immunofluorescence staining using an antibody to SC-35. (*A*) Collagen Iα1 transcripts (*red*), visualized using a full length genomic probe, colocalize with SC-35 domains (*green*, overlap of red and green appears yellow) (2,500×). (*B*) Collagen Iα1 transcripts (*green*) hybridized with a cDNA probe (2,500×). (*C*) Collagen Iα1 gene (*red*) localizes at the edge of SC-35 domains (*green*) (2,500×). (*D* and *E*) Optical sections along the Z-axis of a single SC-35 domain (*green*) illustrating the location of (*D*) collagen Iα1 transcripts (*red*) that are coincident in all planes with the domain, and (*E*) the collagen Iα1

Table I. Localization of Genes and/or RNAs Relative to SC-35 Domains

	No. scored	% Associated	% Separate	Comments
Albumin	160	14	86	Nuclear periphery
c-Myosin Heavy Chain	217	22	78	Near nucleolus
Neurotensin	154	11	89	Nuclear periphery
Actin Total	290	89	11	
DNA or DNA/RNA	155	94	6	Edge of SC-35 domain
RNA	135	83	17	Edge of SC-35 domain
Collagen Total	239	99	1	
RNA	135	100	0	Within domain interior
DNA	104	98	2	Edge of domain

in keeping with a random distribution or a preferential separation from domains that falls within the limitations of our light microscopic technique (Materials and Methods).

Cardiac MHC. The human cardiac myosin heavy chain (cMHC) gene is expressed exclusively in cardiac or skeletal muscle and thus is inactive in fibroblasts (Lompre et al., 1984; Saez et al., 1987). Unlike the more peripheral albumin gene, the MHC gene distributed more internally and frequently near the nucleolus, consistent with its chromosomal location close to the ribosomal RNA genes on chromosome 14. As illustrated in Fig. 1 *G*, despite the fact that this gene is in a nuclear region where there is a preponderance of poly(A)$^+$ RNA/SC-35–rich domains in human fibroblasts (Carter et al., 1991, 1993), the vast majority (78%) of cMHC genes were not associated with any SC-35 domain. Under the optics used, 22% appeared to be in contact with a domain, and were hence scored as associated (Table I). While this frequency is slightly elevated compared to albumin, it does not differ clearly from random, especially given the localization of the gene in a region enriched in domains. Localization of cMHC near the nucleolus is likely related to its linkage to the rRNA genes, but is also consistent with a localization with extra-nucleolar heterochromatin.

Neurotensin. The neurotensin gene, transcribed only in certain types of neural cells (Kislauskis et al., 1988), is transcriptionally inactive in human diploid fibroblasts. Analysis of over 100 cells showed 89% of neurotensin gene loci visibly separate from the SC-35 nuclear domains (Table I), with only 11% appearing to contact domains. Most neurotensin signals were localized to the extreme periphery of the nucleus, close to the nuclear envelope. Hence the inactive neurotensin gene is not preferentially associated with SC-35 domains, and may reside in or near peripheral heterochromatin.

β-actin. The location of the human β-actin gene and RNA, which is expressed in fibroblasts, was examined using a 14-kb genomic probe that specifically detects the locus encompassing the 6-kb β-actin gene as confirmed by hybridization to metaphase chromosomes (Fig. 1 *H*). Nuclear hybridization to the β-actin gene produced two small spots, and hybridization to RNA produced two small foci only slightly larger than the DNA. Detection of DNA and RNA simultaneously showed the same number of signals as when RNA was detected alone, confirming that the DNA and RNA signals overlapped. Initial analysis of 155 β-actin gene or gene and RNA signals showed a very high degree of association (94%). Subsequent hybridizations detecting β-actin RNA alone (non-denatured samples) confirmed that the gene was transcriptionally active and that the foci of nuclear RNA showed a high degree of association. Of 135 RNA signals scored, 83% were associated. In total 290 actin signals were scored by three investigators, with an average of 89% association with SC-35 domains (Table I).

For several reasons the details of the spatial association are important (see Discussion). All associated signals were categorized as either within, meaning the signal overlapped the interior or central SC-35 domain, or border, indicating it localized at the extreme periphery or edge of the domain (Xing et al., 1993). As illustrated in Fig. 1, *I–L*, ~94% of the associated actin gene or RNA signals localized at the border, directly adjacent or abutting the boundary of the SC-35 rich region, whereas only 6% appeared inside the domains (Table I). Most signals associated with a large domain, with ~16% associated with a small but discrete SC-35 domain (see Fig. 1, *I* and *J*, *inset*). The gene or RNA signals from homologous chromosomes generally did not associate with the same domain but with two different domains, and while the two alleles generally were both associated, in

gene (*red*) that localizes to the edge of the domain (5,000×). (*F*) Hybridization of the albumin gene (*red*) shows that it is not associated with SC-35 domains (*green*) and frequently appears at the nuclear periphery. (*G*) MHC gene (*red*) does not associate with the SC-35 domains (*green*) and frequently localizes to the inner regions of the nucleus near the nucleolus (2,500×). (*H*) Metaphase chromosomes hybridized with a genomic β-actin probe (*red*) demonstrating this probe detects a single locus. The location on Chr. 7p is consistent with the reported location of the functional gene, although using fluorescence mapping we observe the gene in a slightly more telomeric position. (2,500×). (*I–K*) Foci of actin RNA (*green*) occur most frequently at the edge of the SC-35 domains (*red*), but occasionally within smaller domains (overlap of red and green appears yellow). The lower insets show detail of overlap (5,000×). (*I*) three nuclei with four of five signals associated (arrow points to the nonassociated signal) (1,250×). (*J–K*) images were processed to enhance contrast (2,500×). (*L*) Actin RNA/DNA (*red*) associated with SC-35 domains (*green*) in a nonenhanced image (2,500×). Bar: (*A–C*, *F–H*, *J–L*) 2.5 μm; (*I*) 5 μm; (*D*, *E*, insets *I–K*) 1.25 μm.

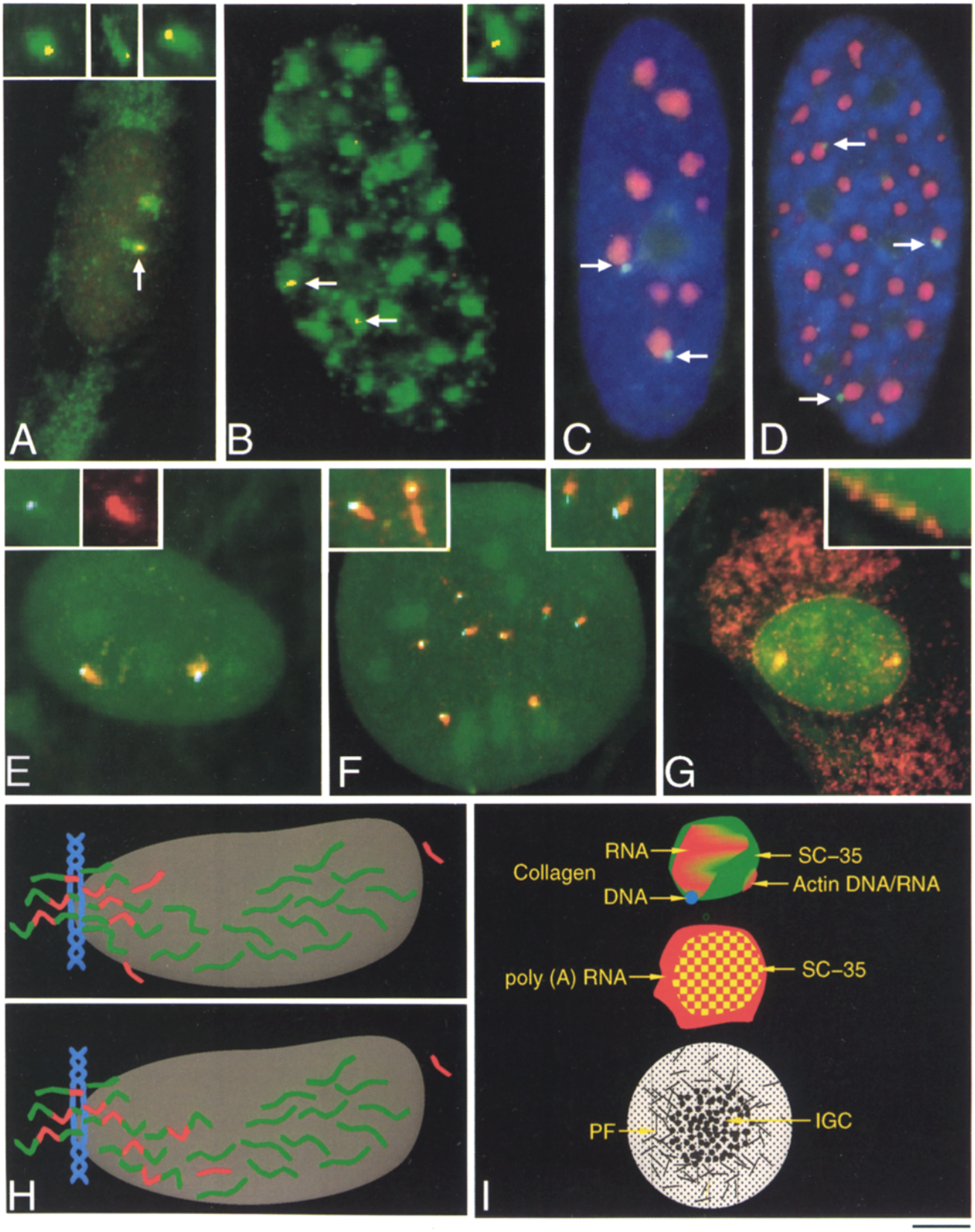

Figure 2. (*A*) Collagen Iα1 intron 26 probe (*red*) and collagen Iα1 cDNA (*green*, overlap of red and green appears yellow) simultaneously hybridized in human fibroblasts (2,500×). Insets show additional examples of intron sequences (*red*) and cDNA sequences (*green*) from other cells (4,000×). Note that intron sequences are generally localized to the edge of the track. (*B*) Colocalization of collagen Iα1 intron 26 sequences (*red*) and SC-35 domains (*green*) (2,500×). Inset shows enlarged example from a different cell, illustrating intron signal at the inner periphery of the domain (4,000×). (*C* and *D*) Collagen Iα1 gene (*blue-green*) remains associated with SC-35

occasional cells one allele associated and the other did not (see bottom cell, Fig. 1 *I*).

Collagen Iα1. The distribution of the transcriptionally active collagen type Iα1 gene/RNA was investigated using a genomic probe encompassing the transcription unit for this 18 kb gene (Barsh et al., 1984), which detected two specific DNA or concentrated RNA signals in most nuclei. Simultaneous hybridization to DNA and RNA, versus DNA or RNA alone, confirmed that the gene signal was a small spot coincident with a larger accumulation of RNA. Although not highly elongated tracks, the RNA accumulations were substantially larger than the small foci observed for actin RNA, and were track-like in that they often had some longitudinal axis. As illustrated in Fig. 1, *A* and *B* and quantitated in Table I, essentially all collagen RNA signals spatially associated with large distinct SC-35 domains. In contrast to actin RNA, collagen RNA consistently appeared to occupy the inner regions of the domains.

Using high resolution optics which allow discrimination of objects more than 0.5 microns apart in the Z-axis, the collagen RNA signal consistently appeared to be within the SC-35 domain. This conclusion was confirmed by optical sectioning of several cells by digital imaging microscopy (Fig. 1 *D*) that showed the collagen RNA occupied the same focal planes as the SC-35 enriched region, and therefore overlapped it in 3-D space. The domains with which collagen RNA associated were among the largest within the nucleus (~1–3 microns; Carter et al., 1993). Some collagen RNA foci or tracks had a similar size and shape to their associated domains (compare Fig. 1 *B*), although sometimes (in some experiments as high as 80%) the collagen nuclear RNA appeared to concentrate in one part of the SC-35 domain (Fig. 1, *A* and *D*).

It is important that sequence distributions be determined in many cells (see Discussion). However, even in nonsynchronized cultures, the gene distributions were reproducible from experiment to experiment and distinct distributions were obvious from just brief viewing blind through the microscope. Irrespective of the function of domains, these results demonstrate a nonrandom organization of genomic loci relative to splicing factor rich domains. While we do not suggest that all transcriptionally active genes or pre-mRNAs associate with domains (see Discussion), these and other results support that the association of the two active genes with SC-35 domains, as contrasted with the three inactive genes, is not coincidental but related to their production or maturation of pre-mRNA.

Structural Relationship of Collegen Iα1 Transcription and Splicing with SC-35 Domains and RNA Tracks

The collagen Iα1 gene is one of the cell's most highly active genes, accounting for approximately 4% of total mRNA in fibroblasts and producing an RNA with 51 introns (Genovese et al., 1989). From some perspectives it might be considered surprising to find pre-mRNA from such a highly active gene associated with SC-35 domains (see Discussion). The collagen Iα1 gene and RNA were analyzed in much greater detail to address several key points as to the relationship of structure and function between RNA foci or tracks and the SC-35 domains.

Transcripton at the Border of SC-35 Domains

To determine the precise spatial relationship of the collagen gene relative to the SC-35 domain, hybridizations were performed under conditions which remove RNA and denature cellular DNA, using the same full-length genomic probe. As shown directly in Fig. 1 *C*, the collagen gene signal was associated with the prominent SC-35 domains in ~99% of cells. But unlike the RNA that localized within the domain, the DNA signals localized as a spot at the domain border, with no visible separation between the gene and the edge of the SC-35 region. It was not possible to discern unequivocally on which side of the border the gene was positioned, since it sometimes appeared on the outer border, but frequently appeared just within the border. Interestingly, when the domain with which the gene was associated was elongated, in most cases the gene positioned at one end of the linear axis (see Fig. 1 *C*).

A curious and interesting aspect of our results is that in the vast majority of cells the collagen gene appeared adjacent to the domain in the X-Y plane, such that the gene was in focus with the boundary of the domain. The gene only rarely appeared to be above or below the domain. This impression was confirmed by optical sectioning of nu-

---

domains (*red*) after three hours inhibition with actinomycin D in (*C*) diploid and (*D*) tetraploid (three signals visible) fibroblasts (2,500×). (*E* and *F*) Collagen Iα1 gene (*blue-green*) with transcripts (*red*) colocalized by sequential hybridization with the same genomic probe encompassing the collagen 1A1 gene in primary mouse fibroblasts (2,500×). Even in the polyploid cell (*F*) where RNA accumulations appear as relatively small foci, there are several examples of gene/RNA polarity. Insets show enlarged examples illustrating the RNA signal extending beyond the gene signal. (*G*) Hybridization to cytoplasmic and nuclear collagen Iα1 RNA (*red*); and collagen Iα1 gene (*green*, overlap of red and green appears yellow) in human fibroblast (1,250×). Inset shows magnification of RNA associated with the nuclear envelope (see text). (*H*) Schematic diagram summarizing spatial arrangements of collagen Iα1 gene transcription and RNA splicing relative to domains identified by immunofluorescence to splicing assembly factor, SC-35. Transcription localizes at the boundary of the SC-35–rich domain. Splicing, as studied for intron 26, occurs at the domain periphery, and spliced RNA enters the central domain. Since gene and intron RNA are both peripheral, splicing may be either co-transcriptional (*top*) or posttranscriptional (*bottom*). (*I*) Model for the subcompartmentalization of individual domains, as it relates to earlier light and electron microscopic observations, proposing that transcription of some specific pre-mRNAs occurs at the border of discrete concentrations of SC-35. The SC-35 core may have additional functions such as in the assembly or distribution of splicing components or mRNA transport. Upper image summarizes results of this work. Collagen and β-actin RNA are shown associated with the same domain for purposes of illustration only. Results illustrate their differential spatial arrangement relative to the SC-35 domains, and the position of gene and RNA. Middle image (from Carter et al., 1993), shows SC-35 as an inner core of a slightly larger domain defined by hybridization to poly(A)$^+$ RNA. Lower image illustrates the potential correlation to electron microscopic structures, with IGC corresponding to the SC-35 core, and perichromatin fibrils to the rim where transcription occurs, for some specific genes. Bar: (*A–F*) 2.5 μm; (*G*) 5 μm; (*insets A, B, E–G*) 1.67 μm.

clei, as illustrated in Fig. 1 *E* which showed the gene next to the domain, as opposed to the RNA that lay within the domain (Fig. 1 *D*). These results clearly demonstrate that the transcription site of collagen is not only associated with SC-35 domains but displays a reproducible spatial position at the boundary or periphery of the SC-35 rich region. This point is pivotal to building a model of the structure and function of SC-35 domains (Discussion and Fig. 2 *I*).

Reproducible Polarity and Orientation of Collagen RNA Tracks

Since the collagen gene signal positions at the edge of the domain and the collagen RNA concentrates within it, this strongly implies a polarity of the gene position relative to the RNA accumulation. The question of whether the RNA accumulation extends beyond the dimensions of the gene is important to understanding whether it comprises strictly nascent transcripts or a posttranscriptional accumulation of RNA. The ~18-kb collagen gene provides an ideal model to examine this, since overwhelming evidence indicates that sequences of this relatively small size are below the resolution of the light microscope, and produce just a single round spot of signal in unfractionated interphase nuclei (Lawrence et al., 1988, 1990; Trask et al., 1991; Wiegant et al., 1992; Gerdes et al., 1994). As expected, in numerous experiments the genomic probe encompassing the entire transcribed region of collagen Iα1 consistently appeared as a small round spot (Figs. 1 *C* and 2, *E* and *F*). That the gene was not packaged as a large extended loop was further confirmed by experiments in which hybridization of a 5.6-kb sequence near the middle of the gene produced overlapping signal with hybridization to the rest of the sequences from the full-length genomic probe (not shown).

To compare the DNA and RNA signals by the most direct means possible, a novel procedure involving sequential hybridization was developed which allowed simultaneous visualization of DNA and RNA in different colors using the same sequence as a probe (see Materials and Methods). A 24-kb mouse genomic probe encompassing almost exclusively the full transcription unit of the collagen gene (~20 kb) was used for this purpose in mouse fibroblasts. Results were striking. As shown in Fig. 2, *E* and *F*, this revealed that the collagen gene was clearly positioned at or very near one end of the RNA "track." This was the case in 124 out of 145 cases examined. As illustrated in the polyploid cell shown in Fig. 2 *F*, even in a cell where the RNA signals are just small foci with no apparent linear axis, a prominent accumulation of RNA is most frequently observed to the side of the gene signal, producing an appearance of a comet consisting of a gene and a trailing track of RNA.

This clear polarity of the gene relative to the RNA track is consistent with and confirms the above observations which independently showed the collagen gene at the boundary of an SC-35 domain and the RNA within the domain. These experiments support that collagen RNA tracks or foci are formed by an accumulation of RNA which extends from and beyond the dimensions of the gene; hence even small foci may be RNA accumulations at some posttranscriptional step, rather than solely nascent transcripts directly on the gene. These RNA tracks extend vectorially from the gene into the SC-35 domain.

Collagen Iα1 RNA Splicing in the Domain Periphery

Why would there be an accumulation of RNA adjacent to the gene? Does the RNA track contain both introns and exons, spliced mRNA, or might this be an accumulation of excised introns? To address these questions and determine directly whether RNA splicing localized with the domain, a series of experiments was done using a combination of genomic, cDNA and intron probes. Sequences detected with a cDNA probe showed high concentrations throughout much of the SC-35 domain, similar to RNA detected with the genomic probe (Fig. 1 *B*). Direct comparison of the signal from the cDNA and full-length genomic probes in two colors showed that these completely overlap (not shown). This demonstrates that no appreciable portion of the collagen Iα1 RNA tracks consisted of accumulated introns alone, ruling out the possibility that RNA within the domain consisted of excised introns.

Analysis was next done using an intron-only probe which detected sequences from intron 26 (143 bp). As illustrated in Fig. 2 *B*, colocalization of intron sequences relative to SC-35 demonstrated that, in contrast to results with the cDNA and genomic probes, the intron was detected as a small focal signal in the periphery of the domain, rather than broadly throughout the central domain. This peripheral concentration was observed in the vast majority of cells; only infrequently was there a weaker more dispersed signal within the domain. Comparison of intron signal directly with cDNA showed that intron 26 sequences occupied only a portion of the track defined by the cDNA (Fig. 2 *A*), with intron consistently concentrated towards one end.

The absence of intron from a portion of the cDNA track indicates that splicing of this intron is occurring within the track. Since collagen RNA tracks associate with SC-35 domains, both splicing and transcription of collagen RNA occurs in intimate association with the domain, generally near the periphery, as schematically illustrated in Fig. 2 *H*. Two versions of the model are presented, illustrating cotranscriptional splicing of introns restricted to nascent transcripts on the gene (*top*) or posttranscriptional splicing at the periphery of the domain (*bottom*). Given that both gene and intron RNA localize near the domain periphery, this indirectly indicates that splicing of intron 26 is spatially proximal to the gene, if not at the gene itself. Cotranscriptional splicing is known to occur for some RNAs (Beyer and Osheim, 1988); however because the intron may extend slightly further into the domain, and because it is possible that the RNA's entry into a splicing factor-rich region is related to its splicing, from our data either version of the model is plausible. In both versions, the bulk of the RNA throughout the domain has been spliced, specifically with respect to intron 26.

RNA Tracks Extend into the Domain but Not to the Nuclear Envelope

The polar relationship between the collagen gene and its RNA accumulation suggests that there is movement of the RNA and gene with respect to one another, such that

the RNA extends from the gene into the domain in a vectorial manner. Consistent with the observation that SC-35 domains do not position adjacent to the nuclear envelope (Carter et al., 1993), collagen RNA tracks, which occupy SC-35 domains, do not appear to extend to the nuclear envelope. As shown in Fig. 2 *G*, dispersed collagen RNA signal is sometimes observed which may represent nuclear RNA in transport towards the nuclear envelope. However, no one clear track is seen to the envelope, suggesting that this step may occur very quickly and/or the RNA could disperse in multiple directions, consistent with previous results for endogenous fibronectin RNA (Xing and Lawrence, 1993; Xing et al., 1993). Fig. 2 *G* shows a ring of collagen RNA studding and encircling the nuclear envelope (*inset*); to the extent that this may reflect RNA as, or shortly after, it exits the nucleus, this would suggest that collagen RNA exits many or all of the nuclear pores and not just a specific region. While it was occasionally observed that there was more RNA in the cytoplasm nearest the nuclear RNA track, this was not commonly observed, and the distribution of the RNA in the cytoplasm was most consistent with random association of the mRNA with the endoplasmic reticulum (Fig. 2 *G*). As viewed in two dimensions, no specific orientation of the gene and RNA tracks relative to the nuclear envelope was noted, although more analysis would be required to demonstrate this rigorously.

In summary our results indicate that visible nuclear collagen RNA tracks lead into SC-35 domains, but do not directly extend to the nuclear envelope.

The Collagen Iα1 Gene Remains Associated with a Prominent SC-35 Domain after Transcriptional Inhibition

Finally we investigated whether the collagen gene remained associated with SC-35 domains in cells treated to inhibit transcription. As considered under discussion, such studies have produced variable and enigmatic effects on domain structure (Lawrence et al., 1993). Several studies report that IGC become fewer and larger and correlate these with the round prominent SC-35 domains observed after inhibition (Davis et al., 1993; Visa et al., 1993; Huang et al., 1994). In one case these studies have been interpreted to support that the large prominent SC-35 domains or IGC which exist in uninhibited cells are unassociated with pre-mRNA metabolism, but are sites devoted to storage of splicing factors (Huang et al., 1994; reviewed in Mattaj, 1994). To provide a critical test of this interpretation of inhibition studies, we colocalized the collagen gene and SC-35 after transcriptional inhibition. This experiment simultaneously addresses two other key questions: (*a*) Is the domain of SC-35 an accumulation of splicing factors dependent on the continued presence of accumulated collagen RNA? (*b*) Is the association of the collagen gene with SC-35 domains dependent on its continued transcription?

As shown in Fig. 2, *C* and *D*, transcriptional inhibition for 3 h resulted in prominent round SC-35 domains in many though not all cells. Under these conditions, collagen RNA tracks are generally undetectable or extremely reduced (Clemson et al., in press; and C. Clemson, data not shown). Hybridization to the collagen gene in denatured samples revealed that the gene remained precisely positioned at the edge of these round large structures, as shown in Fig. 2, *C* and *D*. In the vast majority of cells the collagen genes remained directly abutting the SC-35 domain. Neither the presence of the domain nor the gene's relationship to it is dependent upon the continued presence of accumulated collagen RNA. These results also show that inhibition studies cannot be reliably interpreted to show that domains which remain or rearrange after inhibition are solely storage sites, as our results demonstrate that they correspond to the transcription and splicing site of one of the cell's most active genes.

Discussion

This work details many new observations which substantially enhance an emerging view of a compartmentalized and structured nucleoplasm. We demonstrate that individual genes display distinct and nonrandom distributions relative to domains enriched in SC-35 and poly(A)$^+$ RNA, with two transcribed sequences specifically associating with discrete domains. Three transcriptionally inactive sequences do not. β-actin RNA associated at the periphery of the domain, whereas collagen Iα1 RNA accumulates within the central region of the domain. The focus of collagen RNA contains exon sequences, is not an accumulation of excised introns, and has undergone splicing, as shown for intron 26. Although collagen RNA accumulates within the SC-35 domain, gene transcription occurs just at its border. RNA splicing of intron 26 occurs at the periphery of the domain, likely proximal to the gene. Two-color DNA/RNA hybridization with a full-length genomic probe shows that collagen RNA tracks or foci are polar accumulations of RNA which extend beyond the dimensions of the gene, suggesting a posttranscriptional RNA accumulation beyond the nascent transcript "tree." This accumulation extends vectorially from the gene into the domain, and therefore may involve transport into the domain, but the track does not appear to extend to the nuclear envelope. The results presented also clearly show that neither the presence of the SC-35 domain nor the gene's association with it is dependent on the continued transcription or accumulation of collagen RNA. The prominent domains that remain after inhibition with actinomycin D correspond to the transcription and processing site of one of the cell's most highly active genes, and therefore cannot reliably be interpreted as solely storage sites unassociated with pre-mRNA metabolism.

Although gene position in overall nuclear space is more variable (Ward, W. S., J. A. McNeil, and J. B. Lawrence, manuscript submitted for publication; Lawrence et al., 1993), when viewed relative to SC-35 domains a statistically nonrandom distribution of specific sequences is clearly demonstrated. Irrespective of the relationship to function, these clear differences in gene distribution point to a previously unappreciated aspect of interphase genome packaging and a structural integration of two distinct nuclear compartments. We do not suggest that all active genes are associated with SC-35 domains (see below); however the two examined here, collagen and β-actin clearly show a specific preferential association. In contrast, three transcriptionally inactive genes (cMHC, neurotensin, and

albumin) showed a much lower association rate, consistent with a random distribution, possibly reflecting a preferential avoidance of the domains.

Transcription and Splicing of Some pre-mRNAs Is Directly Associated with SC-35 Domains

While it has been thought that prominent domains, which correspond largely to IGC, were not involved in short lived pre-mRNA metabolism (Fakan and Bernhard, 1971; Fakan and Puvion, 1980; Spector, 1993; Mattaj, 1994; reviewed in Fakan, 1994), the finding of poly(A)$^+$ RNA (Carter et al., 1991, 1993; Visa et al., 1993) and microinjected intron-containing globin RNA (Wang et al., 1991) in all of these domains justified a reevaluation of this view. In this work, we show that pre-mRNA metabolism for two of the most active genes in fibrolasts, collagen Iα1 and β-actin, is directly associated with SC-35 domains. These results, together with the preferential association of fibronectin transcription/splicing with domains (Xing et al., 1993), and the proximity of induced c-fos nuclear RNA to domains (Huang and Spector, 1991) argues strongly that there is pre-mRNA metabolism associated with these regions, with transcription at the periphery. Localization to the central domain is not unique to COL1α1 RNA, as thus far two other pre-mRNAs have been found which show similar localization (J. Coleman, C. Johnson, P. Moen, and J. Lawrence, unpublished data).

While we believe the association is related to function of these specific genes, we do not suggest this to be a property of all transcriptionally active DNA, or even of all intron-containing protein-coding RNAs (Lawrence et al., 1993). Of 10 active genes examined thus far in our lab, 7 have been found associated and 3 have not; in contrast, none of several inactive genes have yet been found associated (this work and Carter et al., 1991; Xing et al., 1993; Clemson et al., in press; J. Coleman and P. Moen, unpublished data). A cell cycle-regulated histone gene showed a mixed result (Xing, 1993), which is under investigation in synchronized cells. Further support that association is related to expression is provided by work showing that two genes which associate with domains in expressing muscle cells do not associate with domains in nonexpressing myoblasts (P. Moen, unpublished data).

Recently, a report focusing on Adenovirus localization included what appears to be a cursory analysis of the β-actin RNA transcription site relative to SC-35 domains in five cells (Zhang et al., 1994). Reporting 4 of 10 signals associated, the authors conclude that the distribution is random, using a random expected frequency that apparently does not correct for Z-axis considerations. Based on our analysis of hundreds of signals from the identical gene/RNA relative to the same structures detected with the same SC-35 antibody (Xing, 1993 and this work), the very small sample size of Zhang et al. (1994) is insufficient to draw a conclusion, and may not be representative. Most importantly, the general conclusion these authors draw from their data, that there is no compartmentalization of gene transcription and splicing with nuclear domains or speckles, is not only contradicted directly by results presented here but argued against by a variety of earlier observations (Walton et al., 1989; Carter et al., 1991; Huang and Spector, 1991; Wang et al., 1991; Xing, 1993; Xing et al., 1993; Jiménez-García and Spector, 1993).

Total Uridine Incorporation: Transcription of What?

The tenet or expectation that prominent SC-35 domains would not be associated with short-lived pre-mRNA or mRNA stems largely from uridine incorporation studies, which clearly show that general transcription occurs primarily in the interdomain (or inter-IGC) space (Fakan and Puvion, 1980; Wansink et al., 1993). However, a key question is, what classes of RNA does this transcription represent (Moen et al., 1995)? Uridine labeling has in some cases been equated with pre-mRNA for protein coding genes (Mattaj, 1994). However, there are many caveats to this assumption, as considered previously (Carter et al., 1993; Visa et al., 1993; Moen et al., 1995). For instance, most uridine label (possibly as much as 75–90%) will not be in mRNA precursors, but in other classes of RNA, especially noncoding hnRNAs of unknown function, and in intron sequences that may be rapidly spliced and disperse from the pre-mRNA molecule (for example, Salditt-Georgieff et al., 1981; Lewin, 1975; Brawerman, 1981; Beyer and Osheim, 1988; Lewin, 1990; Xing et al., 1993). Some labeling of IGC and SC-35 domains does occur (Fakan and Bernhard, 1971; Raska et al., 1991; Jackson et al., 1993; Wansink et al., 1993), and studies using Br-UTP are impacted by the fact that the Br-UTP inhibits splicing (Wansink, 1994). Most importantly, despite the impression of random transcription, these studies cannot address the question of whether transcription and splicing of specific genes is randomly localized. Our results provide direct evidence that it is not.

Two exceptions to the generalization that poly(A)$^+$ RNAs encode mRNAs have recently been described; Omega-n in Drosophila (Hogan et al., 1994) and XIST RNA, which associates with the inactive X chromosome of mammals (Brown et al., 1992; Clemson et al., in press). XIST RNA represents a precedent for an apparently long-lived structural RNA that functions in the nucleus; other nuclear RNAs likely contribute significantly to the label observed with uridine incorporation. Although most are not poly-adenylated (Salditt-Georgieff et al., 1981; reviewed in Lewin, 1990), some long-lived nuclear RNAs could potentially contribute to the poly(A)$^+$ RNA distributions observed.

SC-35 Domains Are Not Just Storage Sites: Inhibition of What?

Differing effects of transcriptional inhibition have been observed which result in complex and variable rearrangements or loss of nuclear poly(A)$^+$ RNA (Lawrence et al., 1993; Moen et al., 1995). Contradictory conclusions have been reported which support that poly(A)$^+$ RNA in discrete domains (or IGC) is short-lived pre-mRNA (Visa et al., 1993) or, alternatively, argue that it is long-lived structural RNA (Huang et al., 1994). In many cases, domains break down or dissipate. Even where there clearly is a retention and redistribution of nuclear poly(A)$^+$ RNA after inhibition (Lawrence et al., 1993; Huang et al., 1994; Moen et al., 1995), interpretation of such results is compromised by the secondary effects on global nuclear function (reviewed in Brasch, 1990), such as inhibition of RNA trans-

port and message-specific increases in RNA stability and poly(A)$^+$ tail length (for example see Herman and Penman, 1977; Herman et al., 1976, Brawerman, 1987; Hogan et al., 1994). The recent observation that XIST RNA is largely retained after inhibition whereas collagen RNA is not lends some credence to the idea that two classes of poly(A)$^+$ RNA with different stabilities may be detected (Xing, 1993; Clemson et al., in press). However XIST RNA does not localize to poly(A)$^+$ RNA-rich domains (Clemson et al., in press). It is an intriguing possibility that there may be some structural poly(A)$^+$ RNAs in domains and potentially elsewhere in the nucleus (Lawrence et al., 1993; Huang et al., 1994), but this would by no means preclude the presence of pre-mRNA in those same regions.

An important result presented here clearly indicates that the subset of domains which remain after transcriptional inhibition cannot simply be interpreted solely as storage sites or as not directly involved in pre-mRNA metabolism, as has been suggested (Huang et al., 1994; O'Keefe et al., 1994; Mattaj 1994). We demonstrate here that domains remaining in actinomycin inhibited cells correspond to the transcription and splicing site of one of the cell's most highly active genes.

A Model for Domain Substructure: Transcription at SC-35 Domain Borders

A key point of our results is that transcription occurs at or near the border of SC-35 domains, as suggested by studies quantitatively demonstrating the peripheral position of the fibronectin gene/RNA (Xing et al., 1993). Subcompartmentalization of light microscopic domains was suggested by Carter et al. (1993) (Fig. 2 *I*). We propose a model showing a high-level of transcription and splicing of some specific genes in the outer rim of the light microscopic domain, abutting the SC-35 core (Fig. 2 *I*). Although indirect, several results support that this SC-35 core corresponds to IGC; particularly our finding that the gene remains associated with a domain in inhibited cells, since such domains have been directly correlated to IGC (Davis et al., 1993; Visa et al., 1993; Huang et al., 1994). The subcompartmentalization model (Fig. 2 *I*) potentially reconciles results from our lab and others with earlier electron microscopic reports, some of which showed heavy uridine labeling at the borders of IGC, as well as the peripheral association of perichromatin fibrils (Fakan and Puvion, 1980; Spector, 1993; Hendzel and Bazett-Jones, 1995).

Several points indicate that SC-35 domains in general cannot simply be explained by the presence of splicing components wherever there are individual intron-containing RNAs. (*a*) Some specific RNAs (e.g., β-actin and fibronectin) clearly do not overlap most of the domain, but are smaller and peripheral to it. (*b*) While the association of the collagen gene likely depends upon its transcriptional history within the cell, the large SC-35 domain with which the gene associates remains after collagen RNA synthesis is inhibited and the RNA accumulation is essentially undetectable. (*c*) We have found large accumulations of specific intron-containing RNAs that do not associate with concentrated accumulations of SC-35, hence RNAs are likely processed by more diffusely distributed splicing factors (Clemson et al., in press; Coleman, J. R., P. T. Moen, and J. B. Lawrence, manuscript in preparation). (*d*) The number of discrete domains of SC-35 (up to ~50) is far fewer than the number of active genes. We note that points 3 and 4 rely on an assumption that essentially the same splicing factors are involved in processing all pre-mRNAs, which may not be the case (Fu, 1993).

It will be important to determine directly whether multiple pre-mRNAs associate with an individual domain. Recognizing that not all prominent domains (or IGC) may be the same functional entities, we suggest two general ideas for what some or all domains might be if they are indeed more than the accumulation of splicing components associated with transcripts of a single gene: (*a*) Domains may represent the association of some transcriptionally active genes with centers involved in maturation and transport of RNA and/or RNA metabolic components. (*b*) Domains may reflect the accumulation of splicing components on pre-mRNA from multiple specific genes, which associate due to either the mutual affinity of their RNAs for splicing components or their genomic organization, or both. Either possibility could facilitate the expression of an mRNA by raising the concentration of RNA metabolic components in the immediate vicinity and would involve nonrandom compartmentalization of specific genes relative to regions of enhanced RNA metabolic activity, as previously suggested (Carter et al., 1991; discussed in Lawrence et al., 1993).

Sequence-specific Structural Arrangements within the Nucleoplasm

Results presented here are indicative of a structured nuclear interior. While splicing factor/poly(A)$^+$ RNA rich domains reside in regions of little to no DNA (Spector, 1990; Carter et al., 1991; Cremer et al., 1993; Visa et al., 1993; Zirbel et al., 1993), we view it unlikely that localizations observed merely reflect diffusion of splicing components or RNAs within inter-chromosome space (see for example Zachar et al., 1993, Cremer et al., 1993). The nuclear splicing factor/poly(A)$^+$ RNA-rich compartment has itself a specific, reproducible 3-D topography that can vary with cell type (Carter et al., 1993; Lawrence et al., 1993). Collagen RNA tracks are highly resistant to triton extraction and remain after nuclear matrix fractionation (Clemson et al., in press). There may well be DNA and RNA or splicing components which randomly distribute and freely diffuse within the nucleus. However, the results presented here show nonrandom distributions of specific genes, a polar configuration of DNA and RNA in RNA tracks, and specific spatial relationships of different pre-mRNAs to domains, all suggest structural relationships with a substantial degree of order.

We appreciate the contributions of clones and antibodies from several scientists which greatly facilitated this work, including X. D. Fu and T. Maniatis, L. Leinwand, B. Strauss, D. Rowe, J. Stein, and P. Dobner. We thank K. Smith for critical comment on the manuscript. We thank H.-L. Haley and P. Sculthorpe for help with manuscript preparation.

This work was supported by National Institutes of Health GM 49254, HG00251, and an RCDA to J. B. Lawrence.

Received for publication 14 August 1995 and in revised form 22 September 1995.

References

Al-Shawi, R., J. Kinnaird, J. Burke, and J. O. Bishop. 1990. Expression of a foreign gene in a line of transgenic mice is modulated by a chromosomal position effect. *Mol. Cell. Biol.* 10:1192–1198.

Barsh, G. S., C. L. Roush, and R. E. Gelinas. 1984. DNA and chromatin structure of the human a1(I) collagen gene. *J. Biol. Chem.* 259(23):14906–14913.

Bernhard, W. 1969. A new staining procedure for electron microscopical cytology. *J. Ultrastruct. Res.* 27:250–265.

Beyer, A. L., and Y. N. Osheim. 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes & Dev.* 2:754–765.

Blencowe, B. J., J. A. Nickerson, R. Issner, S. Penman, and P. A. Sharp. 1994. Association of nuclear matrix antigens with exon-containing splicing complexes. *J. Cell Biol.* 127:593-607.

Blobel, G. 1985. Gene gating: a hypothesis. *Proc. Natl. Acad. Sci. USA.* 82: 8527–8529.

Brasch, K. 1990. Drug and metabolite-induced perturbations in nuclear structure and function: a review. *Biochem. Cell. Biol.* 68:408–426.

Brawerman, G. 1987. Determinants of messenger RNA stability. *Cell.* 48:5–6.

Brawerman, G. 1981. The role of the poly (A) sequence in mammalian messenger RNA. CRC *Crit. Rev. Biochem.* 10:1–38.

Bregman, D. B., L. Du, S. van der Zee, and S. L. Warren. 1995. Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J. Cell Biol.* 129:287–298.

Brown, C. J., B. D. Hendrich, R. G. Lafreniere, L. Rupert, Y. Xing, J. B. Lawrence, and H. F. Willard. 1992. The human Xist gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell.* 71:527–542.

Carter, K. C. 1994. Spatial localization of pre-mRNA transcription and processing within the nucleus. *Curr. Opin. Biotech.* 5:579–584.

Carter, K. C., K. L. Taneja, and J. B. Lawrence. 1991. Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus. *J. Cell Biol.* 115:1191–1202.

Carter, K. C., D. Bowman, W. Carrington, K. Fogarty, J. A. McNeil, F. S. Fay, and J. B. Lawrence. 1993. A three-dimensional view of precursor messenger RNA metabolism within the mammalian nucleus. *Science (Wash. DC).* 259: 1330–1335.

Clemson, C. M., J. A. McNeil, H. F. Willard, and J. B. Lawrence. XIST RNA paints the inactive X chromosome at interphase: evidence for a structural association. *J. Cell Biol.* In press.

Comings, D. 1980. Arrangement of chromatin in the nucleus. *Hum. Genet.* 53: 131–143.

Cremer, T., C. Cremer, T. Schneider, H. Baumann, L. Hens, and M. Kirsch-Volders. 1982. Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. *Hum. Genet.* 62:201–209.

Cremer, T., A. Kurz, R. Zirbel, S. Dietzel, B. Rinke, E. Schröck, M. R. Speicher, U. Mathieu, A. Jauch, P. Emmerich et al. 1993. Role of chromosome territories in the functional compartmentalization of the cell nucleus. *Cold Spring Harbor Symp. Quant. Biol.* 58:777–798.

Davis, L., M. Cadrin, D. L. Brown, and N. Chaly. 1993. Reversible disassembly of transcription domains in lymphocyte nuclei during inhibition of RNA synthesis by DRB. *Biol. Cell.* 78:163–180.

Fakan, S. 1994. Perichromatin fibrils are in situ forms of nascent transcripts. *Trends Cell Biol.* 4:86–90.

Fakan, S., and W. Bernhard. 1971. Localization of rapidly and slowly labeled nuclear RNA as visualized by high resolution autoradiography. *Exp. Cell Res.* 67:129–141.

Fakan, S., and E. Puvion. 1980. The ultrastructural visualization of nuclear and extranucleolar RNA synthesis and distribution. *Int. Rev. Cytol.* 65:255–299.

Fey, E. G., G. Krochmalnic, and S. Penman. 1986. The nonchromatin substructures of the nucleus: the ribonucleoprotein (RNP)-containing and RNP-depleted matrices analyzed by sequential fractionation and resinless section electron microscopy. *J. Cell Biol.* 102:1654–1665.

Fu, X.-D. 1993. Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature (Lond.).* 365:82–85.

Fu, X.-D., and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature (Lond.).* 343: 437–444.

Genovese, C., A. Brufsky, J. Shapiro, and D. Rowe. 1989. Detection of mutations in human type I collagen mRNA in osteogenesis imperfaecta by indirect RNase protection. *J. Biol. Chem.* 264(16):9632–9637.

Gerdes, M. G., K. C. Carter, P. T. Moen, Jr., and J. B. Lawrence. 1994. Dynamic changes in the higher-level chromatin organization of specific sequences revealed by in situ hybridization to nuclear halos. *J. Cell Biol.* 126:289–304.

Hawkins, J. W., and A. Dugaiczyk. 1982. The human serum albumin gene: structure of a unique locus. *Gene (Amst.).* 19:55–58.

Hendzel, M. J., and D. P. Bazett-Jones. 1995. RNA polymerase II transcription and the functional organization of the mammalian cell nucleus. *Chromosoma.* 103:509–516.

Herman, R. C., and S. Penman. 1977. Multiple decay rates of heterogenous nuclear RNA in HeLa cells. *Biochemistry.* 16(15):3460–3465.

Herman, R. C., J. G. Williams, and S. Penman. 1976. Message and non-message sequences to poly(A) in steady state heterogeneous nuclear RNA of HeLa cells. *Cell.* 7:429–437.

Hogan, N. C., K. L. Traverse, D. E. Sullivan, and M.-L. Pardue. 1994. The nucleus-limited Hsr-Omega-n transcript is a polyadenylated RNA with a regulated intranuclear turnover. *J. Cell Biol.* 125:21–30.

Huang, S., and D. Spector. 1991. Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. *Genes & Dev.* 5:2288–2302.

Huang, S., T. J. Derrinck, M. H. Ellisman, and D. L. Spector. 1994. In vivo analysis of the stability and transport of nuclear poly(A)$^+$ RNA. *J. Cell Biol.* 126: 877–899.

Jackson, D. A. 1991. Structure-function relationships in eukaryotic nuclei. *Bioessays.* 13:1–10.

Jackson, D. A., A. B. Hassan, R. J. Errington, and P. R. Cook. 1993. Visualization of focal sites of transcription within human nuclei. *EMBO (Eur. Mol. Biol. Organ.) J.* 12:1059–1065.

Jiménez-García, L. F., and D. L. Spector. 1993. In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism. *Cell.* 73:47–59.

Johnson, C. V., R. H. Singer, and J. B. Lawrence. 1991. Fluorescent detection of nuclear RNA and DNA: implications for genome organization. *In* Methods in Cell Biology. Vol. 35. Academic Press, Inc. 73–99.

Johnson, C. V., D. E. Cool, M. B. Glaccum, N. Green, E. H. Fischer, A. Bruskin, D. E. Hill, and J. B. Lawrence. 1993. Isolation and mapping of human T-cell protein tyrosine phosphatase sequences: localization of genes and pseudogenes discriminated using fluorescence hybridization with genomic versus cDNA probes. *Genomics.* 16:619–629.

Kislauskis, E., B. Bullock, S. McNeil, and P. R. Dobner. 1988. The rat gene encoding neurotensin and neuromedin N. Structure, tissue-specific expression, and evolution of exon sequences. *J. Biol. Chem.* 263(10):4963–4968.

Krause, S., S. Fakan, K. Weis, and E. Wahle. 1994. Immunodetection of poly(A) binding protein II in the cell nucleus. *Exp. Cell Res.* 214:75–82.

Langer-Safer, P. R., M. Levine, and D. C. Ward. 1982. Immunological method for mapping genes on Drosophila polytene chromosomes. *Proc. Natl. Acad. Sci. USA.* 79:4381.

Lawrence, J. B., C. A. Villnave, and R. H. Singer. 1988. Interphase chromatin and chromosome gene mapping by fluorescence detection of in situ hybridization reveals the presence and orientation of two closely integrated copies of EBV in human lymphoblastoid cell line. *Cell.* 52:51–61.

Lawrence, J. B., R. H. Singer, and L. M. Marselle. 1989. Highly localized distribution of specific transcripts within interphase nuclei visualized by in situ hybridization. *Cell.* 57:493–502.

Lawrence, J. B., R. H. Singer, and J. A. McNeil. 1990. Interphase and metaphase resolution of difference distances within the human dystrophin gene. *Science (Wash. DC).* 249:928–933.

Lawrence, J. B., K. C. Carter, and Y. Xing. 1993. Probing functional organization within the nucleus: Is genome structure integrated with RNA metabolism? *Cold Spring Harbor Symp. Quant. Biol.* 58:807–818.

Leavitt, J., P. Gunning, P. Porreca, S.-Y. Ng, C.-S. Lin, and L. Kedes. 1984. Molecular cloning and characterization of mutant and wildtype human beta-actin genes. *Mol. Cell Biol.* 4:1961–1969.

Lentz, T. L. 1971. In cell fine structure: an atlas of drawings of whole-cell structures. W. B. Saunders Pub. Co., Philadelphia, PA. p. 51.

Lerner, E. A., M. R. Lerner, L. A. Janeway, and J. A. Steitz. 1981. Monoclonal antibodies to nucleic acid containing constituents: probes for molecular biology and autoimmune disease. *Proc. Natl. Acad. Sci. USA.* 78:2737–2741.

Lewin, B. 1975. Units of transcription and translation: the relationship between heterogenous nuclear RNA and messenger RNA. *Cell.* 4:11–20.

Lewin, B. 1990. Control of RNA Processing, in Genes IV. Cell Press, Cambridge, MA. 613 pp.

Lichter, P., T. Cremer, J. Borden, L. Manuelidis, and D. C. Ward. 1988. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum. Genet.* 80:224–234.

Lompre, A. M., Nadal-Ginard, B., and V. Mahdavi. 1984. Expression of the cardiac ventricular α- and β-myosin heavy chain genes is developmentally and hormonally regulated. *J. Biol. Chem.* 259:6437–6446.

Manuelidis, L. 1984. Different central nervous system cell types display distinct and non-random arrangements of satellite DNA sequences. *Proc. Natl. Acad. Sci. USA.* 81:3123–3127.

Mattaj, I. W. 1994. Splicing in space. *Nature (Lond.).* 372:727–728.

Moen, Jr., P. T., K. P. Smith, and J. B. Lawrence. 1995. Compartmentalization of specific pre-mRNA metabolism: an emerging view. *Human Molec. Genet.* 4:1779–1789.

Ng, S.-Y., P. Gunning, R. Eddy, P. Ponte, J. Leavitt, L. Kedes, and T. Shows. 1985. Chromosome 7 assignment of the human beta-actin functional gene (ACTB) and the chromosomal dispersion of pseudogenes. *Cytogen. Cell Genet.* 40:712.

Nyman, U., H. Hallan, G. Hadlaczky, I. Pettersson, G. Sharp, and N. R. Ringertz. 1986. Intranuclear localization of snRNP antigens. *J. Cell Biol.* 102:137–143.

O'Keefe, R. T., A. Mayeda, C. L. Sadowski, A. R. Krainer, and D. L. Spector. 1994. Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. *J. Cell Biol.* 124:249–260.

Pinkel, D., J. Landegent, C. Collins, J. Fuscoe, R. Seagraves, J. Lucas, and J. W. Gray. 1988. Fluorescence in situ hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. USA.* 85:9138–9142.

Raska, I., L. E. C. Andrade, R. L. Ochs, E. K. L. Chan, C. Chang, G. Roos, and

E. M. Tan. 1991. Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp. Cell Res.* 195:27–37.

Saez, L. J., K. M. Gianola, E. M. McNally, R. Feghali, R. Eddy, T. Shows, and L. A. Leinwand. 1987. Human cardiac myosin heavy chain genes and their linkage in the genome. *Nucleic Acid Res.* 15:5443–5459.

Salditt-Georgieff, M., M. M. Harpold, M. C. Wilson, and J. E. Darnell, Jr. 1981. Large heterogeneous nuclear ribonucleic acid has three times as many 5' caps as polyadenylic acid segments, and most caps to not enter polyribosomes. *Mol. Cell. Biol.* 1:179–187.

Scherrer, K., G. Spohr, N. Granboulan, C. Morel, J. Groschlaude, and C. Chezzi. 1970. Nuclear and cytoplasmic messenger-like RNA and their relation to the active messenger RNA in polyribosomes of HeLa cells. *Cold Spring Harbor Symp. Quant. Biol.* 35:539–554.

Spector, D. L. 1990. Higher order nuclear organization: three-dimensional distribution of small nuclear ribonucleoprotein particles. *Proc. Natl. Acad. Sci. USA.* 87:147–152.

Spector, D. L. 1993. Macromolecular domains within the cell nucleus. *Annu. Rev. Cell Biol.* 9:265–315.

Spector, D. L., W. H. Schrier, and H. Busch. 1983. Immunoelectron microscopic localization of snRNPs. *Biol. Cell.* 49:1–10.

Spector, D. L., X.-D. Fu, and T. Maniatis. 1991. Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3467–3481.

Trask, B. J., H. Massa, S. Kenwrick, and J. Gitschier. 1991. Mapping of human chromosome Xq28 by two-color fluorescence in situ hybridization of DNA sequences to interphase cell nuclei. *Am. J. Hum. Genet.* 48:1–15.

Visa, N., F. Puvion-Dutilleul, F. Harper, J.-P. Bachellerie, and E. Puvion. 1993. Intranuclear distribution of poly(A) RNA determined by electron microscope in situ hybridization. *Exp. Cell Res.* 208:19–34.

Walton, T. H., P. T. Moen, E. Fox, and J. W. Bodnar. 1989. Interactions of minute virus of mice and Adenovirus with host nuclei. *J. Virol.* 63:3651–3660.

Wang, J., L. G. Cao, Y. L. Wang, and T. Pederson. 1991. Localization of premessenger RNA at discrete nuclear sites. *Proc. Natl. Acad. Sci. USA.* 88: 7391–7395.

Wansink, D. G. 1994. Transcription by RNA polymerase II and nuclear architecture. Ph.D. Thesis. Universiteit van Amsterdam, The Netherlands.

Wansink, D. G., W. Schul, I. van der Kraan, B. van Steensel, R. van Driel, and L. de Jong. 1993. Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J. Cell Biol.* 122:283–293.

Wiegant, J., W. Kalle, L. Mullenders, S. Brookes, J. M. N. Hovers, J. G. Dauwerse, G. J. B. van Ommen, and A. K. Raap. 1992. High-resolution in situ hybridization using DNA halo preparations. *Human Molec. Genet.* 1:587–591.

Xing, Y. 1993. Nuclear structure studied by fluorescence hybridization: Visualization of individual gene transcription and RNA splicing. Ph.D. Dissertation. University of Massachusetts Medical School, Worcester, MA.

Xing, Y., and Lawrence, J. B. 1993. Nuclear RNA tracks: structural basis for transcription and splicing? *Trends Cell Biol.* 3:346–353.

Xing, Y., C. V. Johnson, P. Dobner, and J. B. Lawrence. 1993. Higher level organization of individual gene transcription and RNA splicing: integration of nuclear structure and function. *Science (Wash. DC).* 259:1326–1330.

Zachar, Z., J. Kramer, I. P. Mims, and P. M. Bingham. 1993. Evidence for channeled diffusion of pre-mRNAs during nuclear RNA transport in metazoans. *J. Cell Biol.* 121:729–742.

Zhang, G., K. L. Taneja, R. H. Singer, and M. R. Green. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature (Lond.).* 372:809–812.

Zirbel, R. M., U. R. Mathieu, A. Kurz, T. Cremer, and P. Lichter. 1993. Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries. *Chrom. Res.* 1:93–106.