1 The Nucleus – An Overview

Dean A. Jackson

Faculty of Life Sciences, University of Manchester,
Sackville Street, Manchester, UK

1.1 Introduction

Eukaryotic cells, unlike those of prokaryotic bacteria and archaea, possess a membrane-bound organelle called the nucleus. In general, nuclei are the most conspicuous organells in somatic animal cells (see Figure 1.1), occupying roughly one third of the cell volume. The major function of the nucleus is to act as a repository for the genetic information, encoded in DNA. Somatic mammalian cells generally carry two copies of the DNA genome, one inherited from the female parent and the other from the male. In humans, this information is held within 46 chromosomes, two copies each of 22 autosomes and two sex chromosomes – XY in males and XX in females. Nuclei contain the machinery needed to ensure the fidelity (repair) and duplication (replication) of the genetic code. Nuclei also house the complex machinery needed to ensure that the genetic information can be selectively decoded so that the requisite genes are expressed in the many different cell types present in multicellular eukaryotes. Activating programmes of gene expression is a highly complex process that relies on the interplay between activating transcription factors, the synthetic machinery and different chromatin states that arise through the action of a variety of chromatin remodelling systems. Some chromatin states are permissive for gene expression and allow transcription factors to access DNA so that the programme of events that lead to RNA synthesis can be initiated. Other types of chromatin have a structure that is operationally immune from the action of transcription factors. Many factors are known to influence and stabilize these chromatin states. Local and long-range features of chromosome architecture in particular play a fundamental role in defining both chromatin or gene domains and nuclear compartments where

Viruses and the Nucleus Edited by Julian A. Hiscox
© 2006 John Wiley & Sons, Ltd

NEOGENOMICS EX.1028.001

c01.fm Page 2 Wednesday, March 15, 2006 8:05 AM

transcription is permitted or excluded. As well as establishing chromatin compartments, nuclei are also believed to support a wide variety of other compartments that probably influence the efficiency of chromatin function. Understanding the importance of these specialized nuclear regions, how they work and the purpose they serve could be of fundamental importance to understanding the critical subtleties of chromatin function.

Much of our present understanding of nuclear structure and function in eukaryotic cells has evolved as a result of experiments performed over the past few decades. Viruses, in contrast, have experienced many millions of years of co-evolution with these cells, during which time they have grown to utilize and subvert their hosts. This volume is dedicated to understanding how viruses use host cells to meet their own needs. In many cases viruses are selfish and the infected cell is destined to die. Nevertheless, it is interesting to investigate how viruses use the host machinery to their own ends, and in particular to consider what studies of viral infection are able to teach us about the fundamental principles of nuclear architecture in eukaryotic cells. This chapter sets the scene and explains what confronts a virus as it enters the eukaryotic cell nucleus.

1.2 The gene expression pathway in mammalian cells

In mammals, some 250 distinct cell lineages can be defined by the specialized roles that each performs. Different cell lineages have quite distinct properties and appearance, even though almost all of these cells contain the same genetic information. The phenotype of each of these lineages is determined by patterns of protein synthesis that in turn reflect levels of gene expression. Patterns of gene expression are determined by the dynamic interplay between chromatin, chromatin remodelling machinery, transcription factors and synthetic machinery that drives RNA synthesis. Three different RNA polymerase complexes perform RNA synthesis in the nuclei of mammalian cells. In most cells, RNA polymerase II (pol II) is the major activity, transcribing all protein-coding genes to generate patterns of gene expression that determine cell type. Synthesis is performed by a ~ 4 MD holoenzyme that contains the pol II core enzyme associated with other activities required during RNA synthesis and processing. RNA polymerase I (pol I) is dedicated to the synthesis of the repeated ribosomal RNA (rRNA) genes, within specialized nuclear sites – nucleoli – and RNA polymerase III (pol III), a minor nucleoplasmic activity, transcribes transfer RNA (tRNA) and 5S rRNA genes. Small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) genes encode structural RNAs needed for RNA processing; some are transcribed by pol II and others by pol III.

Proliferating mammalian cells have roughly 30 000, 60 000 and 3000 engaged pol I, II and III complexes/cell, respectively (Jackson *et al.*, 2000). RNA is polymerized at ~ 2–2.5 kbp/min; this rate can be estimated from the time taken to transcribe RNA molecules of known length. The average length of a primary transcript is about 15 kbp in mammalian cells, although the primary transcripts are processed, during splicing,

NEOGENOMICS EX.1028.002

c01.fm Page 3 Wednesday, March 15, 2006 8:05 AM

to generate mRNAs that are roughly one-tenth of this size. The mammalian genome (haploid) has an estimated 35 000 genes and $\sim 3 \times 10^9$ bp DNA. Proliferating cells contain roughly 3.5×10^6 copies of the 18 and 28S ribosomal RNAs, 3×10^7 tRNA and 0.4×10^6 mRNA molecules; of these, a few mRNAs may have many thousand copies per cell, although most have less than 10 (Jackson *et al.*, 2000). In human cells, an active ribosomal RNA gene transcribed by pol I gene has ~120 engaged synthetic complexes and typical pol II transcribed genes has very few; most active genes have only one to three transcription complexes. With these facts in mind, it is important to remember that in mammalian cells the expression of single copy genes can vary by many orders of magnitude. This will inevitably reflect rates of RNA synthesis, although it is important to note that factors such as mRNA turnover and rates of protein synthesis also play a role. The amplification afforded by protein synthesis ensures that very few templates are required for most proteins and, in corollary, that very few genes are expressed continually, so that most genes in euchromatin remain transcriptionally inert for long periods of time.

This raises obvious questions about the dynamics of gene expression. If transcription factor recruitment to a promoter allows productive association of the transcription machinery, RNA synthesis will ensue. Most transcripts are synthesized within 5–10 min. During this time RNA processing and maturation will begin. After roughly 15 min processing and maturation are complete so that the fully mature mRNA is released from the transcription sites. As a general rule, protein complexes involved in maturation and particularly splicing ensure that only mature mRNAs pass into the RNA export pathway. During export, mRNAs are believed to move by diffusion, so that within a few minutes mRNAs pass to the nuclear periphery and through nuclear pores to the cytoplasm (Femino *et al.*, 1998).

1.2.1 Euchromatin and heterochromatin

During embryonic development, defined cell-type specific patterns of gene expression are established through developmentally controlled expression of relevant transcription factors. The patterns of gene expression that arise are inherited throughout subsequent cell division cycles (Cavalli and Paro, 1999). A fundamental feature of this process involves chromatin status and the ability with which DNA sequence motifs are able to engage the factors needed to activate expression. This process is complex and multifactorial. At one level, modifications in DNA, such as methylation, and changes in nucleosome structure, generated by covalent histone modifications, correlate with the expressional capabilities of different classes of chromatin. Highly specialized nucleosome assembly factors and numerous sophisticated chromatin-remodelling complexes contribute to the maintenance and modulation of chromatin structure. But on top of this, more stable chromatin forms that characterize heterochromatin can have a dramatic impact on gene expression.

In most situations heterochromatin is the dominant stable chromatin form. It has been known for many years that active chromatin translocated adjacent to heterochromatin

NEOGENOMICS EX.1028.003

c01.fm Page 4 Wednesday, March 15, 2006 8:05 AM

will generally acquire a heterochromatic structure and lose the ability to support gene expressions. This phenomenon is emphasized if genes are introduced close to the centromeric heterochromatin in transgenic mice. In this case, modulating the concentration of heterochromatin proteins has been shown to influence the efficacy of gene silencing (Festenstein *et al.*, 1999). In *Drosophila*, the suppressive influence of heterochromatin was characterized in the study of position effect variegation, where the position of a gene in the genome can dramatically influence its level of transcription. These early studies characterized chromatin proteins that are known to regulate gene expression, such as HP1 (heterochromatin protein 1) and Pc (polycomb). Mechanistically, a chromatin organization-modifying chromodomain (CD) within both HP1 and Pc is known to play a critical role in heterochromatin formation. Recent experiments in mammalian cells have shown that the chromodomain of HP1 α binds to histone H3 and also to the linker histone H1 (Breiling *et al.*, 1999; Nielsen *et al.*, 2001). The polycomb CD also binds to histone H3. These observations suggest that interactions between HP1 and the histone proteins, together with the self-association properties of HP1, are crucial factors in HP1-mediated heterochromatin assembly. In addition to this, specific modification of chromatin – such as methylation of histone H3 lysine 9 – creates a binding site for HP1 proteins and stabilizes the heterochromatic state (Lachner *et al.*, 2001). Polycomb group proteins and trithorax proteins are thought to act antagonistically to control dynamic transitions between inert and open chromatin states (Muyrers-Chen and Paro, 2001).

1.2.2 Chromatin status and gene activation

In heterochromatin, specific chromatin structures restrict the ability of transcription factors and the transcription machinery to access binding sites within gene promoters and so prevent gene expression. Before gene expression can be activated, the appropriate chromatin structure must be established. Changes that occur in chromatin in response to gene activation and transcription have been reviewed in detail (Workman and Kingston, 1998; Jenuwein and Allis, 2001; Becker 2002). For the purposes of the present discussion, it will be sufficient to recognize that a combination of ATP-dependent chromatin remodelling machines, together with histone acetylase and deacetylase complexes, cooperate to generate chromatin states that are permissive for gene expression. For genes located in inert chromatin, the activation process is stochastic. The initial process will involve activating factors that recruit appropriate ATP-dependent chromatin remodelling machinery and open chromatin sufficiently for the promoter to now be accessibe to further factors (Becker, 2002). Remodelling complexes of the ISWI and SWI/SNF families are recruited by a variety of activating factors and induce nucleosme sliding to reveal binding sites that were previously masked by chromatin structure. These early changes then allow the binding of other transcription factors, other remodelling machines and histone acetylases that stabilize the open chromatin status further to facilitate assembly of the transcription machinery.

NEOGENOMICS EX.1028.004

c01.fm Page 5 Wednesday, March 15, 2006 8:05 AM

Finally, as transcription proceeds, proteins associated with the RNA polymerase holoenzyme generate an extended region of modified, transcriptionally competent, chromatin. This is a complex process, but it is also worth remembering that the process is dynamic, so that these different chromatin modifications will be in a constant state of flux.

This pathway has many interesting features. The transition from inert to active chromatin involves structural changes within chromatin at promoters and the spatial relocation of the relevant gene (away from heterochromatin) prior to the onset of RNA synthesis (Lundgren *et al.*, 2000). Decondensation of compact chromatin is seen during the activation of gene expression, although the process of RNA synthesis appears to have little further impact on the dynamic properties of the active domains (Tsukamoto *et al.*, 2000). Many proteins – perhaps 100 or more – are subsequently required to activate gene expression and many of these are dynamic in nature. The dynamic properties of glucocorticoid receptor coupled to green fluorescent protein (GR-GFP) emphasize this point (McNally *et al.*, 2000). A mouse cell line, 3617, containing a single locus with a 200-copy MMTV LTR array (and some 1000 GR-GFP binding sites), was used. Dexamethasone treatment of GR-GFP cells induces the nuclear accumulation of the activated receptor, so that in most cells a single region of intense staining correlates with the location of the MMTV array. Under these conditions, a FRAP (fluorescence reactivation after photobleaching) analysis showed that bound receptor molecules were only transiently associated with the promoter, with almost all bound factors exchanging in ~10 s.

1.2.3 Gene location and expressional competence

Gene expression from ectopic chromosomal sites is extremely unpredictable. Predictably, genes introduced into heterochroamtin are generally inactive, although expression can be modulated by the availability of heterochromatin proteins (Festenstein *et al.*, 1999). Ectopic genes in active chromatin also behave unpredictably in cultured cells or transgenic animals (Wilson *et al.*, 1990). For example, it is possible to analyse gene expression from artificial gene constructs introduced in predetermined target sites by homologous recombination. In one case mouse ES cells were used to express an Oct4/*lacZ* transgene (Wallace *et al.*, 2000). Using homologous recombination, four developmentally controlled promoters were introduced into the target locus of two selected clones and their ability to drive *lacZ* expression in transgenic animals determined. Animals derived from one clone showed appropriate regulation of the four different promoters. However, animals derived from the second showed additional, unexpected patterns of ectopic expression, confirming that gene location can have profound effects on expression throughout development.

While chromosome location can influence gene expression, it transpires that nuclear location might also play a role. Studies on the lymphoid lineage of mammalian cells have demonstrated that the inactivation of gene expression correlates with

NEOGENOMICS EX.1028.005

c01.fm Page 6 Wednesday, March 15, 2006 8:05 AM

relocation of a gene to heterochromatic nuclear sites, while the reverse process occurs during gene activation. During this process, the sequence-specific transcription factor Ikaros becomes associated both with the silenced gene and local centromeric heterochromatin. Ikaros is able to bind both a target promoter and sites within the centromeric satellite repeats, so providing a means of driving appropriate genes into nuclear sites that are not permissive for gene expression (Brown *et al.*, 1997). Other examples are known where nuclear location influences gene function. During B lymphocyte development, the immunoglobulin H (IgH) and IgK loci are located at the nuclear periphery in haematopoietic progenitors and pro-T cells and in the nuclear interior in pro-B nuclei (Kosak *et al.*, 2002). The inactive loci are associated with the nuclear lamina and must move to active sites within the nuclear interior before recombination and transcription of the IgH and IgK loci can occur.

1.2.4 Chromatin domains as units of gene expression

The unpredictable behaviour of gene expression from ectopic sites implies that natural gene loci have evolved mechanisms to protect the required levels and patterns of gene expression. One way of achieving this is to place genes in chromatin domains that are spatially isolated for local epigenetic factors. Clear evidence for the existence of domains of chromatin function comes from the ease with which chromatin is cut by nucleases (Workman and Kingston, 1998); transcriptional status correlates with a generalized nuclease 'sensitivity' that results from an open (10 nm) chromatin fibre. Moreover, chromatin modifications that correlate with gene activity, such as histone acetylation, play a role in maintaining chromatin status that can be perpetuated once the initial developmental transactivators are removed (Cavalli and Paro, 1999). Hence, chromatin modifications are capable of generating epigentic signals that could play a role in the control gene expression, provided that this information can be reproduced when chromatin is duplicated (Taddei *et al.*, 1999).

As chromatin remodelling accompanies transcription, RNA synthesis will play a role in maintaining chromatin status. However, it is well known that the general nuclease sensitivity at active loci is not restricted to genes. Chromatin immunoprecipitation experiments (using antibodies to histone variants found in active chromatin) confirm that zones of active chromatin might spread well beyond the coding region of a particular gene. Moreover, regions of transition from active to inactive chromatin structure imply that chromatin domains are demarcated by elements that determine the boundaries of functional genetic units (Elefant *et al.*, 2000; Litt *et al.*, 2001). A high-resolution analysis of acetylation across three gene loci in the vicinity of the chicken β-globin locus has emphasized this point (Litt *et al.*, 2001). Across this region, three levels of chromatin acetylation can be defined. The condensed chromatin maintains very low levels of histone acetylation at all developmental stages, with similar levels maintained in inactive genes. Much higher levels of acetylation are seen throughout transcribed gene domains, while chromatin in the vicinity of upstream regulatory sites maintains

NEOGENOMICS EX.1028.006

c01.fm Page 7 Wednesday, March 15, 2006 8:05 AM

the highest levels of acetylation. Most significantly, a very strong constitutive focus of hyperacetylation corresponds with an insulator element that appears to demarcate the globin and adjacent folate receptor domains. This demonstrates how epigenetic modifications might influence chromatin domain structure and also implies that different classes of histone acetyl transferase, with different chromatin targets, control gene expression.

Interestingly, experiments analysing expression of human β-globin gene constructs in transgenic mice have shown that the locus is composed of three functional sub-domains, which acquire an active chromatin status at the time of their expression, at the appropriate stage of development (Gribnau *et al.*, 2000). Intriguingly, the appearance of these nuclease-sensitive domains correlates with the activity of intergenic, non-globin transcripts that extend throughout the corresponding sub-domains. However, the maintenance of chromatin status does not demand continual transcription of the non-genic sequences, as most synthesis is seen in the G_1 cell cycle phase (see Chapter 2 for description of the cell cycle) of appropriate cells. With this in mind, it is notable that a substantial fraction of primary transcripts synthesized in mammalian cells by RNA pol II are not destined to produce mature mRNA (Jackson *et al.*, 2000). Perhaps non-productive RNA synthesis acts to keep chromatin domains 'open', so preventing the formation of inactive chromatin states.

1.2.5 Domains of gene expression

We have seen that in many eukaryotes, individual genes are expressed to quite different extents and many are expressed in specific cell types and at precise times during development. This begs the question: how are different levels of gene expression maintained? For a specific gene, chromatin status and the availability of activating transcription factors will combine to establish an engaged transcription complex that drives RNA synthesis. Following activation, certain combinations of promoter-bound proteins must drive many cycles of synthesis, while others must decay and so generate many fewer transcripts. Gene promoters define the location of the transcription start site, but in the chromosomal context, promoters are rarely sufficient to define natural levels of expression. Operationally, expression is augmented and modulated by genetic elements such as enhancers, locus control regions (LCRs), scaffold/matrix attachment regions (S/MARs) and insulators. Together these elements help us to define a unit or domain of gene expression.

Enhancers generally contain arrays of transcription factor recognition motifs that bind appropriate factors to augment levels of gene expression by directly upregulating rates of transcription (Blackwood and Kadonaga, 1998). Enhancer function is dependent on factor binding and is likely to contribute to the protein complex that drives assembly of the transcription machinery. Enhancers may be responsible for establishing a productive spatial configuration in chromatin or for delivering chromatin to specific nuclear sites. Related elements called locus control regions (LCRs) are also complex

NEOGENOMICS EX.1028.007

c01.fm Page 8 Wednesday, March 15, 2006 8:05 AM

arrays of factor binding motifs that make critical contributions to levels of gene expression at the natural chromosomal locus and allow position-independent expression from ectopic sites (Engel and Tanimoto, 2000; Bulger *et al.*, 2002). LCRs contribute to the maintenance of expressional status and are believed to achieve this by maintaining adjacent chromatin in an active configuration; probably by recruiting chromatin remodelling machinery. However, the view that LCRs directly influence chromatin status has been challenged, using recombination strategies to investigate LCR function by deletion of elements at the natural locus (Bender *et al.*, 2000). This approach appears to suggest that LCR elements operate as very efficient enhancers at the natural locus that do not themselves directly influence chromatin status across the locus. An explanation for this apparent controversy is that LCR function is dependent on chromosome context.

1.2.6 The nuclear matrix and gene domains

The nuclear matrix and related scaffold are believed to play important roles in different aspects of chromatin function (Bode *et al.*, 2000). Scaffold/matrix attachment regions (S/MARs) clearly augment transcription and may achieve this using AT-rich sequences to modulate superhelical stress that arises during transcription. S/MARs may also play a role in targeting gene domains to matrix-associated transcription centres prior to gene expression. As an example, MAR elements play a critical role in orchestrating the temporal and spatial expression of many genes during T cell development (Alvarez *et al.*, 2000). In a comprehensive comparison of different classes of potential boundary elements, a strong MAR was shown to support efficient and long-term gene expression from ectopic chromosomal sites (Zahn-Zabal, 2001).

Insulator or boundary elements are believed to function as chromatin domain boundaries (Bell and Felsenfeld, 1999; Udvardy, 1999). A possible role for boundary elements was first suggested by their ability to uncouple linked promoters and enhancers, presumably by 'blocking' enhancer function. This property supports their role as barriers to the influence of neighbouring DNA elements. Many elements that are capable of providing insulator activity and proteins that bind to them have now been described. Characterized boundary element attachment factors (BEAF-32A and -32B) and their recognition motifs within the specialized chromatin structures of *Drosophila* may be important determinants of chromosome structure (Zhao *et al.*, 1995). In vertebrates, recent studies have confirmed that sequence elements at the upstream border of the chicken β-globin LCR act as classical insulators (Bell *et al.*, 1999). A protein, CTCF, has been shown to bind to critical sequence motifs within this region and is thought to play a role in determining insulator function (Farrell *et al.*, 2002). Intriguingly, this same protein appears to play a role in the process of 'genomic imprinting' to control the differential expression of maternally and paternally inherited genes (Ferguson-Smith and Surani, 2001).

NEOGENOMICS EX.1028.008

c01.fm Page 9 Wednesday, March 15, 2006 8:05 AM

1.3 Chromosome structure and DNA loops

Gene distribution within the chromosomes of mammalian genomes is not uniform. Even within individual chromosomes, the majority of genes are clustered within chromatin that can be defined using cytological criteria. These chromosomal R-bands are interspersed with G-bands that contain very few genes and are rich in inert, repetitive sequences. The arrangement of these elements will clearly impact on both local (Gasser, 2002) and long-range (Cremer and Cremer, 2001) chromatin organization. It is important to realise that mammalian chromosomes occupy discrete regions within interphase nuclei. These chromosome 'territories' do not occupy specific nuclear positions, although gene activity can influence interphase chromosome location, so that chromosomes with a high density of active genes tend to be located towards the nuclear centre (Bridger *et al.*, 2000). Interestingly, there is some evidence that different classes of chromatin within individual territories may be spatially separated – perhaps through aggregation of heterochromatin – to generate functional 'polarity' within territories (Sadoni *et al.*, 1999). This potential to develop local compartments of active and inert chromatin must reflect the path of the chromosome axis within each territory and the organization of chromosomal sub-domains represented by DNA foci. These structures are thought to represent structurally related gene clusters that are stable over many cell generations (Jackson and Pombo, 1998); they are derived for individual chromosome bands and typically contain ~ 1 Mbp DNA.

In mammals, a typical interphase chromosome territory might be represented as a sphere that measures roughly 2.5 μm in diameter. Within this volume, chromatin is only locally dynamic; for most chromatin relative movements of > 500 nm are seen only rarely (Abney *et al.*, 1997). This implies that long-range movements are constrained by chromosome structure. The borders of individual territories appear distinct, implying that interaction of chromatin for adjacent chromosomes is spatially restricted. However, in some instances, chromatin can be seen to escape from the confines of an individual territory to form loops that spread ~ 1–2 μm into the adjacent nucleoplasm. Chromatin loops of this sort probably reflect the transcriptional status of highly expressed gene-rich chromosomal regions, and their extruded appearance supports the impression that they expand to occupy local chromatin channels. The functional implications of this type of chromatin movement are unclear.

1.4 Duplicating the genetic code

In multicellular organisms it is self-evident that many cells must arise by duplication of a single precursor, the fertilized egg; the adult human has roughly 5×10^{13} somatic cells. Moreover, it is critical that both the genetic code and epigenetic information are faithfully reproduced during each cell division cycle (see Chapter 2 for further details of the cell cycle). In fact it is reasonable to suggest that the precision with which DNA

NEOGENOMICS EX.1028.009

c01.fm Page 10 Wednesday, March 15, 2006 8:05 AM

replication occurs is one of the greatest feats performed by higher eukaryotic cells (Berezney *et al.*, 2000; Gilbert, 2002). Control of DNA synthesis relies on the complex interplay of chromosome structure, chromatin organization and perhaps even global nuclear architecture. This complex process has a number of steps that occur at different points of the cell cycle. Initially, replicons that are activated at the onset of S phase and most probably the downstream replication programme are established during the first 2 h of the G_1 phase of the cell cycle. This 'temporal decision point' coincides with the period when chromosomal territories re-establish their interphase structure and full transcriptional potential. The process of engaging an S phase programme that duplicates the genome in a predetermined order is distinct from the process that establishes active origins of DNA replication. This second process is thought to involve the binding of the origin recognition complex (orc) to particular chromosomal sites. This process occurs during the mid-G_1 phase and is designated the 'origin decision point'. About 2 h after this, mammalian cells pass a restriction point, at which time they are committed to proceed through S phase and on to cell division.

During S phase, replication takes place in dedicated replication factories. These sites contain all the machinery need to duplicate the DNA of small replicon (replication unit) clusters with about 1 Mbp DNA. Like transcription factories, these structures appear to be assembled on the nucleoskeleton; many components of the replication machinery have defined nucleoskeleton/replication factory-targeting sequences. At the onset of S phase, a specific sub-set of DNA foci is selected for replication. These foci are from chromosomal R bands and are rich in transcribed genes. Typical replicons within these early replicon clusters are ~150 kbp in length, so that most replicons activated at the onset of S phase will be completed 30–60 min later. As replication of these replicons begins to terminate, DNA synthesis is activated at new sites. This process continues until after about 10 h the entire genome has been duplicated. Although the precise mechanisms are unknown, chromosome architecture is believed to play a pivotal role in defining this S phase programme.

1.5 Nuclear compartmentalization

The cytoplasm of a mammalian cell contains various membranated organelles, which provide compartments that are dedicated to specific roles. The nucleus contains no equivalent structures and appears ill-organized in comparison. A classical electron micrograph of a typical human cell emphasizes this impression (Figure 1.1). The nuclear membrane, nucleolus and regions rich in condensed heterochromatin stand out in contrast to the relatively amorphous nuclear interior. Detailed inspection of this interior reveals an interchromatin space, which is rich in hnRNP (an RNA-binding protein) and contains characteristic structures, such as perichromatin fibrils, perichromatin granules, interchromatin granules and interchromatin granule clusters. These are known to be related to different aspects of RNA metabolism (Fakan, 1994). The major message from this type of image is that chromatin-rich regions are generally distinct

NEOGENOMICS EX.1028.010

c01.fm Page 11 Wednesday, March 15, 2006 8:05 AM

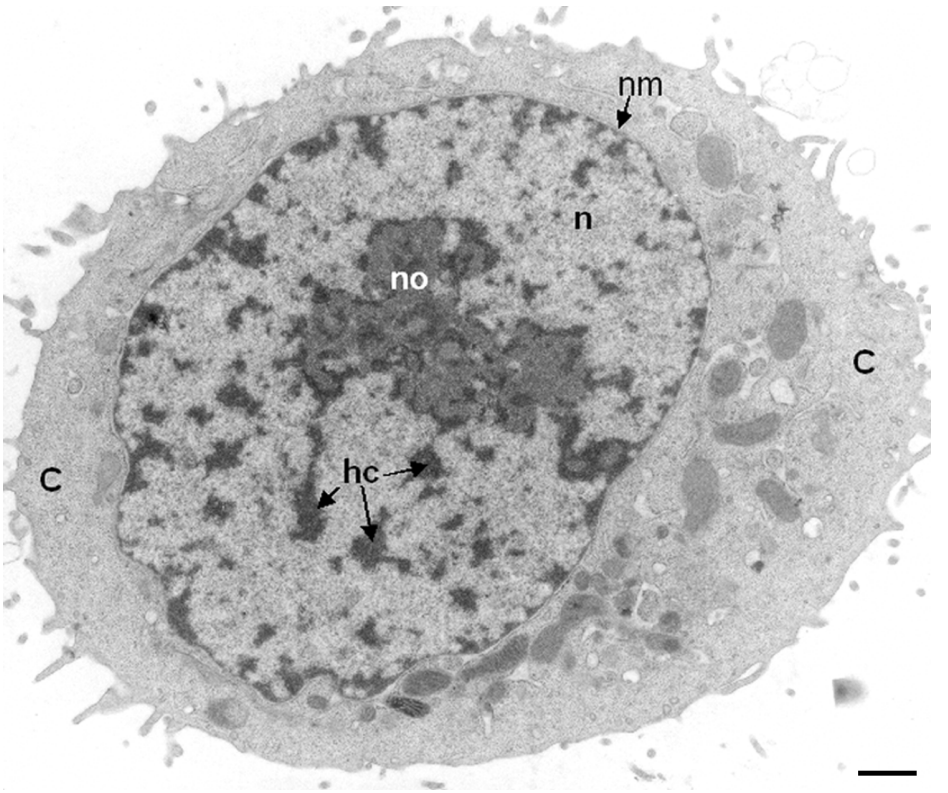

Figure 1.1 Morphology of a mammalian cell. Proliferating HeLa cells (this is a transformed human cell line) were fixed, stained and processed using standard electron microscopy techniques. Thin sections were then prepared and inspected. This typical cell section shows a central nucleus (n), typically of about 10 μm diameter in an equatorial section, which is separated from the cytoplasm (c) by a double nuclear membrane (nm). The nuclear membrane is studded with a few thousand nuclear pores that allow molecules to pass between the two major cell compartments. Note that the cytoplasm contains a large number of membrane-bound organelles. The most conspicuous of these, the mitochondria, are dark staining and typically about 1 μm in diameter in this example. The nucleus does not contain any membrane-bound sub-compartments. Even so it is highly organized. In this type of image, the most obvious substructure is the nucleolus (no). In this example a single large nucleolus occupies the centre of the nucleus. The nucleolus is densely staining and has three distinct subcompartments. The most densely-staining regions are called the dense fibrillar component; these structures are generally crescent-shaped and a few hundred nm in size. The dense fibrillar components represent the nascent transcripts that surround the much paler fibrillar centres. The processed ribosomal RNAs are assembled into ribosomal subunits in the granular component, which comprises the bulk of the nucleolus. Dark-staining clumps of material that lie against the surface of the nucleolus, the nuclear face of the membrane and are scattered throughout the centre of the nucleus are dense heterochromatin (hc). The remainder of the nucleus has an amorphous appearance and little discenrable structure. This nucleoplasmic region contains the transcribed genes in euchromatin and is composed of a mixture of this open chromatin and the interchromatin material, which is predominantly RNA and protein. Bar = 1 μm

NEOGENOMICS EX.1028.011

c01.fm Page 12 Wednesday, March 15, 2006 8:05 AM

from the RNA-rich interchromatin regions and that gene transcription occurs predominantly at the interface between these two zones.

The nucleolus provides the best example of a nuclear compartment with a well-characterized role (see Chapter 7 for details of the nucleolus). The anatomy of a mammalian cell nucleolus is shown in Figure 1.1. Fibrillar centres are rich in the synthetic machinery – RNA polymerase I – and are surrounded by a dense fibrillar component, which contains the nascent RNAs. An individual fibrillar centre is likely to be associated with the genes and nascent products from a single active rDNA cluster, typically three to five genes. Following transcription, the nascent transcripts are processed and assembled into the preribsomal subunits that form the granular component of the nucleolus. The fibrillar centres of nucleoli are the most active transcription factories in mammalian cells. A proliferating cell produces at least 2.5×10^6 ribosomes in each cell cycle. Each ribosome contains the 28S, 18S and 5.8S RNA molecules that are derived from the 45S primary transcript of a ribosomal RNA gene. A proliferating cell has roughly 25 active fibrillar centres with roughly 100 active genes. Each of these has about 120 engaged polymerase complexes at any time. A recent study has used many nucleolar proteins coupled to a green fluorescent protein (GFP) to provide a detailed insight into the dynamic features of ribosomal RNA gene transcription (Dundr *et al.*, 2002).

The complexity of nuclear organization is emphasized when the nuclear distribution of different proteins is analysed by immunostaining. This approach shows that many proteins involved in different aspects of nuclear function are found to be concentrated at discrete nuclear sites, often with a punctate distribution (van Driel *et al.*, 1996; Hendzel *et al.*, 2001). This general impression implies that different proteins accumulate at sites where specific functions are performed but, perhaps more importantly, suggests that many proteins have specific mechanisms that ensure that their distribution throughout nuclear space is not uniform. The nuclear speckles provide an excellent example of this. The term 'nuclear speckle' was used to describe the appearance of the structures that immunostained with a human autoimmune Sm serum. The structures were later shown to be equivalent to interchromatin granule clusters that can be visualized by electron microscopy. The reacting protein and many others that stain the same structures are now known to be involved in RNA metabolism and particularly RNA processing. However, the structures themselves are not active centres; instead they appear to serve as storage depots. Interestingly, while speckles form a distinct nuclear compartment, which can be purified, the proteins within them are very dynamic. Hence, speckles represent a steady-state snapshot of the distribution of a protein population that rapidly moves throughout most of the available nuclear space.

The fact that high concentrations of particular proteins exist at defined nuclear sites raises obvious questions about the role of these compartments, particularly with respect to structure–function relationships. With this in mind, careful studies have demonstrated that many very active genes are located at the periphery of nuclear speckles (Smith *et al.*, 1999). However, this is clearly not the case for all transcribed genes, most of which are not located at borders of the speckle domains; at these

NEOGENOMICS EX.1028.012

c01.fm Page 13 Wednesday, March 15, 2006 8:05 AM

remote sites, proteins that are found at high concentrations in speckles are supplied from the dynamic nucleoplasmic pools. The fact that highly active genes and speckles undergo function-dependent interactions gives us a valuable insight into the principles of nuclear architecture. Here it seems that the major structures are fundamentally plastic, so that changes in the structure of chromosome territories and local changes in the arrangement of dynamic compartments, like speckles, determine how the different components might interact. In this way, highly active genes might be located to the periphery of a chromosome territory, where the splicing components can accumulate in close proximity to sites of greatest demand (Misteli and Spector, 1998). Mechanistically this is not outrageous, as the... carboxy-terminal domain (CTD) of pol II is known to couple RNA synthesis and processing (McCraken *et al.*, 1997). If the cross-talk between RNA polymerase and splicing is eliminated by removing most of the CTD, the association between highly active genes and nuclear speckles is lost (Misteli and Spector, 1999). More generally, disrupting global RNA synthesis with inhibitors causes a dramatic alteration in the appearance of speckles, which are seen to lose surface texture, become more spherical and aggregate. This confirms the dynamic nature of this compartment and the fact that its appearance is related to function. In addition to this, recent studies have indicated that the efficiency of RNA processing is disrupted if the components normally found within speckles are artificially dispersed (Sacco-Bubulya and Spector, 2002). This observation suggests that the speckle compartment plays an important role in the cycle of events that is required to generate active spliceosomes and then splice the nascent RNA; if the speckles are not essential, they certainly influence the efficiency of the splicing process.

Speckles are large subnuclear structures, many of which will measure hundreds of nanometres across. However, it is clear that the same principles apply to other aspects of nuclear organization. For example, nascent RNA can be labelled in mammalian cells using Br-UTP (Jackson *et al.*, 1993; Wansink *et al.*, 1993) and Biotin-CTP (Jackson *et al.*, 1998; Figure 1.2) and sites of incorporation detected by immunolabelling to reveal transcripts associated with the dense fibrillar component of nucleoli and many (~5000) sites, dispersed throughout the nucleoplasm. High-resolution immunostaining techniques show that the active transcription sites contain transcript-rich zones measuring ~50 nm with adjacent regions rich in chromatin, transcription factors, the active form of RNA polymerase II or proteins required for RNA processing (Pombo *et al.*, 1999; Iborra *et al.*, 1996). As these sites perform all steps needed to generate mature mRNA, they have been called transcription 'factories'. Perhaps surprisingly, detailed studies have shown a typical factory contains many active RNA polymerase complexes (typically about five), although the majority of transcribed genes contain fewer engaged polymerase complexes. This raises the possibility that groups of genes are transcribed together within specialized nuclear compartments that are dedicated to performing all the steps needed to generate mature mRNA–protein complexes. Protein–nucleic acid interactions within these compartments are thought to account for the majority of DNA loops found within mammalian cells (Cook, 1999).

NEOGENOMICS EX.1028.013

c01.fm Page 14 Wednesday, March 15, 2006 8:05 AM

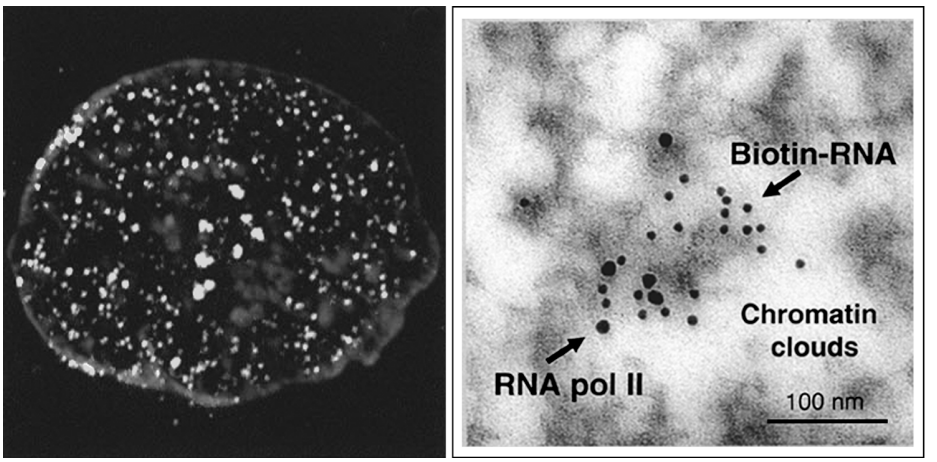

Figure 1.2 Transcription centres in human cells. Patterns of transcription in human cells can be visualized by either light (left) or electron (right) microscopy. In these examples, the light microscopy is used to demonstrate the global distribution of the transcription centres and the electron microscopy a more detailed image of the anatomy of a single transcription factory. For the first image, Hela cells were permeabilized, nascent transcripts extended in Br-UTP and cryosections (~ 100 nm) prepared. Br-RNA was indirectly immunolabelled with Cy3 and nucleic acids were counterstained with TOTO-3. Red and far-red images were collected using a 'confocal' microscope. The transcription sites are bright (shown green) foci. Note the distribution of very bright nucleolar foci in the centre and many smaller sites – RNA polymerase II/III transcription centres – throughout the nucleoplasm. The very bright foci will typically contain ~ 4 ribosomal RNA genes and ~ 500 nascent transcripts. The nucleoplasmic sites have ~ 5–10 active transcription complexes. TOTO-stained heterochomatin is clearly visibly around the nuceolus and along the nuclear periphery. Reproduced from Pombo *et al.* (1999) with *permission* of Oxford University Press. The second image is an electron microscopy section that has been prepared in LR-white, treated using the EDTA regressive staining technique and immunolabelled using antibodies to RNA polymerase II and Biotin-RNA (labelled *in vitro* using Biotin-CTP). Note the relative distributions of RNA polymerase II (large, 10 nm, gold particles), Biotin-RNA (small, 5 nm, gold particles), chromatin clouds (pale areas) and RNA rich interchromatin channels (grey areas). The transcription centre of ~ 100 nm in this example corresponds to 1 nucleoplasmic focus in the light microscopy image. Image courtesy Francisco Iborra

The idea that active RNA polymerase complexes might be localized within discrete and dedicated nuclear sites can have profound implications for gene expression. In particular, the fact that groups of engaged polymerases are found at each site means that the enzymes must be 'fixed' within the sites and the chromatin mobile. How chromatin interacts with the active centre would then dictate levels of gene expression. Moreover, it is also interesting to consider how chromatin structure and genetic elements outside promoters might influence gene activity. *A priori*, if chromatin is dynamic the fluid properties inherent in different chromatin states might have a profound influence on the potential of a promoter to engage any local active

NEOGENOMICS EX.1028.014

c01.fm Page 15 Wednesday, March 15, 2006 8:05 AM

centres – genes in or close to heterochromatin would be more inert and their movement locally constrained, while sequences such as enhancers, LCRs and S/MARs might contribute to the overall efficiency of gene expression by targeting genes to the active centres. All of these components are likely to be in a finely balanced dynamic equilibrium that reflects both local and long-range chromatin organization and contributes in some way to levels of gene expression.

While many components involved in chromatin function are extremely dynamic or mobile, others appear to be rather static or fixed. We have seen that many RNA-binding proteins have a preferred location within the speckle compartment, even though they are normally in a continuous state of flux (Misteli and Spector, 1998). The same applies for many transcription factors and the way they interact with both promoters and the nuclear substructure. Intriguingly, many nuclear proteins appear to be stably associated with the nucleoskeleton. The most comprehensive study of this question has been performed on the DNA replication protein PCNA (proliferating cell nuclear antigen). PCNA is a major component of the replication factories and in living cells has been shown to be associated with sites that behave as though they are associated with a much larger nuclear structure (Sporbert *et al*., 2002).

1.5.1 The nucleoskeleton, nuclear matrix and lamina

We have seen that nuclei can be plastic, so that they appear to have a complex structure of ill-defined nuclear compartments, which generally represent centres of nuclear function. These centres have components that are inherently dynamic and in general are not stable entities. But while these compartments might be plastic, it is also interesting to note that the spatial organization of sites such as transcription factories and speckles is unaffected if nuclei are depleted of chromatin (Figure 1.3). This suggests that the compartments are spatially constrained and raises the possibility that a nuclear 'solid phase' exists to provide a platform upon which the active sites might assemble. A 'nucleoskeleton' has been proposed to fulfil this role.

As the molecular structure of a filament-based network equivalent to the cytoskeleton has not been unequivocally demonstrated in nuclei, the nature of the nucleoskeleton remains a matter for debate. Core filaments that could provide a form of continuity throughout nuclei have been described (Nickerson, 2001). These filaments have characteristic morphological features of the intermediate filaments seen throughout the cytoplasm. The best-characterized intermediate filaments within nuclei are undoubtedly the nuclear lamins, which form a mesh-like structure at the nuclear periphery, where they play important roles in maintaining nuclear shape and controlling chromatin organization. Recently, it has been suggested that the lamin proteins spread further into the nucleus as a veil of lamin filaments that pervades the nucleoplasm. Interestingly, the integrity of nuclear lamin filaments has been shown to be a major determinant of nuclear function. For example, using dominant negative mutant lamin proteins to disrupt lamin assembly, it has been demonstrated that intact lamin filaments are

NEOGENOMICS EX.1028.015

c01.fm Page 16 Wednesday, March 15, 2006 8:05 AM

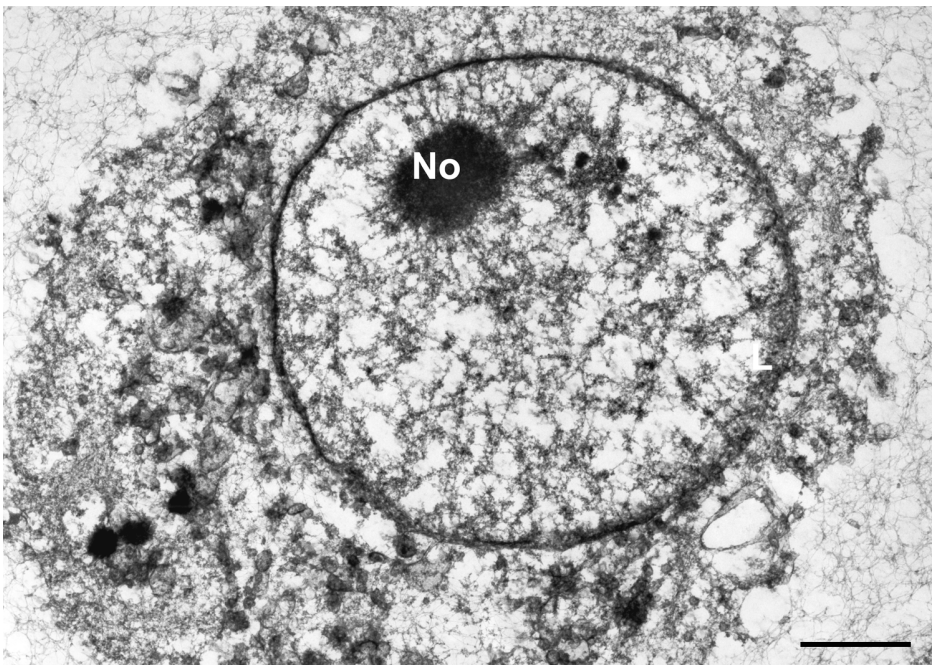

Figure 1.3 The nucleoskeleton in mammalian cells. To reveal the nucleoskeleton, HeLa cells encapsulated in agarose microbeads were permeabilized and chromatin cut with nucleases, so that ~90% DNA could be removed before preparing a resinless electron micrograph. In this ~250 nm thick section, note that a diffuse agarose mesh surrounds the densely stained cell. The spherical, central nucleus is demarcated by a dense nuclear lamina (L). It is important to recognize that nuclear architecture is preserved even though most chromatin – half of the nuclear mass – is removed. The residual nucleoskeleton can be visualized as a diffuse network of coated filaments that pervade the nuclear volume and provides a 'solid phase' upon which nuclear compartments such as nucleoli (No), replication and transcription factories and interchromatin granule clusters are assembled. The distribution of transcription and replication centres is preserved in preparations such as this, even though almost all chromatin has been removed. Bar = 2 μm. Reproduced from Hozak *et al*. (1994) with permission of the Company of Biologists

required for both DNA and RNA synthesis. The disruption of RNA synthesis (Spann *et al*., 2002) is especially compelling, as experiments have demonstrated that destroying the lamin filaments can ablate synthesis by RNA polymerase II in the nucleoplasm, while synthesis by other classes of polymerase is unaffected. This approach supports the existence of a network that is required for expression of protein coding genes but which is either related to or dependent on the nuclear lamin filaments for its function. With this in mind, it is worth reflecting on the many examples where transcription factors are recruited to the nuclear matrix prior to activating differential gene expression in cell lineages where expression is maintained (Zaidi *et al*., 2001).

NEOGENOMICS EX.1028.016

c01.fm Page 17 Wednesday, March 15, 2006 8:05 AM

The nuclear matrix (Berezney *et al.*, 1996) is classically described as an amorphous fibro-granular structure that can be isolated from nuclei by hypertonic treatment following nuclease digestion. The matrix typically contains many hundreds of different proteins, many of which have been studied in detail. Two of the best-characterized, SAF-A (scaffold attachment factor A) and ARBP (attachment region binding protein), were discovered because of their association with matrix-attached DNA elements. SAF-A (Romig *et al.*, 1992) turns out to be a major RNA binding protein (hnRNP-U) and ARBP (Stratling and Yu, 1999) a protein that binds methylated DNA (MeCP2). The behaviour of this protein is undoubtedly complex. SAF-A, for example, has been shown to bind p300, a major transcriptional co-activator, and in this way can recruit active genes to the nuclear matrix. ARBP, in contrast, interacts with methylated DNA in MAR elements and, through an intermediary that contains Sin3A protein, recruits a histone deacetylase to generate a silenced chromatin state (Stratling and Yu, 1999). Another protein, SAF-B (scaffold attachment factor B) specifically binds to S/MAR regions, interacts with RNA polymerase II (RNA pol II) and a subset of serine/arginine-rich RNA processing factors (SR proteins). It was proposed that these interactions allow SAF-B to provide a surface for the assenbly of the transcription apparatus (Nayler *et al.*, 1998). As with the vast majority of characterized matrix proteins, these are involved in different aspects of chromatin function. Extending this idea implies that the nucleoskeleton and nuclear matrix (which forms during salt extraction) are an expression of different processes that are performed within the interchromatin space. In this way, the complexity of the matrix simply reflects that fact that many of the proteins involved in DNA and RNA metabolism are engaged in multiple interactions within protein megacomplexes that, through direct or indirect associations (via a filament network), are able to form a stable structure in extracted cells. This implies that interaction and associations between these structures and chromatin will be the major determinants of nuclear structure in mammalian cells.

1.5.2 Other nuclear compartments

Many other discrete nuclear compartments have been described (Spector, 2001; Wang *et al.*, 2002). A prominent example is the coiled or Cajal bodies, which are generally spherical, measure ~1 μm across and in cross section appear as a network of coiled fibres (Matera, 1999; see also Chapter 3). Coiled bodies are found in both animals and plants, and contain many different proteins, including p80 coilin (the diagnostic marker), various transcription factors, as well as small ribonucleoproteins. While these bodies are not transcriptionally active, nascent snRNAs are found immediately next to them in mammalian cells. PML bodies are defined by their content of promyelocytic leukemia protein (Matera, 1999; see also Chapter 8). They are roughly similar in size to coiled bodies but are usually more abundant, so a mammalian cell in tissue culture may have one to three coiled bodies and 10–20 PML-bodies, and the PML bodies split up when cells experience the stresses associated with heat-shock and

NEOGENOMICS EX.1028.017

c01.fm Page 18 Wednesday, March 15, 2006 8:05 AM

inflammation. PML bodies have also been referred to as the ND10 compartment. One report suggests that they are transcriptionally active and another that their surface can provide a site of efficient viral transcription and replication in infected cells (Maul, 1998).

Other sites have been described that might be classified as 'transcription super-factories'. The OPT (Oct1/PTF/transcription) domain provides the best example (Pombo *et al*., 1998). Immunofluorescence shows that in HeLa cells the transcription factors Oct1 and PTF are found in one to three large domains (diameter ~ 1.3 μm) that appear during G_1 phase and disappear in S phase. Each OPT domain typically contains two or three factories where Br-UTP is incorporated into nascent transcripts, as well as RNA polymerase II, TBP and Sp1. Some chromosomes, notably chromosome 6, are found to be closely associated with the OPT domains. A variety of other structures, such as the perinucleolar compartment and gemini bodies, have been described (Spector, 2001). However, none of the extranucleolar domains discussed above yet have well-defined roles. They all seem to be dynamic structures. Most importantly, these domains are all defined by a local concentration of a particular marker. However, these markers are also found elsewhere in the nucleus, usually in many thousands of smaller sites. Therefore, the major domains discussed above may simply be uninteresting depots where components are stored. Thus, a component like p80 coilin may be stored in a coiled body, but function elsewhere. Note in this context that microinjecting antibodies directed against p80 coilin disrupts coiled bodies, without obvious detriment to the cell (Almeida *et al*., 1998).

1.6 Nuclear architecture – implications for viral infection

The preceding discussion gives an overview of nuclear architecture in higher eukaryotes and emphasizes how chromatin domains and specialized nuclear compartments play a vital role in defining programmes of gene expression. With these features in mind, we now need to ask how these aspects of nuclear architecture influence the behaviour of a viral genome as it enters the nucleus soon after infecting the host cell (for the sake of clarity, I will not discuss viruses that only occupy the cytoplasm). As a generalization, it is safe to assume that the major functions of a viral genome are performed predominantly by host cell components. This is not surprising, as most viral genomes are too small to code for each of the proteins that are needed to generate the very sophisticated machines required for DNA replication and RNA synthesis and processing. Following infection, the virus must enter the nucleus. Many different mechanisms for nuclear targeting have been described (see Chapter 3). Once inside the nucleus, the viral genome must engage the relevant host cell apparatus. In most cases, this early phase of the infection cycle requires that the viral genome accesses the nuclear compartment where gene expression occurs. Hence, viruses target transcription sites and the nucleoskeleton. In reality, nuclear structure ensures that this is almost inevitable. Hence, the location of transcription sites at the interface

NEOGENOMICS EX.1028.018

c01.fm Page 19 Wednesday, March 15, 2006 8:05 AM

between chromatin domains and the interchromatin channels ensures that mature mRNPs can pass efficiently to the cytoplasm just as the same organization ensures that viruses entering the nucleus will pass efficiently to transcription sites via the same channels. Some viruses might seek preferred nuclear target sites, such as PML bodies, but these too are established in the highly accessible interchromatin domains. Once the early phase of transcription is established, viruses commonly engage other mechanisms to ensure that expression of viral genes will proceed at the expense of host expression. In some instances host synthesis shuts off and the host chromatin architecture is destroyed. Clearly, different viruses use different mechanisms, which it is not appropriate to catalogue here. At a later stage of infection the focus of activities might switch to viral genome synthesis. Once again, those viruses with DNA in their infectious particles will subvert the host replication machinery to reproduce large quantities of the viral genome. During this process, some features of host cell replication are retained while others are discarded.

1.7 Conclusion

In higher eukaryotes, many layers of organization contribute to the complex processes that control chromatin function. We have seen how changes in chromatin structure modulate the access of transcription factors that activate pathways of gene expression. For genes in chromatin that are competent for gene expression, simple recognition motifs in DNA can interact with transcription factors that eventually recruit the transcription machinery so that RNA synthesis can begin. During this process, the constellation of factors assembled on promoters is likely to be in a continual state of flux until some dynamic equilibrium is established that allows transcription to proceed. This dynamic process explains some of the complexity of transcriptional regulation, as individual factors will usually interact weakly with DNA but operate cooperatively with other factors in the promoter complex. Many other features of genome architecture and nuclear structure can have profound effects on gene expression. This is clear from many experiments that demonstrate that different chromosomal locations have quite different expressional capabilities. *A priori*, we might assume that genes expressed from their natural chromosomal sites would operate with the required efficiency. It is now obvious that the combination of genetic factors determined by the organization of DNA sequence elements and different epigenetic factors ensure that the expression of any particular gene is a very complex process.

In higher eukaryotes, global features of nuclear organization also appear to influence gene expression. Over recent years a significant body of literature has described the compartmentalization of different nuclear functions within dedicated sites. In mammalian cells the density of transcription sites is far lower than expected from estimates of active transcripts and transcription unit complexity, implying that multiple transcription units are active in each site. Morphological analyses, together

NEOGENOMICS EX.1028.019

c01.fm Page 20 Wednesday, March 15, 2006 8:05 AM

with immunostaining, indicate that these sites have a zonal organization, with different regions performing specialized roles. Furthermore, as the transcript-containing regions occupy only 0.2% of a HeLa nucleus (Iborra *et al.*, 1996), this arrangement is inconsistent with the view that transcripts will be uniformly dispersed throughout euchromatin (occupying ~10% of the nucleus) bound to tracking RNA polymerases. Instead it appears that many chromatin domains may be served by a single, active transcription compartment.

Although many ideas have been proposed to explain the observed arrangement of active sites in mammalian cells, it is still unclear whether this level of organization is an essential component of nuclear structure. Nevertheless, it does at least appear reasonable to argue that by concentrating components required to perform coupled functions in a limited number of sites, it will be possible to execute these functions with optimal efficiency. It is clear that active chromatin domains must maintain their active configuration under circumstances where the majority of chromatin is inert. Although different mechanisms will tend to stabilize these forms, it is important that they are both sufficiently stable to maintain expression and sufficiently unstable to allow reprogramming. Interestingly, proteins capable of remodelling chromatin have been shown to form an integral part of the transcription machinery. Restricting the spatial distribution of the complex in this way will tend to preserve existing chromatin states.

How, then, does this impact on the events that follow virus infection? For the point of view of a virus, the critical requirement is to ensure that the viral genome has an efficient mechanism to engage the active nuclear domains and evade the inert chromatin compartments – where they too might acquire a condensed chromatin status. I have described how nuclear architecture ensures that the active sites are linked to the cytoplasm through a network of interchromatin channels. It is proposed that this interchromatin compartment will represent a 'route of least resistance', so that mRNP complexes passing from transcription sites into the channel network will move quickly to the nuclear periphery. Inevitably, the same route will be used in the reverse direction – from cytoplasm to active sites – by the substrates for RNA synthesis and other dynamic components of the active centres. By engaging this very pathway, a virus will ensure that the active compartments are encountered with optimal efficiency.

References

Abney, J. R., Cutler, M. L., Axelrod, D. and Scalettar, B. A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. *J Cell Biol* **137**, 1459–1468.

Alvarez, J. D., Yasui, D. H., Niida, H., Joh, T., Loh, D. Y. and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes Dev* **14**, 521–535.

Almeida, F., Saffrich, R., Ansorge, W. and Carmo-Fonseca, M. (1998). Microinjection of anti-coilin antibodies affects the structure of coiled bodies. *J Cell Biol* **14**, 2899–2912.

NEOGENOMICS EX.1028.020

c01.fm Page 21 Wednesday, March 15, 2006 8:05 AM

Becker, P. B. (2002). Nucleosome sliding: facts and fiction. *EMBO J* **21**, 4749–4753.

Bell, A. C. and Felsenfeld, G. (1999). Stopped at the border: boundaries and insulators. *Curr Opin Genet Dev* **9**, 191–198.

Bell, A. C., West, A. G. and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396.

Bender, M. A., Bulger, M., Close, J. and Groudine, M. (2000). β-globin gene switching and Dnase I sensitivity of the endogenous β-globin locus in mice do not require the locus control region. *Mol Cell* **5**, 387–393.

Berezney, R., Dudey, D. D. and Huberman, J. A. (2000). *Chromosoma* **108**, 471–484.

Berezney, R., Mortillaro, M. J., Ma, H., Wei, X. and Samarabandu, J. (1996). The nuclear matrix: a structural milieu for genomic function. *Int Rev Cytol* **162A**, 1–65.

Blackwood, E. M. and Kadonaga, J. T. (1998). Going the distance: a current view of enhancer action. *Science* **281**, 60–63.

Bode, J., Benham, C., Knopp, A. and Mielke, C. (2000). Transcriptional augmentation: modulation of gene expression by scaffold/matrix attached regions (S/MAR elements). *Crit Rev Eukaryotic Gene Expr* **10**, 73–90.

Breiling, A., Bonte, E., Ferrari, S., Becker, P. B. and Paro, R. (1999). The Drosophila polycomb protein interacts with nucleosomal core particles *in vitro* via its repression domain. *Mol Cell Biol* **19**, 8451–8460.

Bridger, J. M., Boyle, S., Kill, I. R. and Bickmore, W. A. (2000). Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. *Curr Biol* **10**, 149–152.

Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M. and Fisher, A. G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **9**, 845–854.

Bulger, M., Sawado, T., Schubeler, D. and Groudine, M. (2002). ChIPs of the β-globin locus: unraveling gene regulation within an active domain. *Curr Opin Genet Dev* **12**, 170–177.

Cavalli, G. and Paro, R. (1999). Epigenetic inheritance of active chromatin after removal of main transactivator. *Science* **286**, 955–958.

Cook, P. R. (1999). The organization of replication and transcription. *Science* **284**, 1790–1795.

Cremer, T. and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**, 292–301.

Dundr, M., Hoffmann-Rohrer, U., Hu, Q. Y., Grummt, I., Rothblum, L. I., Phair, R. D. and Misteli, T. (2002). A kinetic framework for a mammalian RNA polymerase *in vivo*. *Science* **298**, 1623–1626.

Elefant, F., Cooke, N. E. and Liebhaber, S. K. (2000). Targeted recruitment of histone acetyltransferase activity to a locus control region. *J Biol Chem* **275**, 13827–13834.

Engel, J. D. and Tanimoto, K. (2000). Looping, linking and chromatin activity: new insights into β-globin locus regulation. *Cell* **100**, 499–502.

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

Farrell, C. M., West, A. G. and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human β-globin loci. *Mol Cell Biol* **22**, 3820–3831.

Femino, A. M., Fay, F. S., Fogarty, K. and Singer, R. H. (1998). Visualization of single RNA transcripts *in situ*. *Science 280*, 585–590.

Ferguson-Smith, A. C. and Surani, M. A. (2001). Imprinting and the epigenetic asymmetry between parental genomes. *Science* **293**, 1086–1089.

NEOGENOMICS EX.1028.021

c01.fm Page 22 Wednesday, March 15, 2006 8:05 AM

Festenstein, R., Sharghi-Namini, S., Fox, M., Roderick, K., Tolaini, M., Norton, T., Saveliev, A., Kioussis, D. and Singh, P. (1999). Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal context-dependent manner. *Nat Genet* **23**, 457–461.

Gasser, S. M. (2002). Visualizing chromatin dynamics in interphase nuclei. *Science* **296**, 1412–1416.

Gilbert, D. M. (2002). Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* **14**, 377–383.

Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R. and Fraser, P. (2000). Intergenic transcription and developmental remodeling of chromatin subdomains in the human β-globin locus. *Mol Cell* **5**, 377–386.

Hendzel, M. J., Kruhlak, M. J., MacLean, N. A. B., Boisvert, F. M., Lever, M. A. and Bazett-Jones, D. P. (2001). Compartmentalization of regulatory proteins in the cell nucleus. *J Steroid Biochem* **76**, 9–21.

Hozak, P., Jackson, D. A. and Cook, P. R. (1994). Replication factories and nuclear bodies: the ultrastructural characterization of replication sites during the cell cycle. *J Cell Sci* **107**, 2191–2202.

Iborra, F. J., Pombo, A., Jackson, D. A. and Cook, P. R. (1996). Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci* **109**, 1427–1436.

Jackson, D. A. and Pombo, A. (1998). Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S-phase in human cells. *J Cell Biol* **140**, 1285–1295.

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

Jackson, D. A., Iborra, F. J., Manders, E. M. M. and Cook, P. R. (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. *Mol Biol Cell* **9**, 1523–1536.

Jackson, D. A., Pombo, A. and Iborra, F. (2000). The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells. *FASEB J* **14**, 242–254.

Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* **293**, 1074–1080.

Kosak, S. T., Skok, J. A., Medina, K. L., Riblet, R., Le Beau, M. M., Fisher, A. G. and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296**, 158–162.

Lachner, M., O'Carroll, N., Rea, S., Mechtler, K. and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120.

Litt, M. D., Simpson, M., Recillas-Targa, F., Prioleau, M. N. and Felsenfeld, G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J* **20**, 2224–2235.

Lundgren, M., Chow, C. M., Sabbattini, P., Georgiou, A., Minaee, S. and Dillon, N. (2000). Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell* **103**, 733–743.

Matera, A. G. (1999). Nuclear bodies: multifaceted subdomains of the interchromatin space. *Trends Cell Biol* **9**, 302–309.

Maul, G. G. (1998). Nuclear domain 10, the site of DNA virus transcription and replication. *Bioessays* **20**, 660–667.

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

NEOGENOMICS EX.1028.022

c01.fm Page 23 Wednesday, March 15, 2006 8:05 AM

McNally, J. G., Muller, W. G., Walker, D., Wolford, R. and Hager, G. L. (2000). The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287**, 1262–1265.

Misteli, T. and Spector, D. L. (1998). The cellular organization of gene expression. *Curr Opin Cell Biol* **10**, 323–331.

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

Muyrers-Chen, I. and Paro, R. (2001). Epigenetics: unforeseen regulators in cancer. *Biochim Biophys Acta* **1552**, 15–26.

Nayler, O., Stratling, W., Bourquin, J. P., Stagljar, I., Lindemann, L., Jasper, H., Hartmann, A. M., Fackelmayer, F. O., Ullrich, A. and Stamm, S. (1998). SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements. *Nucleic Acids Res* **26**, 3542–3549.

Nickerson, J. A. (2001). Experimental observations of a nuclear matrix. *J Cell Sci* **114**, 463–474.

Nielsen, A. L., Oulad-Abdelghani, M., Ortiz, J. A., Remboutsika, E., Chambon, P. and Losson, R. (2001). Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol Cell* **7**, 729–739.

Pombo, A., Cuello, P., Schul, W., Yoon, J.-B., Roeder, R. G., Cook, P. R. and Murphy, S. (1998). Regional and temporal specialization in the nucleus: a transcriptionally-active nuclear domain rich in PTF, Oct1 and PIKA antigens associates with specific chromosomes early in the cell cycle. *EMBO J* **17**, 1768–1778.

Pombo, A., Jackson, D. A., Hollinshead, M., Wang, Z., Roeder, R. G. and Cook, P. R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J* **18**, 2241–2253.

Romig, H., Fackelmayer, F. O., Renz, A., Ramsperger, U. and Richter, A. (1992). Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix scaffold attachment DNA elements. *EMBO J* **11**, 3431–3440.

Sacco-Bubulya, P. and Spector, D. L. (2002). Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. *J Cell Biol* **156**, 425–436.

Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B. M. and Zink, D. (1999). Nuclear organization of mammalian genomes: polar chromosome territories build up functionally distinct higher order compartments. *J Cell Biol* **146**, 1211–1226.

Smith, K. P., Moen, P. T. Jr, Wynder, K. L., Coleman, J. R. and Lawrence, J. B. (1999). Processing of endogenous pre-mRNAs in association with SC-35 domains is gene specific. *J Cell Biol* **144**, 617–629.

Spann, T. E., Goldman, A. E., Wang, C., Huang, S. and Goldman, R. D. (2002). Alteration of nuclear lamin organization inhibis RNA polymerase II-dependent transcription. *J Cell Biol* **156**, 603–608.

Spector, D. L. (2001). Nuclear domains. *J Cell Sci* **114**, 2891–2893.

Sporbert, A., Gahl, A., Ankerhold, R., Leonhardt, H. and Cardoso, M. C. (2002). DNA polymerase clamp shows little turnover at established replication sites but sequential *de novo* assembly at adjacent origin clusters. *Mol Cell* **10**, 1355–1365.

Stratling, W. H. and Yu, F. (1999). Origin and roles of nuclear matrix proteins. Specific functions of the MAR-binding protein MeCP2/ARBP. *Crit Rev Eukaryot Gene Expr* **9**, 311–318.

Taddei, A., Roche, D., Sibarita, J.-B., Turner, B. M. and Almouzni, G. (1999). Duplication and maintenance of heterochromatin domains. *J Cell Biol* **147**, 1153–1166.

Tsukamoto, T., Hashiguchi, N., Janicki, S. M., Tumbar, T., Belmont, A. S. and Spector, D. L. (2000). Visualization of gene activity in living cells. *Nat Cell Biol* **2**, 871–878.

NEOGENOMICS EX.1028.023

c01.fm Page 24 Wednesday, March 15, 2006 8:05 AM

Udvardy, A. (1999). Dividing the empire: boundary chromatin elements delimit the territory of enhancers. *EMBO J* **18**, 1–8.

van Driel, R., Wansink, D. G., van Steensel, B., Grande, M. A., Schul, W. and de Jong, L. (1996). Nuclear domains and the nuclear matrix. *Int Rev Cytol* **162A**, 151–189.

Wang, I. F., Reddy, N. M. and Shen, C. K. J. (2002). Higher order arrangement of the eukaryotic nuclear bodies. *Proc Natl Acad Sci USA* **99**, 13583–13588.

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

Wallace, H., Ansell, R., Clark, J. and McWhir, J. (2000). Pre-selection of integration sites imparts repeatable transgene expression. *Nucleic Acids Res* **28**, 1455–1464.

Workman, J. L. and Kingston, R. E. (1998). Alterations of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* **67**, 545–579.

Wilson, C., Bellen, H. J. and Gehring, W. J. (1990). Position effects on eukaryotic gene expression. *Annu Rev Cell Biol* **6**, 679–714.

Zahn-Zabal, M., Kobr, M., Girod, P. A., Imhof, M., Chatellard, P., de Jesus, M., Wurm, F. and Mermod, N. (2001). Development of stable cell lines for production or regulated expression using matrix attachment regions. *J Biotechnol* **87**, 29–42.

Zaidi, S. K., Javed, A., Choi, J. Y., van Wijnen, A. J., Stein, J. L., Lian, J. B. and Stein, G. S. (2001). A specific targeting signal directs Runx2/Cbfa1 to subnuclear domains and contributes to transactivation of the osteocalcin gene. *J Cell Sci* **114**, 3093–3102.

Zhao, K., Hart, C. M. and Laemmli, U. K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* **81**, 879–889.

NEOGENOMICS EX.1028.024